Peer Review File

Manuscript Title: Adeno-associated virus 2 infection in children with non-A-E hepatitis

Reviewer Comments & Author Rebuttals

Reviewer Reports on the Initial Version:

Referee #1 (Remarks to the Author):

Ho et al. report the results of an observational study on 9 cases of unexplained hepatitis (nonA-E, transaminases >500IU, <10yrs) and 58 controls.
An increase in cases of non-A-E hepatitis in children under the age of 10 has recently been detected in several countries. With approximately 5-6% of cases requiring transplantation approximately 2% mortality this is a serious condition, yet the etiology remains currently unknown. There is a high need to identify causes of pediatric hepatitis in order to provide better diagnostics and develop therapeutic approaches. Human adenovirus infection (hAdV) and SARS-CoV-2 infection has been discussed as possible causes, yet the evidence has not been convincing or conclusive.

Here, Ho et al. present evidence that human adeno-associated virus 2 (AAV2) might be associated with unexplained pediatric hepatitis.
The key finding of the study is a highly significant difference in positivity rates for AAV2 in the pediatric hepatitis cohort (9/9 in blood, and 4/4 cases in liver) compared to 0/13 sera/plasma of age-matched healthy controls, 0/12 children with adenovirus (HAdV) infection and normal liver function and 0/33 children admitted to hospital with hepatitis of other aetiology. In addition, the class II HLA-DRB1*04:01 allele was identified in 17/20 cases (85%), compared with a background frequency of 10/64 (15.6%) in Scottish blood donors.
This might indicate an involvement of AAV2 in pediatric hepatitis, potentially through immune activation in susceptible donors. The authors argue that AAV2 might be a causative agent of pediatric hepatitis or a sensitive biomarker of previous hAdV infection, which cannot be ruled out as a cause of hepatitis.

Overall, these are interesting and potentially important results that might contribute to reveal the etiology of unexplained pediatric hepatitis.
The main limitation of the study is its limited sample size and the retrospective design and the heterogeneity of the case- and control cohorts.

Specific comments
The authors need to better define inclusion and exclusion criteria for the control group

One of the main limitations is sample size. It would be preferable to increase sample size in the case cohort. Given that 20 cases were available for HLA-genotyping, it may be possible to increase the sample size in the case cohort also for viral analysis (AAV2).

Was the liver damage as assessed in Fig, 1e-j scored and compared to other forms of (viral) hepatitis?

Minor
Overall the quality of the figures could be improved (higher quality data plots, alignment of panels, adding missing scale bars in Fig. 1e-j, legible font size).

Referee #2 (Remarks to the Author):

A study that examines for viral infections in 9 children in Scotland who developed otherwise unexplained hepatitis during the March and April of 2022. The context is that there was an unusual number of cases of hepatitis in children that were non-Hepatitis A-E virus; other cases have been reported elsewhere, but were not investigated. As controls, relatively similar samples were collected from three groups of: nominally healthy children; human adenovirus infected but no signs of liver disease; and other children with raised transamidases. While reported as a cluster, it is not clear that this is an outbreak (e.g. of a connected transmission chain) or some other event. The other epidemiological and demographic data is a little hard to see, but may have been reported elsewhere(?) Unusual circumstances relate to the prior lock-docks associated with SARS CoV-2 and recent reduction in restrictions - and an increase in human adenovirus infections in diagnostic samples around the time of the cluster was seen. Some association with SASR CoV-2 is not easy to exclude, but no direct association among the affected children was not obvious.

The approach taken was to conduct deep metagenomic sequencing of the RNA and DNA from liver samples (in some cases) and from blood samples in all. The results show the presence of varying levels of AAV2 DNA in the samples of the affected children, at a higher proportion and level than seen in the control samples. The various adenoviral and herpes viral DNAs were found in some cases. AAV2 is a common infection of people worlwide, with most having been infected early in life, and DNA is found to persist in tissues for many years, or even for life, with some integrated into the genome. The AAV2 infection has not been known to be associated with any adverse effects, and the virus rarely replicates in the absence of a helper adenovirus or herpesvirus. When delivered in very large amounts for gene therapy studies AAV capsids have been associated with liver disfunction and disease. Some adenovirus and HHV6 DNA was also reported, but at lower levels than the AAV2 DNA, or sometimes are absent. The suggestion made in the end is that the hepatitis is associated with an increase in human adenovirus infections, with a possible contributing or exacerbating role for the AAV2 coinfection. Phylogenetic analysis of the AAV2 sequences suggests that those fall into 2 clades which suggests possible epidemiological linkages between those viruses, and some mutations were present - there is no way to assess the significance of those mutations - there may be none.

Overall this study is quite well done, and presents a case for a role of adenovirus and AAV2 in the liver disease seen, by an unknown mechanism. It seems unlikely that AAV2 has suddenly emerged as a new pathogen in humans, so the title (and some of the text) is going to be misinterpreted by interested clinicians and others who do not read the study carefully or are not knowledgeable about the nuances of AAV and Ad biology. The main issues are around the presentation and not creating confusion in the future, particularly in the clinical fields - don't want everyone suddenly finding AAV2 DNA (which is quite ubiquitous in the human population), and diagnosing that as the cause of the hepatitis or other diseases.

Some specific comments:
1) The title might be more accurate: "A cluster of hepatitis cases in Scottish children possibly due to adenovirus and adeno-associated virus infection"
2) Is the model that this is a hit-and-run infection by one or more virus that progressed to more severe hepatitis after those viruses were largely cleared? If so, what is the pathological mechanisms causing the tissue damage, and can that be elaborated more clearly, including the association with HLA type explained?
3) The Conclusion/final statements section is quite long and a little hard to follow. The key take-away seems ot be saved to the last paragraph, and this could be used as the lead in to this section. In part the difficulty here is due to the uncertain nature of the disease and the associations that need more study. This could be tightened up quite a bit and made more focused on the certainties and uncertainties of the situation. The division into logical subsection with heading may help to make this clearer. A clear statement of what data or additional studies are needed to confirm or disprove the hypotheses about the cause of these type of heptitis cases would also be helpful.

Referee #3 (Remarks to the Author):

Ho et al. provide a spectacular observation of high levels of AAV2 in plasma and liver tissue of pediatric patients with acute severe hepatitis thus adding a new layer of information to this mystery. While the emergence of AAV2 is certainly novel, the study fails to make a causal link, merely making an interesting association. In its present form, the study lacks the rigorous scientific workup of this finding including proper hepatologic workup and immunologic studies to make any conclusion if we are dealing with an infectious hepatitis or an immune-mediated liver disease. I worry that in the present for the data will cause further confusion.

1. General comment: The current wave of acute severe hepatitis of unknown etiology is a pressing issue in the (pediatric) hepatology community due to high rate of liver transplantation (5%) in these kids. The majority of case studies thus far found an association with adenovirus, particularly the F type HAdV41. This however is puzzling since adenovirus induced hepatitis is rare and usually restricted to severely immunocompromised individuals. Furthermore, on histology adenovirus hepatitis is hallmarked by nuclear viral inclusion particles, something that has not been detected in the current outbreak. Furthermore, the current cases occur in otherwise healthy children and it was thus far difficult to reconcile adenovirus infection with fulminant hepatitis leading to liver failure. As it currently stands there are four major unresolved questions: i.) is there a novel adenovirus strain that might cause acute liver failure, ii.) is another pathogen involved not yet detected, iii.) what is the role of SARS-Cov2, which has been described to cause hepatitis and persistence in organ systems such as the intestine and iv.) do these patients show an autoimmune-like (AIH-like) liver failure or is it truly an infectious hepatitis. The study now adds valuable information, identifying AA2 as a potential new infectious agent, however due to the superficial workup of this finding, they do not

provide meaningful further evidence towards solving the mystery.

Specific criticisms:

2. AAV2 requires a helper virus to replicate as the authors correctly point out. As such the authors find HAdV and HHV6 in 6/9 and 3/9 plasma samples, 3/4 and 2/4 liver biopsies. How do the authors explain this? Does it not seem of that at least one patient must have had 3 viruses detectable?

3. Clinical characteristics are minimal. The authors should add a detailed table with clinical parameters (mean ALTs, ASTs, IgG, histology, fibrosis, autoantibody titers, medication, BMI)

4. The authors write that autoimmune markers were negative. More details here should be provided. Unlike adults, for pediatric AIH, ANA titers of 1:20, SMA is 1:10-1:20 and LKM-1 1:10 is considered the threshold (Sebode et al. Front Immunol 2018). Furthermore, AIH 2 is often presented with normal IgG. How were autoantibodies detected? IFT or ELISA? What were the cut-offs? What was the ANA pattern? Were Hep2 cells used? This is critical here to rule out autoimmune hepatitis.

5. Along those lines, while liver histology samples are shown, much information is missing. HE stains clearly show interface hepatitis. What was the mHAI? Were the samples scored by liver pathologists, the author must have that data. What was the grading. Follwing the simplified AIH score at least the patient with ANA 1:80 would score a probable AIH if liver histology was typical (Hennes et al. 2008) and thus there need to be much more information to follow the authors argument that autoimmunity was excluded.

6. As the authors nicely discuss, the findings of AAV2 detection and the DRB1*04 HLA allele point to a CD4 T cell mediated pathology. However this should be analyzed. Did the authors stain for CD4 T cells? Are there data of a broad immune phenotyping? TCR phenotyping is also possible from FFPE tissue or could at least be attempted.

7. Can the authors provide serology for AAV2? Are these cases acute infection or reactivation? Of note, up to 80% of adults can have seroconversion and La Bella et al. (Gut 2020) found AAV2 in 20% of individuals undergoing liver biopsy. Are the authors suggesting the cases are first time infected, which would align with the age group?

8. Data in Fig. 4 is really important, indicating that there is not a new strain of AAV2 but matching known strains. This seems somehow underemphasized in the manuscript and should be expanded.

More general comments:

9. While AAV2 is know to have liver tropism, it is not commonly known to cause hepatitis. It is unclear why the current outbreak is linked to the presence of AAV2. Simple experiments such as infection of primary human hepatocytes or coinfections with adenovirus and AAV2 are lacking which would help to identify if liver damage would really be due to AAV2. Such experiments could show whether AAV2 infects hepatocytes and impacts hepatocyte viability / cellular metabolism (as opposed to trigger an immune response causing an “autoimmune-like” hepatitis).

10. Almost all cases were actively or previously infected with SARS-CoV-2. Furthermore, both the US and UK have the highest reported cases of unknown acute hepatitis and both countries had high numbers of all previous SARS-CoV-2 variants and have a high rate of seroprevalence compared to other places such as continental Europe. This at least hints at a potential role for SARS-CoV-2 in these cases. While discussed, both studies take little efforts to rule out Covid-19 as a potential explanation. Scientific questions that need to be addressed are SARS-CoV-2 persistence in stool, T-cell receptor skewing in the liver and systemic or hepatic IFN-γ upregulation. These analyses could provide evidence of a SARS-CoV-2 superantigen mechanism in an adenovirus-41F- or AAV2 sensitised host. While I appreciate that Covid must not be the explanation per se, of all the potential pathogens, Covid is the new kid on the block and both studies do not provide sufficient evidence by a mile to rule out a contribution of Covid in this epidemic of cases.

11. The evidence provided that this is indeed a viral hepatitis is limited. The association of high levels of AAV2 in affected individuals both in liver and blood is striking. However, the study did not find evidence of virus proteins, signs of infected hepatocytes on histology, or even provide principal evidence of AAV2 directed hepatotoxicity from model systems. This strongly suggests an immune-mediated pathology as discussed. To investigate this further, a proper immunologic workup, including TCR sequencing, immunofluorescence, immune phenotyping etc. should be done. This is of highest clinical relevance as the result for clinical management will be either immunosuppression or directed novel antivirals.

12. After exclusion of known etiologies, autoimmune hepatitis is often the cause of acute liver failure in both adults and children. AIH can be triggered by drugs or infections (molecular mimicry) and it has been speculated that these acute severe hepatitis cases are due to AIH-like inflammation triggered by e.g. adenovirus. Both studies undertake little efforts to rule out or investigate virus-triggered AIH. It would be paramount to provide histology of the liver samples, scoring by a liver pathologist, autoantibody titers, quantitative immunoglobulins etc. I assume that the authors have those datasets and they should be provided. In fact, the genetic association with HLA-DRB1*04:01 is strong for AIH and is even part of the international AIH study groups (IAIHG) diagnostic scoring, further suggesting these kids could have an AIH-like hepatitis.

13. One potential clue to understand if AAV2-induced hepatitis is occurring vs. an autoimmune hepatitis (potentially AA2-triggered) could result from analyzing complement factors. Complement does not play an important role in AIH (Biewenga et al. 2020, PMID 32234403), however, complement activation and deposition is an essential part of the immune response to AAVs as seen in rodents (Zaiss et al. 2008, PMID 32234403) and in the clinical trials that used AAVs as vectors that reported severe side effects (ClinicalTrials.gov identifiers: NCT03362502 and NCT03368742, reviewed by Muhuri et al. 2021, PMID 33393506). Could the authors stain complement in liver tissue and measure complement factors in blood?

Minor comment:
14. Is there additional data on why the cases developed severe hepatitis vs. controls? More C-sections? Breastfeeding? Environmental factors that could explain higher incidence such as daycare exposure, more or less quarantine?

Referee #4 (Remarks to the Author):

The paper by Ho et al., reports the detection of AAV2 in cases of acute hepatitis of unknown etiology (also referred to as non-A-E hepatitis). The authors present sequencing and molecular diagnostics to support the presence of AAV2 and helper viruses (HAdV and HHV6) in cases but rarely in several different control groups. The authors also show an increased frequency of HLA-DRB1*04:01 in cases compared to controls, suggestive of a role of human genetics in predisposing cases to hepatitis.

The paper presents a compelling result regarding the potential relevance of AAV2 and HLA alleles in the recent outbreak of non-A-E pediatric hepatitis, despite the small sample size. While this result is novel and timely, the presentation of the data is confusing and, at times, repetitive, which makes obscures the contribution of this work. Additionally, very little context is given for the motivation of this study and how to interpret the results until the discussion.

I will focus my technical comments on the molecular methods used (PCR and sequencing), which are in general appropriate and analyzed using reasonable metrics. However, the presentation makes interpretation challenging. The genome sequencing sub-section is very long compared to all other sections and would benefit from both shortening and restructuring to more clearly define the purpose of each paragraph and the analyses being presented within each, including more clearly separating sequencing and PCR results and removing redundancy / clarifying where the same result is being presented in a different way.

Below I give several specific comments and questions regarding the data presented in this section;
- Can the authors clarify if all numbers of viral reads reported in text are from targeted enrichment libraries and from DNA or RNA extracts? It seems that RNA in general was much less sensitive than DNA sequencing for these pathogens. If this is not used in any analysis it may be better presented only in supplementary material.
- Are the full AAV2 genomes reported all from plasma/serum or other compartments? Also, in the mutation analyses why are only an arbitrarily defined group of closely related sequences compared for the presence of these mutations? I think it is important to also know if these mutations are found in other clades
- Line 177 – I am surprised that so many different pathogens are detected in such a small cohort of samples. It is noted that these are detected by sequencing but I couldn't see where this data was shown nor details of the sample types, which would have been interesting.
- Are these 9 genomes of the sub-lineage AAVv66 or they are classified as a different sub-lineage but share these mutations with that sub-lineage? If the latter, is there any evidence of recombination that could explain this?
- Is there a reason why control group 3 was not sequenced? Their inclusion for qPCR but not sequencing feels a little sudden; it would be helpful to explain why the addition of this control group via PCR only is beneficial to the analysis.
- It was also striking from Figure 3, though not mentioned in the text, the one HAdV+ control with

many coinfections. Is there anything clinically that would be consistent with this observation?
- Please specify sample type for control samples. It seems it is a mix of serum and plasma but I was unsure which was which.
- Should qPCR actually be referred to as RT-qPCR as a reverse transcription was performed?
- In the methods, please add details of how RPM is calculated.

Regarding the figures for this section (Figures 2 and 3), I found the representation of control data in heatmaps to be hard to read and potentially misleading for a couple of reasons. First, it implies there is something relevant about the ordering of individuals relative to cases which is not the case. Secondly, low read counts (especially on Figure 3) are hard to read. I think it would be more clear to show controls as aggregate data via a barplot, scatterplot or similar. Thirdly, presenting this in a binary format will allow to more directly compare the statistics between sequencing and qPCR.

Below are several comments on clarity of presentation and structure relating to the genome sequencing section.
- It seems that the first paragraph is intended to be a general overview of the sequencing and reporting just information from plasma. This was clear but then the PCR liver result for HHV6 is mentioned at the end which would be better placed elsewhere
- The next two paragraphs are diving more deeply into the genomics results on AAV2 and then HAdV and HHV6. This again makes sense but it would be helpful to have topic sentences on these paragraph to set-out this goal so the reader can more easily follow. Additionally, though the AAV2 paragraph focuses exclusively on genomic data, the HAdV/HHV6 paragraph starts to bring in PCR data. This is inconsistent and hard to follow what the main result being presented is.
- In some cases the same information appears to be presented multiple times but with subtle differences. This is particularly true of the PCR results, which is challenging to follow. E.g. lines 64-65 vs lines 114-115 lists routine blood tests run but HHV 6/7 is only in the latter list. As similar statement is again made on line 172 but it is unclear if this is a separate round of PCR testing for these pathogens or the same as the initial clinical diagnostics. Relatedly, are the PCR results that are presented alongside genome sequencing data (e.g. see lines 141, 163, 173) different to those presented as part of the “real-time PCR” sub-section? If not this should be made clear. If so then these should be presented just once and then clearly referred to elsewhere. Detection of AAV2 by sequencing in 9/9 plasma is also mentioned on lines 137 and 144.
- As the case-control comparison is one of the major analyses of the study, I suggest giving this a separate sub-heading. Initially presenting detection in the controls and then the statistical results would improve readability.

I found the uses of reads/million for the cases vs control comparison surprising and potentially problematic. In order for this to be robust I believe the authors need to more clearly demonstrate that their sequencing results are truly quantitative. In particular, that there are – as I understand it – a mix of serum and plasma samples here may influence the amount of host background. Additionally, I am inferring that sample collection and storage was not the same across case and control cohorts given that these are from distinct studies and sites. All of these variables can influence the relative amounts of human vs viral reads as well as the success of hybrid capture, which further alters this relationship. I strongly suggest that this be done as a qualitative analysis (just presence/absence of the virus) in order to be more robust to such variation.

Although in line 181 it is noted that case-control analysis was also done for real-time PCR results this analysis is only briefly referred via citation of the figures 2c-e and this data is not explored further. Given that all control groups are included here it seemed surprising that the significant difference here was not mentioned. I believe this is a more reasonable way to look at differences in load than the sequencing data. However, if copy number is available for qPCR data (which it seems from the written results that it may be) then it would be preferable to present this than CT values, which are not comparable across runs.

Overall, the paper appropriately credits previous work. However, in line with my comments below about clarity, the introduction, in particular, would benefit from more citations in particular around what is known about these viruses and the motivations for the study. One place where a citation is lacking is on line 263 where the statement about T-helper responses should have a citation. Potentially those given on lines 327-328 could be moved here as this section also feels out of place.

The presentation of the manuscript feels overall heavily geared towards a clinical audience and the manuscript overall would benefit from more clear framing of the key questions and differentiation between what was known of these samples prior to this study vs what is newly reported here.
- In the first section (Hepatitis outbreak in Scottish children) it would be helpful to include more background that would help the reader understand the relevance of the viruses being mentioned and also the central hypothesis the paper is looking to resolve. In particular, it is not clear to non-experts why if the majority of cases were already found to have HAdV this is insufficient as an explanation for these cases. Similarly, HHV6A and 6B incidence over 2021-2022 are mentioned in the final sentence but it does not become clear why this is relevant to the current cases until the results of this study later on.
- On a related note, the first mention of AAV2 as needing a helper virus such as HAdV or HHV6 is on line 179 when the case control analysis is introduced. This is something that would be better brought up earlier to explain why this analysis is being done
- Presentation of the control datasets is not consistent and there is variability in whether they are referred to as groups 1-3 or by clinical traits. When clinical traits are used, these too can be variable (e.g reference to transaminase levels, liver function, ALT, or LFT). Overall, my suggestion is to ascribe simple but descriptive identifiers to these 3 groups when first presented and use them consistently throughout in both text and figures. This variability makes it further very hard to trace other sub-groupings (e.g. the 12 children in clinical care on line 191 and 197).
- I wonder if presenting the clinical presentation before the control cohorts would flow more naturally. With the current format comparisons between cases & controls in age are presented before the cases have been described.
- In the conclusion, the relevance of the paragraph lines 302-308 was not clear. Though the clinical descriptions are helpful context I feel this would be better placed earlier in the paper where background on why HAdV and HHV6 are unlikely to be sufficient explanations for the observed hepatitis
- The limitations paragraph seems to also include additional interpretation of results. I suggest more clearly separating these two things. Specifically, everything from line 329 onwards should probably be moved sooner as they form part of the paper's hypothesis not its limitations. Which should be discussed separately.

Below are several smaller technical comments that would also aid clarity of presentation;
A specific definition of non-A-E hepatitis / hepatitis of unknown etiology would be very helpful. This could be as simple as a parentheses in the sentence that ends on line 65.

I found it hard to understand when the authors use HHV6 vs HHV6B and HAdV vs HAdV41 or Adenovirus species C and F. Could this be clarified?

Is the case shown in figure 1e-j representative? Can other images be shown in the supplement?

The (n=12) on line 94 is unnecessary as since the number was already written in the sentence

Line 145, it would be better to say AAV2 genomes were assembled rather than detected

Line 147, it would be better to say seven distinct sequences rather than lineages. Lineages implies more clear separation and structure than is observed here where all samples are highly genetically similar though not identical

What does "(TE)" mean on line 181?

Line 197 – I was unsure where the 2/12 with detectable HHV6B reads was already described above.

Line 252 – should "patients" be changed to "cases" to make it clear this does not include other unwell children from control groups?

Line 278 – it would be good to clarify that this was also already shown and is not specifically a result of this paper

Line 287-288 & 324; it was not apparent to me why in each instance HHV6 was not likely based on other information the authors have presented. Please expand on this comment for non-specialists

Line 345 – the word "indicated" is somewhat ambiguous in this context.

Figure 2a: please label AAV2 and HAdV here rather than only on B

Figure 2d & e: why are thresholds not shown?

Line 407 – I think cases are in green not red

Line 636: something is missing after "="

Author Rebuttals to Initial Comments:

We thank the editors and reviewers for their detailed comments. We note that these were broadly supportive of the work done at the time of review.

Referee #1 Overall, these are interesting and potentially important results that might contribute to reveal the etiology of unexplained pediatric hepatitis. **Referee #2** Overall this study is quite well done, and presents a case for a role of adenovirus and AAV2 in the liver disease seen, by an unknown mechanism. **Referee #3** Ho et al. provide a spectacular observation of high levels of AAV2 in plasma and liver tissue of pediatric patients with acute severe hepatitis thus adding a new layer of information to this mystery. **Referee #4** The paper presents a compelling result regarding the potential relevance of AAV2 and HLA alleles in the recent outbreak of non-A-E pediatric hepatitis, despite the small sample size

However, we note that the referees wanted to see further detail on the likely associated mechanism of disease and higher case numbers. Both of these points have now been addressed by the additional prospective recruitment of cases and by several additional experiments as detailed below. The mechanistic question was a common theme, and in this resubmission, we have added additional experimental data to address the points brought up by the referees.

We have now recruited 32 cases and 74 controls to our study and our initial findings remain consistent, both with the association with AAV2 and with the class II HLA allele DRB1*04:01.

Further detailed comments are outlined below in our point by point response.

Referee #1

"Is the model that this is a hit-and-run infection by one or more virus that progressed to more severe hepatitis after those viruses were largely cleared? If so, what is the pathological mechanisms causing the tissue damage, and can that be elaborated more clearly, including the association with HLA type explained?" "A clear statement of what data or additional studies are needed to confirm or disprove the hypotheses about the cause of these type of hepatitis cases would also be helpful."

Referee #3 "While the emergence of AAV2 is certainly novel, the study fails to make a causal link, merely making an interesting association. In its present form, the study lacks the rigorous scientific workup of this finding including proper hepatologic workup and immunologic studies to make any conclusion if we are dealing with an infectious hepatitis or an immune-mediated liver disease."

We have now conducted further studies that have shed light on these questions.

Firstly, we have now clearly demonstrated AAV2 in diseased hepatocytes in all five liver biopsies by *in situ* hybridisation.

Secondly, liver function over time appears to correlate with AAV2 viral load in several sampled cases.

Thirdly, the MHS class II HLA association is strongly indicative of a CD4+ immune-mediated pathology directed against AAV2-infected hepatocytes.

Finally, using immunohistochemistry and multiplex RNAScope technology, we have now shown the presence of MHC class II-expressing cells (likely CD68+ Kupfer cells), and a T cell infiltrate in all five case liver biopsies. In an explant liver, these were composed of CD4+ and CD8+ cells, with an activated phenotype. There were also CD20+ cells present and increased expression of Mx-1 indicating an increase in B cell and innate immune activity.

Further questions brought up by the reviewers are answered in detail below.

Referee #1

The main limitation of the study is its limited sample size and the retrospective design and the heterogeneity of the case- and control cohorts… One of the main limitations is sample size. It would be preferable to increase sample size in the case cohort. Given that 20 cases were available for HLA-genotyping, it may be possible to increase the sample size in the case cohort also for viral analysis (AAV2).

We have increased the sample size to 32 cases and have included a fourth control group (n=16) from Scotland that includes children attending hospital sampled during the time period of the study. Cases were recruited both retrospectively and prospectively – this has been clarified in the Methods text.

The authors need to better define inclusion and exclusion criteria for the control group

Inclusion and exclusion criteria added.

Was the liver damage as assessed in Fig, 1e-j scored and compared to other forms of (viral) hepatitis?

Yes – we have now added modified hepatic activity index score to the text (Supplementary Table 3).

Minor. Overall the quality of the figures could be improved (higher quality data plots, alignment of panels, adding missing scale bars in Fig. 1e-j, legible font size).

Figures now improved and expanded with additional data.

Referee #2

A study that examines for viral infections in 9 children in Scotland who developed otherwise unexplained hepatitis during the March and April of 2022. The context is that there was an unusual number of cases of hepatitis in children that were non-Hepatitis A-E virus; other cases have been reported elsewhere, but were not investigated. As controls, relatively similar samples were collected from three groups of: nominally healthy children; human adenovirus infected but no signs of liver disease; and other children with raised transamidases. While reported as a cluster, it is not clear that

this is an outbreak (e.g. of a connected transmission chain) or some other event. The other epidemiological and demographic data is a little hard to see, but may have been reported elsewhere(?)

We have now expanded the demographic and clinical data substantially to include further detail. The cases appear related by AAV2 phylogeny but are not identical (Figure 3a). Three cases were identical, two of which had a known epidemiological link. We have also added a new contemporaneous control group (Group 4).

The approach taken was to conduct deep metagenomic sequencing of the RNA and DNA from liver samples (in some cases) and from blood samples in all.

NB – we also sequenced throat swabs, faecal and rectal samples.

The results show the presence of varying levels of AAV2 DNA in the samples of the affected children, at a higher proportion and level than seen in the control samples. The various adenoviral and herpes viral DNAs were found in some cases. AAV2 is a common infection of people worldwide, with most having been infected early in life, and DNA is found to persist in tissues for many years, or even for life, with some integrated into the genome.

We agree with the reviewer that there is good evidence from various studies that people are exposed to AAV2 by adulthood. However, we now have good evidence that AAV2 infection is not ubiquitous in children of the age of this cohort. Indeed, we detected AAV2 at low level in only one patient in our Group 1 to 3 control groups. We also detected AAV2 in 4/16 Scottish contemporaneous (Group 4) controls at just above the detection threshold, indicating circulating virus at the time of the study. This motivated us to investigate if the AAV2 infection was acute in the cases. We now present more recent data demonstrating that most cases had evidence of recent AAV2 infection by the presence of AAV2-specific IgM, and variable titres of IgG over time. We now hypothesise that the AAV2 infection was acute. This is in keeping with previous literature on the age of exposure to AAV2.

The AAV2 infection has not been known to be associated with any adverse effects, and the virus rarely replicates in the absence of a helper adenovirus or herpesvirus. When delivered in very large amounts for gene therapy studies AAV capsids have been associated with liver disfunction and disease. Some adenovirus and HHV6 DNA was also reported, but at lower levels than the AAV2 DNA, or sometimes are absent. The suggestion made in the end is that the hepatitis is associated with an increase in human adenovirus infections, with a possible contributing or exacerbating role for the AAV2 coinfection.

As the reviewer points out, AAV2 has not previously been associated with hepatitis in children except in the context of gene therapy. Indeed, the delivery of AAV-vectored gene therapy is *commonly* associated with hepatitis in the context of T cell responses directed against viral capsid (used to deliver the gene of interest). Sometimes such responses have resulted in hepatic failure and death – this has now been observed with high and low doses of AAV8-vectored gene therapy. Because severe hepatitis is strongly associated with a class II HLA allele (DRB1*04:01), we

hypothesise that sporadic cases may have occurred in the past, undetected, as a cause of unexplained hepatitis in those with a genetic susceptibility. We have added this hypothesis to the text. We also propose that studies to carry out HLA typing of gene therapy recipients with severe hepatitis should be considered in the future. This has also been added to the updated text.

Phylogenetic analysis of the AAV2 sequences suggests that those fall into 2 clades which suggests possible epidemiological linkages between those viruses, and some mutations were present - there is no way to assess the significance of those mutations - there may be none.

We found that all cases occurred in one clade, within which there were two sub-lineages. The mutations we describe happen to have been investigated previously in some detail in a previous study that described phenotypic changes *in vitro* and in mouse models that might be relevant. It is possible that this lineage of AAV2 is more likely to cause hepatitis. We think that further characterisation of these mutations is indicated, ideally in an animal model.

Overall this study is quite well done, and presents a case for a role of adenovirus and AAV2 in the liver disease seen, by an unknown mechanism. It seems unlikely that AAV2 has suddenly emerged as a new pathogen in humans, so the title (and some of the text) is going to be misinterpreted by interested clinicians and others who do not read the study carefully or are not knowledgeable about the nuances of AAV and Ad biology. The main issues are around the presentation and not creating confusion in the future, particularly in the clinical fields - don't want everyone suddenly finding AAV2 DNA (which is quite ubiquitous in the human population), and diagnosing that as the cause of the hepatitis or other diseases.

We agree that we do not want the data to be misinterpreted - we haven't looked at adults in this study for example and so we have made sure to include paediatric in the title and also have now included the importance of the association with MHC class II – as most of the population are not likely to be susceptible.

We now present new data in support of a more direct association between AAV2 and hepatitis, including evidence of acute infection (as described above) in the majority of cases by IgM assay, the absence of high titres of AAV2 in multiple control groups and the presence of AAV2 in diseased hepatocytes in the liver by *in situ* hybridisation. While integrated AAV2 may be common in adults (we did not look at adults), we did not find AAV2 DNA detection to be ubiquitous in children, hence we think that AAV2 detection *in susceptible children* may be an informative test in the future (in combination with the HLA DRB1*04:01 allele). However, we acknowledge that further work is required to establish the pathogenesis of hepatitis in these children. We propose that humanised mouse studies may be useful and further studies of the epidemiology of AAV2 circulation in the community (in adults and children) and T cell ELISpots and flow cytometry are also indicated.

1) The title might be more accurate: "A cluster of hepatitis cases in Scottish children possibly due to adenovirus and adeno-associated virus infection"

We agree the title needed to be clarified and have now improved it to emphasise AAV2 in children and the importance of MHC class II susceptibility (we think that the Scottish bit is less relevant and we have had to keep this succinct due to Nature restrictions on permitted length of the title).

3) The Conclusion/final statements section is quite long and a little hard to follow. The key take-away seems ot be saved to the last paragraph, and this could be used as the lead in to this section.

The conclusions section has been rewritten to take account of these comments.

Referee #3 (Remarks to the Author):

"Ho et al. provide a spectacular observation of high levels of AAV2 in plasma and liver tissue of pediatric patients with acute severe hepatitis thus adding a new layer of information to this mystery." "The current wave of acute severe hepatitis of unknown etiology is a pressing issue in the (pediatric) hepatology community due to high rate of liver transplantation (5%) in these kids. The majority of case studies thus far found an association with adenovirus, particularly the F type HAdV41. This however is puzzling since adenovirus induced hepatitis is rare and usually restricted to severely immunocompromised individuals"

We agree with this – F type HAdV41 is not usually associated with hepatitis and hence AAV2 may be the missing piece in the puzzle. The ISH confirming the presence of AAV2 in ballooned hepatocytes is indicative of a direct causal link, while the HLA association is highly suggestive of an immune-mediated pathology.

Furthermore, on histology adenovirus hepatitis is hallmarked by nuclear viral inclusion particles, something that has not been detected in the current outbreak.

While we detected HAdV at very low level in the liver and other tissues in our initial investigation (by PCR and by sequencing), we did not demonstrate significant infection of hepatocytes by ISH. We also observed occasional inclusion bodies in the study and we have now made this clearer in the histology figure (Figure 4g) and in the text. Figure 4g: left panel, small, dark basophilic intranuclear inclusion body (arrow) in a hepatocyte next to the nucleolus; right panel bottom right corner with a hepatocyte showing a large, pale-basophilic, diffuse intranuclear inclusion body (suggestive of adenovirus infection, arrow) next to a multinucleated giant cell in the liver (*). We suspect that HAdV is performing a "helper" function – as the presence of AAV2 was far more prominent than either HAdV or HHV6.

Furthermore, the current cases occur in otherwise healthy children and it was thus far difficult to reconcile adenovirus infection with fulminant hepatitis leading to liver failure. As it currently stands there are four major unresolved questions:

i.) is there a novel adenovirus strain that might cause acute liver failure

We have shown that the HAdV41 strain detected in the cases is not a novel HAdV strain, therefore this hypothesis has been excluded. The closest sequence was HAdV41 isolate MU22 from Iraq in

2017 which differed by only 10 mutations, two other recent sequences reported in Germany were a further two mutations away from the detected sequence.

ii.) is another pathogen involved not yet detected

We carried out metagenomic sequencing and included host depletion methods, generating approximately 14 million reads per sample. While we concentrated on methods to identify viruses, we also used BLAST and Kraken-based methods to identify bacterial and parasite reads and did not find alternative pathogens of interest. We also applied target enrichment sequencing using a probe-design based on publicly available sequences of all known viruses to infect vertebrates until 2015 (VirCap) which has high sensitivity. Additionally, there was extensive clinical workup on multiple sample types, hence we believe that this is extremely unlikely.

iii.) what is the role of SARS-Cov2, which has been described to cause hepatitis and persistence in organ systems such as the intestine…

We do not think that SARS-CoV-2 is a likely cause of disease in cases of hepatitis for the following reasons.

Firstly, COVID-19 cases preceded cases of unexplained hepatitis by two years. Further, the spike of infections present several months before the cases occurred does not appear to have a similar pattern (**Figure 1a**). In contrast, HAdV diagnoses increased immediately before the onset of hepatitis cases. This was likely accompanied by AAV2 co-infection in the cases.

Secondly, we examined throat, plasma, liver, faecal and rectal samples for the presence of SARS-CoV-2 nucleic acid using a highly sensitive enrichment method and did not detect it. In contrast, we detected HAdV41 and AAV2 (also HHV6). Furthermore, we found that evidence of exposure to SARS-CoV-2 by serology was lower than the average in children in Scotland at the time the cases occurred. For an autoimmune hepatitis to emerge relating to SARS-CoV-2, we might expect to detect an over-active immune response targeting viral proteins. SARS-CoV2-directed antibody responses were investigated using a highly sensitive electrochemiluminscence assay Meso Scale Discovery V-PLEX), screening for both spike and nucleoprotein-specific antibody responses exposure.

AAV2 requires a helper virus to replicate as the authors correctly point out. As such the authors find HAdV and HHV6 in 6/9 and 3/9 plasma samples, 3/4 and 2/4 liver biopsies. How do the authors explain this? Does it not seem of that at least one patient must have had 3 viruses detectable?

We did detect three viruses in some patients. This is now clarified in the text and in **Supplementary Table 1**. The virus with a markedly higher number of reads was AAV2. Both HAdV and HHV6B were detected at far lower levels (and were not detected in hepatocytes).

With the new data that we have generated, and the increasing literature showing a strong association with HAdV and paediatric hepatitis, we think that the HAdV infection is likely to have been an acute coinfection alongside recent AAV2, acting as a “helper” virus. It is possible that the presence of HHV6B may also have had a similar function and in some patients, both HAdV and HHV6B may have acted together. We detected all three viruses in five cases and in two liver biopsies.

HHV6 commonly reactivates in the setting of critical illness and is also integrated into chromosomal telomeres in around 1% of the population[1]. In keeping with this, we found that there was HHV6 reactivation in the control subjects with critical illness also.

Clinical characteristics are minimal. The authors should add a detailed table with clinical parameters (mean ALTs, ASTs, IgG, histology, fibrosis, autoantibody titers, medication, BMI)

We have now included a detailed table with key biochemistry and autoimmune markers for all cases (Supplementary Table 2) and modified hepatic activity index score for the 5 cases that had a liver biopsy (Supplementary Table 3). Unfortunately, while we were able to access detailed clinical information on liver function, histology, autoantibodies and co-morbidities, BMI was not recorded in this study.

The authors write that autoimmune markers were negative. More details here should be provided. Unlike adults, for pediatric AIH, ANA titers of 1:20, SMA is 1:10-1:20 and LKM-1 1:10 is considered the threshold (Sebode et al. Front Immunol 2018). Furthermore, AIH 2 is often presented with normal IgG. How were autoantibodies detected? IFT or ELISA? What were the cut-offs? What was the ANA pattern? Were Hep2 cells used? This is critical here to rule out autoimmune hepatitis.

We now include detailed information outlining the workup for autoimmune hepatitis in the paper. All children admitted in the study were investigated for autoimmune liver disease. Five of the 32 included cases had a liver biopsy performed, none of whom had diagnostic features of AIH and in particular, no features of chronicity[2]. The routine work-up included immunoglobulins and autoantibodies against ANA, SMA, LKM1 and mitochondria.

The reviewer noted that lower antibody titres can be seen in children with AIH, and the review paper also highlighted the variation in practice between laboratories. We have therefore now described the testing in this cohort in more detail in the text as follows. ANA screens were performed on HEp-2 cells, an indirect immunofluorescence assay (IIF) with a screening dilution of 1:80. If the ANA screen was positive, an ANA titration was then performed. Results were reported as negative or positive (a positive result includes titre and pattern). Liver autoantibody screen was performed on rodent liver/kidney/stomach tissue (also an IIF). Most Scottish laboratories use a screening dilution of 1:40, though one immunology laboratory which tested samples in this cohort screened paediatric samples at 1:10. Results were reported as either negative or positive (a positive result included a pattern and, where appropriate, titre). The liver immunoblot was performed on 23 patients in the cohort. This assay identifies antibodies to the following antigens: M2, M2-3E, Sp100, PML, gp210, LKM1, LC-1, SLA/LP, and Ro52. Samples are diluted 1:100. This result was reported as negative, equivocal, weak-positive, positive or strong-positive. All subjects tested were negative.

As the aforementioned review paper describes, type II AIH is most common in younger patients (median age in this cohort was between 3 and 4 years old) and is strongly associated with positivity to anti-LKM1. All patients in our cohort were negative for anti-LKM1. Four patients had very low titres of anti-SMA (all 1:40 and all from the same laboratory – that tested samples at 1:10), and four different patients had low (1:80) titres for ANA (two speckled and two homogenous – see new Supplementary Table 3). A publication in 2018 reporting the clinical relevance of positive antibodies

in the diagnosis of AI liver disease describes low titres for ANA associated with other liver disease, in particular viral hepatitis and have low specificity at these levels.[3]

In the patients with positive autoantibodies, all had normal or normalising transaminases at last follow up in the absence of anti-inflammatory or immunosuppressive treatment. Furthermore, they had no evidence of chronic liver disease on ultrasound scan. It should also be noted that in the previously mentioned published comparable cohort of Scottish patients with AIH, all had immunology-based results consistent with AIH (i.e. high IgG with positive antibodies; with 100% of patients with type II AIH being positive for anti-LKM1), were older in age (median age 11.4yrs vs. 4.1yrs in our cohort) and had significantly lower median ALT at diagnosis/presentation (444 IU/L vs. 1756 IU/L) – no patient in the historical Scottish cohort improved without treatment. Lastly, given our accurate historical incidence data from 2013-2018 (0.5/100,000/year) we would have expected to have diagnosed between 4-5 patients in total across the whole of Scotland within the timeframe of the current study, with a median age of 12 years.

We have added these points to the text.

Along those lines, while liver histology samples are shown, much information is missing. HE stains clearly show interface hepatitis. What was the mHAI? Were the samples scored by liver pathologists, the author must have that data. What was the grading. Follwing the simplified AIH score at least the patient with ANA 1:80 would score a probable AIH if liver histology was typical (Hennes et al. 2008) and thus there need to be much more information to follow the authors argument that autoimmunity was excluded.

A table of modified HAI is now provided (Supplmentary Table 3). The samples were scored for mHAI by two pathologists. The histology of all liver biopsies is provided (Figure 1e-t) and was not typical of autoimmune hepatitis.

As the authors nicely discuss, the findings of AAV2 detection and the DRB1*04 HLA allele point to a CD4 T cell mediated pathology. However this should be analyzed. Did the authors stain for CD4 T cells? Are there data of a broad immune phenotyping? TCR phenotyping is also possible from FFPE tissue or could at least be attempted.

We now provide histology with staining for T cells (CD4+ and CD8+), and B cells (CD20+) which we agree has added a lot of useful information to the likely underlying mechanism of liver injury. Unfortunately, there was insufficient liver biopsy material available to carry out TCR sequencing reliably. We would have liked to collect PBMCs to carry out further B and T cell assays including immune phenotyping and also TCR sequencing but unfortunately these samples are not yet available to us and is not feasible within the timeframe of the current manuscript. We strongly agree with the reviewer that these questions will be important in follow-up studies and we plan to examine B and T cell responses in detail in due course. We have added this to the manuscript.

Can the authors provide serology for AAV2? Are these cases acute infection or reactivation? Of note, up to 80% of adults can have seroconversion and La Bella et al. (Gut 2020) found AAV2 in 20% of

individuals undergoing liver biopsy. Are the authors suggesting the cases are first time infected, which would align with the age group?

We now provide evidence that the AAV2 infection which is confirmed by serology in all cases is suggestive of acute infection as shown in our IgM and IgG assays. Further, there is evidence of changes in titre over time.

Data in Fig. 4 is really important, indicating that there is not a new strain of AAV2 but matching known strains. This seems somehow underemphasized in the manuscript and should be expanded.

We have added detail on this to the text. We agree that it is important that this is not a new strain although the lineage that all cases belong to contains mutations that have been associated with phenotypic change *in vitro* and *in vivo*. Instead, our data suggest that occasional cases of AAV2-related hepatitis may occur sporadically in those with HLA DRB1*04:01 and co-infection with HAdV. We have added a comment that future studies should be carried out to investigate historical cases of unexplained hepatitis. We think this is best thought of as a "triple hit phenomenon"; AAV2, HAdV and HLA class II susceptibility.

While AAV2 is know to have liver tropism, it is not commonly known to cause hepatitis. It is unclear why the current outbreak is linked to the presence of AAV2. Simple experiments such as infection of primary human hepatocytes or coinfections with adenovirus and AAV2 are lacking which would help to identify if liver damage would really be due to AAV2. Such experiments could show whether AAV2 infects hepatocytes and impacts hepatocyte viability / cellular metabolism (as opposed to trigger an immune response causing an "autoimmune-like" hepatitis).

We agree that wild-type AAV2 has never previously been associated with hepatitis. However, it is extremely common in recipients of AAV2- and AAV8-vectored gene therapy, most likely as a result of immune responses directed against AAV capsid proteins. Further, there have been fatal cases of hepatitis in children who have received AAV-vectored gene therapy (NB – there have been recent cases in both high-dose and lower dose gene therapies). One important future question that our data raise is whether there may be an HLA association with severe hepatitis cases who have received AAV-vectored gene therapy. It could be helpful in future in guiding the method of gene therapy to use in those who have the HLA DRB1*04:01 allele. We have now added these points to the text.

We agree that further mechanistic work is indicated. However, as there is a strong MHC class II association, we think that we would need to express MHC class II in AAV2 infected cells to facilitate peptide expression, followed by CD8+ cytotoxic killing assays. A humanised mouse model with inducible expression of MHC class II is something we have also considered. Unfortunately, these experiments are not feasible within the timespan of publication of this manuscript before someone else does. However, we will certainly explore these avenues in the future.

10. Almost all cases were actively or previously infected with SARS-CoV-2. Furthermore, both the US and UK have the highest reported cases of unknown acute hepatitis and both countries had high numbers of all previous SARS-CoV-2 variants and have a high rate of seroprevalence compared to other places such as continental Europe. This at least hints at a potential role for SARS-CoV-2 in

these cases. While discussed, both studies take little efforts to rule out Covid-19 as a potential explanation. Scientific questions that need to be addressed are SARS-CoV-2 persistence in stool, T-cell receptor skewing in the liver and systemic or hepatic IFN-γ upregulation. These analyses could provide evidence of a SARS-CoV-2 superantigen mechanism in an adenovirus-41F- or AAV2 sensitised host. While I appreciate that Covid must not be the explanation per se, of all the potential pathogens, Covid is the new kid on the block and both studies do not provide sufficient evidence by a mile to rule out a contribution of Covid in this epidemic of cases.

We argue that SARS-CoV-2 is unlikely to be directly related to this phenomenon and have generated further data that strengthens this position.

Firstly, on testing 32 cases, we found that exposure to SARS-CoV-2 by serology was lower than background prevalence. Nearly half of the children had no evidence of exposure to SARS-CoV-2 – if they were to develop a superantigen response driven by a determinant in the SARS-CoV-2 spike protein, it is reasonable to assume that we might detect a serological response against SARS-CoV-2 using a highly sensitive assay platform (MSD V-PLEX). We examined responses against both spike and nucleoprotein in order to increase sensitivity of detection.

Secondly, the epidemiology is not suggestive of a SARS-CoV-2-related phenomenon as COVID-19 started 2 years before the onset of hepatitis cases.

Thirdly, we looked hard for SARS-CoV-2 in several types of samples and did not find it in any of the cases by NGS. Regarding persistence in stool, we looked for this and did not find it. Nor did we find SARS-CoV-2 in any of the tissues or blood samples that we analysed. In contrast, we found AAV2 in all cases.

While we don't think that there is any evidence that the cases were directly related to SARS-CoV-2, we do suggest in the text that it is very likely that there is an indirect relationship as a result of increased co-circulation of respiratory and gastrointestinal pathogens following the removal of measures put in place to prevent COVID-19 transmission.

The evidence provided that this is indeed a viral hepatitis is limited. The association of high levels of AAV2 in affected individuals both in liver and blood is striking. However, the study did not find evidence of virus proteins, signs of infected hepatocytes on histology, or even provide principal evidence of AAV2 directed hepatotoxicity from model systems. This strongly suggests an immune-mediated pathology as discussed. To investigate this further, a proper immunologic workup, including TCR sequencing, immunofluorescence, immune phenotyping etc. should be done. This is of highest clinical relevance as the result for clinical management will be either immunosuppression or directed novel antivirals.

We have now undertaken a detailed analysis, as suggested by the reviewer and have carried out several critical experiments to investigate this further. As described above, we now provide detailed histological evidence of AAV2 infection of diseased hepatocytes in association with an inflammatory infiltrate and a characteristic histological appearance in all liver biopsies. There is evidence of active

AAV2 replication in diseased “ballooned” hepatocytes in all five liver biopsies, including in the cytoplasm as well as the nucleus of infected cells (ruling out latent infection). This points to an immune-mediated pathology triggered by ongoing AAV2 infection in diseased hepatocytes. The evidence for immune-mediated pathology includes the histology revealing a dense MHC class II infiltrate, the presence of activated CD4+ and CD8+ cells and the strong association with the class II DRB1*04:01 allele. Further, we demonstrate the presence of IgM positive antibody responses in the majority of cases, suggestive of acute infection. Additionally, longitudinal sampling of liver function and AAV viral load showed correlation of these variables in multiple cases.

After exclusion of known etiologies, autoimmune hepatitis is often the cause of acute liver failure in both adults and children. AIH can be triggered by drugs or infections (molecular mimicry) and it has been speculated that these acute severe hepatitis cases are due to AIH-like inflammation triggered by e.g. adenovirus. Both studies undertake little efforts to rule out or investigate virus-triggered AIH. It would be paramount to provide histology of the liver samples, scoring by a liver pathologist, autoantibody titers, quantitative immunoglobulins etc. I assume that the authors have those datasets and they should be provided. In fact, the genetic association with HLA-DRB1*04:01 is strong for AIH and is even part of the international AIH study groups (IAIHG) diagnostic scoring, further suggesting these kids could have an AIH-like hepatitis.

We agree that the strong association with the MHC class II HLA-DRB1*04:01 allele is highly suggestive of an immune-mediated pathology. However, following a series of experiments in the period while this paper was under review, we do not think that this has the histological features of a primary autoimmune hepatitis as outlined above. Crucially, we have now detected a strong signal for the presence of AAV2 in ballooned hepatocytes and endothelial cells. NB while type 1 AIH is associated with the MHC class II HLA-DRB1*04:01 allele in adults, it is not a common association in children with AIH. Further analysis of T cells is indicated when these become available to identify peptides presented via this allele.

One potential clue to understand if AAV2-induced hepatitis is occurring vs. an autoimmune hepatitis (potentially AA2-triggered) could result from analyzing complement factors. Complement does not play an important role in AIH (Biewenga et al. 2020, PMID 32234403), however, complement activation and deposition is an essential part of the immune response to AAVs as seen in rodents (Zaiss et al. 2008, PMID 32234403) and in the clinical trials that used AAVs as vectors that reported severe side effects (ClinicalTrials.gov identifiers: NCT03362502 and NCT03368742, reviewed by Muhuri et al. 2021, PMID 33393506). Could the authors stain complement in liver tissue and measure complement factors in blood?

We took the reviewer’s advice on this and stained for complement in liver biopsies. All biopsies were negative. Complement factors were also measured in blood for 12/32 cases; all were within normal range.

Minor comment:
Is there additional data on why the cases developed severe hepatitis vs. controls? More C-sections? Breastfeeding? Environmental factors that could explain higher incidence such as daycare exposure,

more or less quarantine?

We agree that it would be very interesting to know about these but unfortunately these details were not included in the protocol used to collect data from affected children. At this time, we are not able to comment on daycare exposure or quarantine beyond the national measures taken in Scotland.

Referee #4 (Remarks to the Author):

The paper by Ho et al., reports the detection of AAV2 in cases of acute hepatitis of unknown etiology (also referred to as non-A-E hepatitis). The authors present sequencing and molecular diagnostics to support the presence of AAV2 and helper viruses (HAdV and HHV6) in cases but rarely in several different control groups. The authors also show an increased frequency of HLA-DRB1*04:01 in cases compared to controls, suggestive of a role of human genetics in predisposing cases to hepatitis.

The paper presents a compelling result regarding the potential relevance of AAV2 and HLA alleles in the recent outbreak of non-A-E pediatric hepatitis, despite the small sample size. While this result is novel and timely, the presentation of the data is confusing and, at times, repetitive, which makes obscures the contribution of this work. Additionally, very little context is given for the motivation of this study and how to interpret the results until the discussion.

We have introduced several introductory sentences to explain the motivation for and design of each part of the study throughout the text. We have also reorganised the results section to make interpretation easier.

I will focus my technical comments on the molecular methods used (PCR and sequencing), which are in general appropriate and analyzed using reasonable metrics. However, the presentation makes interpretation challenging. The genome sequencing sub-section is very long compared to all other sections and would benefit from both shortening and restructuring to more clearly define the purpose of each paragraph and the analyses being presented within each, including more clearly separating sequencing and PCR results and removing redundancy / clarifying where the same result is being presented in a different way.

We have now restructured each section and removed redundant text. To clarify the different steps of the investigation, we have separated each section more clearly, including separation of the PCR and NGS results.

Below I give several specific comments and questions regarding the data presented in this section;
- Can the authors clarify if all numbers of viral reads reported in text are from targeted enrichment libraries and from DNA or RNA extracts? It seems that RNA in general was much less sensitive than DNA sequencing for these pathogens. If this is not used in any analysis it may be better presented only in supplementary material.

The viral reads reported in the main text are from target enrichment sequencing which was more sensitive than metagenomic sequencing (included only in the supplementary material). We would like to include the RNA data in the main text because it is suggestive of replication of the DNA viruses that we detected. This is also the case with the RNAScope *in situ* hybridisation. We have now emphasised this point in the main text.

- Are the full AAV2 genomes reported all from plasma/serum or other compartments? Also, in the mutation analyses why are only an arbitrarily defined group of closely related sequences compared for the presence of these mutations? I think it is important to also know if these mutations are found in other clades

The AAV2 genomes presented in the phylogenetic/mutational analysis are those obtained from the plasma samples (using viral target enrichment) as these samples consistently represented the largest read numbers and most complete genomes across samples/runs, however AAV2 sequences (of various degrees of completeness) were obtained from other compartments of the same patient and were found to be the same sequence. For the phylogenetic and mutation analysis, the number of sequences has now been expanded to be more complete and included all those returned via BLAST against the GenBank NT database that were a complete linear genome and had a BLAST query coverage > 75% (Supplementary Figure 3). The mutations highlighted in Figure 3 were those observed in 5 or more of the 9 AAV2 genomes obtained from the Scottish cases, these mutations are observed at various frequencies in other AAV2 genomes.

- Line 177 – I am surprised that so many different pathogens are detected in such a small cohort of samples. It is noted that these are detected by sequencing but I couldn't see where this data was shown nor details of the sample types, which would have been interesting.

The pathogens detected were all also confirmed by PCR during the course of routine investigation and the source of each virus is now outlined in Supplementary Table 1. The herpesviruses detected at low level were likely present as a result of reactivation in the context of severe/critical illness. We also detected other pathogens such as norovirus (also detected by routine testing) which may have been acquired in hospital.

- Are these 9 genomes of the sub-lineage AAVv66 or they are classified as a different sub-lineage but share these mutations with that sub-lineage? If the latter, is there any evidence of recombination that could explain this?

The 9 genomes appear to be in the same sub-lineage as AAVv66 (for which a whole genome has not been published; only a partial genome is available – accession MT496778). There was no evidence of recombination.

- Is there a reason why control group 3 was not sequenced? Their inclusion for qPCR but not

sequencing feels a little sudden; it would be helpful to explain why the addition of this control group via PCR only is beneficial to the analysis.

Control group 3 was included after the initial sequencing study and was included in the case-control comparison to explore whether or not AAV2 may reactivate commonly in patients with severe hepatitis or critical illness of other aetiology as a bystander event. We did not seek to look for other pathogens in this group. We have altered the text to make this much clearer, separating out each stage of the different experiments and question more clearly.

- It was also striking from Figure 3, though not mentioned in the text, the one HAdV+ control with many coinfections. Is there anything clinically that would be consistent with this observation?

We looked into this in more detail.

The HAdV control with many HHV co-infections had evidence of low levels of herpesvirus reads of several species and high numbers of reads only for CMV and HHV6B. Some of the low levels of reads were conserved reads common to several herpesviruses as presented in Supplementary Table 8. In order to provide a better understanding of the herpesviruses present in this case, we carried out further bioinformatic analysis by filtering read mappings to discount those with any alignment clipping and those where both members of the pair do not map. In addition, Kraken was used to classify reads into viral bins based on unique kmers specific to viral species, these hits were further filtered to only consider those of high confidence. In both cases, the hits to the different viruses remained, but it is important to remember that read numbers were low for HHVs other than CMV and HHV6B.

The case was a critically unwell child who had been given immunosuppressant therapies for a rare inflammatory condition (we are unable to disclose the exact nature of the disease to avoid deductive disclosure). It is therefore likely that the HHVs reactivated in this context.

- Please specify sample type for control samples. It seems it is a mix of serum and plasma but I was unsure which was which.

Group 1 healthy controls were all serum samples, Group 2 included 8 plasma and 5 sera, Group 3 consisted of 27 sera and 9 plasma and Group 4 were all plasma. This has been added to Supplementary information. We tested Ct values and GAPDH in a small number of paired serum/plasma samples and found these to be similar (data not shown).

- Should qPCR actually be referred to as RT-qPCR as a reverse transcription was performed?
- In the methods, please add details of how RPM is calculated.

The reviewer is correct that we used RT-qPCR – this has now been clarified in the text.

Reads per million were calculated as the number of viral reads per million reads sequenced to normalise for variation in numbers of reads sequenced per sample. This has been added to the Methods section.

Regarding the figures for this section (Figures 2 and 3), I found the representation of control data in heatmaps to be hard to read and potentially misleading for a couple of reasons. First, it implies there is something relevant about the ordering of individuals relative to cases which is not the case. Secondly, low read counts (especially on Figure 3) are hard to read. I think it would be more clear to show controls as aggregate data via a barplot, scatterplot or similar. Thirdly, presenting this in a binary format will allow to more directly compare the statistics between sequencing and qPCR.

We have retained the heatmaps - the colour scheme has been modified for clarity for low read counts because we feel that it is a useful way of showing the results from multiple sample types in the initial investigation. However, we have also included scatterplot information in the main figure and in supplementary material for qPCR and sequencing data, allowing statistical analysis. This has also been updated to reflect the larger number of cases that we have and also inclusion of serology data.

Below are several comments on clarity of presentation and structure relating to the genome sequencing section.
- It seems that the first paragraph is intended to be a general overview of the sequencing and reporting just information from plasma. This was clear but then the PCR liver result for HHV6 is mentioned at the end which would be better placed elsewhere

Moved/clarified as suggested.

- The next two paragraphs are diving more deeply into the genomics results on AAV2 and then HAdV and HHV6. This again makes sense but it would be helpful to have topic sentences on these paragraph to set-out this goal so the reader can more easily follow.

We have restructured the results section for clarity, as suggested.

Additionally, though the AAV2 paragraph focuses exclusively on genomic data, the HAdV/HHV6 paragraph starts to bring in PCR data. This is inconsistent and hard to follow what the main result being presented is.

We have now separated the NGS and PCR data.

- In some cases the same information appears to be presented multiple times but with subtle differences. This is particularly true of the PCR results, which is challenging to follow. E.g. lines 64-65 vs lines 114-115 lists routine blood tests run but HHV 6/7 is only in the latter list. As similar statement is again made on line 172 but it is unclear if this is a separate round of PCR testing for these pathogens or the same as the initial clinical diagnostics. Relatedly, are the PCR results that are presented alongside genome sequencing data (e.g. see lines 141, 163, 173) different to those

presented as part of the “real-time PCR” sub-section? If not this should be made clear. If so then these should be presented just once and then clearly referred to elsewhere. Detection of AAV2 by sequencing in 9/9 plasma is also mentioned on lines 137 and 144.

We have now resolved the repetition in the text.

- As the case-control comparison is one of the major analyses of the study, I suggest giving this a separate sub-heading. Initially presenting detection in the controls and then the statistical results would improve readability.

Agreed – we have separated this as suggested.

I found the uses of reads/million for the cases vs control comparison surprising and potentially problematic. In order for this to be robust I believe the authors need to more clearly demonstrate that their sequencing results are truly quantitative. In particular, that there are – as I understand it – a mix of serum and plasma samples here may influence the amount of host background.

We used plasma samples for cases and a mixture of plasma and serum samples for controls as outlined above. We did not find a significant difference between paired serum/plasma samples for AAV2 Ct (and have also not found this in previous studies using similar methods). We also found similar GAPDH levels in a small number of paired sera/plasma samples.

Regarding the quantitative aspect of PCR vs target enrichment (TE) sequencing, we have now included a correlation co-efficient analysis in Supplementary Information. TE sequencing reads correlated with viral loads estimated by PCR. We and others have found this in other studies also.[4,5]

Additionally, I am inferring that sample collection and storage was not the same across case and control cohorts given that these are from distinct studies and sites. All of these variables can influence the relative amounts of human vs viral reads as well as the success of hybrid capture, which further alters this relationship. I strongly suggest that this be done as a qualitative analysis (just presence/absence of the virus) in order to be more robust to such variation.

There may be slight variation in storage of samples at different sites but samples obtained for the study were stored at -80 degrees and sent by courier on dry ice. Just in case, we carried out a binary analysis using Fisher’s exact test, as suggested, in addition to the quantitative analysis. The result is similar with a p value also significant at $p<0.0001$.

Although in line 181 it is noted that case-control analysis was also done for real-time PCR results this analysis is only briefly referred via citation of the figures 2c-e and this data is not explored further. Given that all control groups are included here it seemed surprising that the significant difference here was not mentioned.

We have now brought this result into the narrative more clearly and updated with the new cases and controls.

I believe this is a more reasonable way to look at differences in load than the sequencing data. However, if copy number is available for qPCR data (which it seems from the written results that it may be) then it would be preferable to present this than CT values, which are not comparable across runs.

This is a fair point - we now include graphs of adjusted copy number instead of Ct.

Overall, the paper appropriately credits previous work. However, in line with my comments below about clarity, the introduction, in particular, would benefit from more citations in particular around what is known about these viruses and the motivations for the study.

We have added detail on motivation for the study and, as below, further detail on previous T cell work in the context of AAVs being used for gene therapy.

One place where a citation is lacking is on line 263 where the statement about T-helper responses should have a citation. Potentially those given on lines 327-328 could be moved here as this section also feels out of place.

Reorganised and cited.

The presentation of the manuscript feels overall heavily geared towards a clinical audience and the manuscript overall would benefit from more clear framing of the key questions and differentiation between what was known of these samples prior to this study vs what is newly reported here.

We have restructured the manuscript and added more narrative to outline the questions more clearly and what we have found that is novel.

- In the first section (Hepatitis outbreak in Scottish children) it would be helpful to include more background that would help the reader understand the relevance of the viruses being mentioned and also the central hypothesis the paper is looking to resolve. In particular, it is not clear to non-experts why if the majority of cases were already found to have HAdV this is insufficient as an explanation for these cases.

We have added further detail on the thinking here – firstly that HAdV41 is not commonly associated with hepatitis but would be required for AAV2 infection. The additional *in situ* hybridisation experimental data that we have added is more suggestive of causation than the data we presented in our previous version. The diseased (ballooned) hepatocytes in all five biopsies contain AAV2 while it was not detected in control liver samples.

Similarly, HHV6A and 6B incidence over 2021-2022 are mentioned in the final sentence but it does not become clear why this is relevant to the current cases until the results of this study later on.

We have clarified this.

- On a related note, the first mention of AAV2 as needing a helper virus such as HAdV or HHV6 is on

line 179 when the case control analysis is introduced. This is something that would be better brought up earlier to explain why this analysis is being done

Agreed – we have introduced this now earlier in the text.

Presentation of the control datasets is not consistent and there is variability in whether they are referred to as groups 1-3 or by clinical traits. When clinical traits are used, these too can be variable (e.g reference to transaminase levels, liver function, ALT, or LFT). Overall, my suggestion is to ascribe simple but descriptive identifiers to these 3 groups when first presented and use them consistently throughout in both text and figures. This variability makes it further very hard to trace other sub-groupings (e.g. the 12 children in clinical care on line 191 and 197).

We now refer to the control samples as group 1, 2, 3 and 4 (added group) throughout the text to avoid confusion.

- I wonder if presenting the clinical presentation before the control cohorts would flow more naturally. With the current format comparisons between cases & controls in age are presented before the cases have been described.

We agree that this works better. Clinical presentation has been included now in much greater detail and moved to the first part of the text.

- In the conclusion, the relevance of the paragraph lines 302-308 was not clear. Though the clinical descriptions are helpful context I feel this would be better placed earlier in the paper where background on why HAdV and HHV6 are unlikely to be sufficient explanations for the observed hepatitis

Moved as suggested to an earlier part of the text.

- The limitations paragraph seems to also include additional interpretation of results. I suggest more clearly separating these two things. Specifically, everything from line 329 onwards should probably be moved sooner as they form part of the paper's hypothesis not its limitations. Which should be discussed separately.

Separated as suggested.

Below are several smaller technical comments that would also aid clarity of presentation;
A specific definition of non-A-E hepatitis / hepatitis of unknown etiology would be very helpful. This could be as simple as a parentheses in the sentence that ends on line 65.

Added as suggested.

I found it hard to understand when the authors use HHV6 vs HHV6B and HAdV vs HAdV41 or Adenovirus species C and F. Could this be clarified?

This reflects differences in clinical PCR testing and sequencing. The latter allows better discrimination of species type.

Is the case shown in figure 1e-j representative? Can other images be shown in the supplement?

Yes – we have now updated the Figure substantially to add histology, *in situ* hybridisation and inflammatory infiltrate staining.

The (n=12) on line 94 is unnecessary as since the number was already written in the sentence

Removed

Line 145, it would be better to say AAV2 genomes were assembled rather than detected

Corrected

Line 147, it would be better to say seven distinct sequences rather than lineages. Lineages implies more clear separation and structure than is observed here where all samples are highly genetically similar though not identical

Corrected – we actually think that the sequences belong to one clade – this has been made clearer.

What does "(TE)" mean on line 181?

Target enrichment (TE) now defined in the text.

Line 197 – I was unsure where the 2/12 with detectable HHV6B reads was already described above.

Clarified

Line 252 – should "patients" be changed to "cases" to make it clear this does not include other unwell children from control groups?

Yes – this is better. Corrected.

Line 278 – it would be good to clarify that this was also already shown and is not specifically a result of this paper

Clarified**.**

Line 287-288 & 324; it was not apparent to me why in each instance HHV6 was not likely based on other information the authors have presented. Please expand on this comment for non-specialists

We have expanded on this in the text.

Line 345 – the word “indicated” is somewhat ambiguous in this context.

Corrected

Figure 2a: please label AAV2 and HAdV here rather than only on B

Corrected and updated substantially

Figure 2d & e: why are thresholds not shown?

Corrected

Line 407 – I think cases are in green not red

Corrected

Line 636: something is missing after “=”

Corrected

References

1 Pantry, S. N. & Medveczky, P. G. Latency, Integration, and Reactivation of Human Herpesvirus-6. *Viruses* **9**, doi:10.3390/v9070194 (2017).

2 Mieli-Vergani, G. *et al.* Diagnosis and Management of Pediatric Autoimmune Liver Disease: ESPGHAN Hepatology Committee Position Statement. *J Pediatr Gastroenterol Nutr* **66**, 345-360, doi:10.1097/MPG.0000000000001801 (2018).

3 Sebode, M., Weiler-Normann, C., Liwinski, T. & Schramm, C. Autoantibodies in Autoimmune Liver Disease-Clinical and Diagnostic Relevance. *Front Immunol* **9**, 609, doi:10.3389/fimmu.2018.00609 (2018).

4 Fogel, J. M. *et al.* Performance of a high-throughput next-generation sequencing method for analysis of HIV drug resistance and viral load. *J Antimicrob Chemother* **75**, 3510-3516, doi:10.1093/jac/dkaa352 (2020).

5 Thomson, E. *et al.* Comparison of Next-Generation Sequencing Technologies for Comprehensive Assessment of Full-Length Hepatitis C Viral Genomes. *J Clin Microbiol* **54**, 2470-2484, doi:10.1128/JCM.00330-16 (2016).

Reviewer Reports on the First Revision:

Referee #1 (Remarks to the Author):

Ho et al. have added valuable new data into their manuscript. Particularly, the prospective recruitment of additional cases as well as a contemporary control cohort to control for seasonal variations in AAV2 incidence have significantly strengthened the robustness of the results and thge conclusions of the study.

Histological analysis, specifically, ISH and IHC demonstrating the presence of AAV2 in damaged hepatocytes in liver specimen of hepatitis patients, but not observed in controls, further supports the conclusions.

Overall, the manuscript provides solid evidence that acute AAV2 infection (helped by hAdV co-infection) is associated with unexplained pediatric hepatitis in this observational cohort / case-control study.

Additional data also strengthen the observed association of unexplained hepatitis with MHC class II allele HLA DRB1*04:01. However, the authors noted that other HLA alleles were also associated with hepatitis and that the identification of the exact susceptibility allele is hindered by a strong linkage disequilibrium at this locus. The observed HLA-association will require further confirmation in larger cohorts and this should be discussed in the paper.

Based on their data, the authors propose that unexplained pediatric non-A-E hepatitis is caused by AAV2-trigered, MHC-II-dependent CD4+ T cell driven immunopathology. The authors included additional immunofluorescence staining revealing pronounced infiltration of MHC-II+/CD68+ macrophages and activated T cells in the liver tissue of patients with severe hepatitis to underscore this hypothesis.

Overall, the revision has substantially improved the study and I have only few minor comments and suggestions.

1. I suggest to add labels (“Masson”, “H&E” and “case”, “Control”) to Fig. 1f-t for better clarity. Also, panel l lacks a label. The figure legend is hard to follow and it is interpreting the findings rather than stating the sample and method.
2. Figure 2 is so grainy (low resolution) and the labels are so small that it is hard for me to make out details.
3. What is the difference between the top row in 2a and 2b? It appears that both show the same sample IDs, same specimen (Plasma/Serum) and the same results (AAV2 RNA / DNA) – the same for the “cases” further down. It is not quite clear why it is shown twice in a different order.
4. The overall quality of the figures should be improved.

Referee #2 (Remarks to the Author):

This is a major revision of the previous manuscript, and has been changed in a variety of ways - those include the addition of more patients and a large number of additional tests and data which really makes this a completely different manuscript. The authors dismissed my previous concerns about the role of AAV2 versus adenovirus in the liver disease, so I will not repeat those, and their claims of higher viral loads of AAV2 DNA in the liver of clinical cases, and evidence of recent viral infection make a stronger case for a direct connection between the AAV2 and the severe liver disease seen.
There is no marked-up version to compare the changes, but based on the responses (rebuttal) to the previous reviews it appears that this is significantly improved. It is hard to tell if this will hold up on prolonged scrutiny, but this makes a much better case for a role in the disease. The comments by the other reviewers asking how such a (likely) poorly replicating virus may cause such a severe hepatitis are still germane, and likely point to an autoimmune or inflammatory effect that causes the tissue destruction.
A possible useful model for comparison which is not cited here is the discovery of a parvovirus is horses that appears to be associated with severe liver disease, despite replicating in only a few cells in the liver...https://pubmed.ncbi.nlm.nih.gov/35907440/
https://pubmed.ncbi.nlm.nih.gov/32192415/

Referee #3 (Remarks to the Author):

The authors have undertaken substantial efforts to respond to my comments. I still have some doubts regarding the autoimmune phenomena (e.g. autoantibody titers, which are relatively high for children), but I acknowledge the line of arguments favoring the AAV2 (and HLA class II) hypothesis. The additional cases, experiments and analyses of liver biopsies strongly support the original finding.

Referee #4 (Remarks to the Author):

The manuscript by Ho et al is overall greatly improved by the revisions that have been made and the narrative is clear and compelling. The authors have responded adequately to my comments. I have a small number of additional comments on the revised version.

In the abstract the authors suggest an “interplay” of virus co-infection and host genetics. This word to me suggests an interaction between these factors. While I find the individual evidence presented to support the role of i) AAV2 and helper virus infection, and ii) HLA class II DRB1*04:01 allele to be strong, the authors do not test any sort of interaction effect of these different factors and thus I think this would be more accurately presented as both of these factors playing a role.

The structure and presentation of findings is much clearer now. However, the HLA allele association result in particular would benefit from more citations / a couple of sentences clearly laying out the link to CD4+ T cell responses for non-experts. Additionally, on line 369 the use of “demonstrate” is too strong given that the finding is of a statistical association. I suggest “support” or “indicate” as more appropriate alternatives.

The authors indicate that control group 2 all tested positive for HAdV, however sequencing reads were only identified in 9/12 samples. Do the authors suspect this is because the samples were plasma / serum (elsewhere it is noted that whole blood would be preferred but was unavailable), as this seems to conflict with the greater sensitivity of TE sequencing compared to PCR in cases? Elaborating on this in the limitations paragraph would be appropriate. More generally, I suggest modfying the limitations paragraph, which currently focuses more on the HLA typing result than general limitations of the study. In particular, it would be good if the authors could explicitly acknowledge the differences between the case and control cohorts (age, time of collection) as well as differences in detection sensitivities between sample types that could impact the findings.

In general the figures are still not of high quality. Layout and coloring of figure 1 could be modified to improve readability. It would be helpful to see HHV6 on figure 2A. In figure 3, it is hard to see how 3A and 3C relate to each other – could they be put alongside each other with the same ordering? The coloring is also confusing: I was under the impression the colors reflected whether a sample was part of this study but the first sample from this study (about 7 from the top) appears to have some green and some blue mutations. Lastly, p-values on figure 3C would be helpful.

There are several grammatical and formatting errors throughout in both text and figures. Please read very carefully prior to resubmission (e.g. ensuring all acronyms are defined, all figure axes have labels).

Author Rebuttals to First Revision:

Referees' comments:

Referee #1 (Remarks to the Author):

Ho et al. have added valuable new data into their manuscript. Particularly, the prospective recruitment of additional cases as well as a contemporary control cohort to control for seasonal variations in AAV2 incidence have significantly strengthened the robustness of the results and thge conclusions of the study.

Histological analysis, specifically, ISH and IHC demonstrating the presence of AAV2 in damaged hepatocytes in liver specimen of hepatitis patients, but not observed in controls, further supports the conclusions.

Overall, the manuscript provides solid evidence that acute AAV2 infection (helped by hAdV co-infection) is associated with unexplained pediatric hepatitis in this observational cohort / case-control study.

*Additional data also strengthen the observed association of unexplained hepatitis with MHC class II allele HLA DRB1*04:01. However, the authors noted that other HLA alleles were also associated with hepatitis and that the identification of the exact susceptibility allele is hindered by a strong linkage disequilibrium at this locus. The observed HLA-association will require further confirmation in larger cohorts and this should be discussed in the paper.*

We have added another sentence on this in the limitations section of the manuscript – it is also already described in the HLA section.

Based on their data, the authors propose that unexplained pediatric non-A-E hepatitis is caused by AAV2-trigered, MHC-II-dependent CD4+ T cell driven immunopathology. The authors included additional immunofluorescence staining revealing pronounced infiltration of MHC-II+/CD68+ macrophages and activated T cells in the liver tissue of patients with severe hepatitis to underscore this hypothesis.

Overall, the revision has substantially improved the study and I have only few minor comments and suggestions.

1. I suggest to add labels (“Masson”, “H&E” and “case”, “Control”) to Fig. 1f-t for better clarity. Also, panel l lacks a label. The figure legend is hard to follow and it is interpreting the findings rather than stating the sample and method.

We have added the labels to the figure, as suggested and have clarified/shortened the legend.

2. Figure 2 is so grainy (low resolution) and the labels are so small that it is hard for me to make out details.

A higher quality image of Figure 2 is provided as a pdf file.

3. What is the difference between the top row in 2a and 2b? It appears that both show the same sample IDs, same specimen (Plasma/Serum) and the same results (AAV2 RNA / DNA) – the same for the "cases" further down. It is not quite clear why it is shown twice in a different order.

We do show the same data in 2a and 2b – but these are being compared with other sites in panel (a) and control samples and PCR in panel (b) – we think it is easier to look at this way. We think it starts to look very untidy if you take the serum out of panel (a) or the cases out of panel (b) and difficult to interpret. They are not in a different order.

4. The overall quality of the figures should be improved.

These are now provided as high-quality pdf files.

Referee #2 (Remarks to the Author):

This is a major revision of the previous manuscript, and has been changed in a variety of ways - those include the addition of more patients and a large number of additional tests and data which really makes this a completely different manuscript. The authors dismissed my previous concerns about the role of AAV2 versus adenovirus in the liver disease, so I will not repeat those, and their claims of higher viral loads of AAV2 DNA in the liver of clinical cases, and evidence of recent viral infection make a stronger case for a direct connection between the AAV2 and the severe liver disease seen. There is no marked-up version to compare the changes, but based on the responses (rebuttal) to the previous reviews it appears that this is significantly improved. It is hard to tell if this will hold up on prolonged scrutiny, but this makes a much better case for a role in the disease. The comments by the other reviewers asking how such a (likely) poorly replicating virus may cause such a severe hepatitis are still germane, and likely point to an autoimmune or inflammatory effect that causes the tissue destruction.
A possible useful model for comparison which is not cited here is the discovery of a parvovirus is horses that appears to be associated with severe liver disease, despite replicating in only a few cells in the liver...https://pubmed.ncbi.nlm.nih.gov/35907440/
https://pubmed.ncbi.nlm.nih.gov/32192415/

The article mentioned is helpful – and I have cited this alongside a similar level of infected hepatocytes for HCV.

Referee #3 (Remarks to the Author):

The authors have undertaken substantial efforts to respond to my comments. I still have some doubts regarding the autoimmune phenomena (e.g. autoantibody titers, which are relatively high for children), but I acknowledge the line of arguments favoring the AAV2 (and HLA class II) hypothesis. The additional cases, experiments and analyses of liver biopsies strongly support the original finding.

We are very grateful for those helpful comments – we favour the immune pathology angle as a direct effect but agree more work is needed to delineate what epitopes the immune response is directed against – we plan for further B and T cell assays in due course when we have access to PBMCs from the children – this is planned but not feasible in the timespan of this paper.

Referee #4 (Remarks to the Author):

The manuscript by Ho et al is overall greatly improved by the revisions that have been made and the narrative is clear and compelling. The authors have responded adequately to my comments. I have a small number of additional comments on the revised version.

*In the abstract the authors suggest an "interplay" of virus co-infection and host genetics. This word to me suggests an interaction between these factors. While I find the individual evidence presented to support the role of i) AAV2 and helper virus infection, and ii) HLA class II DRB1*04:01 allele to be strong, the authors do not test any sort of interaction effect of these different factors and thus I think this would be more accurately presented as both of these factors playing a role.*

We agree with this comment – while we think that "interplay" is the most plausible hypothesis at this time, we haven't (yet) carried out experiments to link the two directly. We have therefore clarified the text accordingly.

The structure and presentation of findings is much clearer now. However, the HLA allele association result in particular would benefit from more citations / a couple of sentences clearly laying out the link to CD4+ T cell responses for non-experts.

We have added a reference on the MHC class II and CD4+ cell interaction to the text.

Additionally, on line 369 the use of "demonstrate" is too strong given that the finding is of a statistical association. I suggest "support" or "indicate" as more appropriate alternatives.

Altered as suggested

The authors indicate that control group 2 all tested positive for HAdV, however sequencing reads were only identified in 9/12 samples. Do the authors suspect this is because the samples were plasma / serum (elsewhere it is noted that whole blood would be preferred but was unavailable), as this seems to conflict with the greater sensitivity of TE sequencing compared to PCR in cases? Elaborating on this in the limitations paragraph would be appropriate.

The PCR positivity in the controls came from routine clinical testing of higher yield samples that were not available to us (stool and nasopharyngeal aspirate) – we had serum/plasma from these patients only. This has been clarified in the text. It therefore doesn't conflict with the findings so we haven't added this to the limitations section.

More generally, I suggest modfying the limitations paragraph, which currently focuses more on the HLA typing result than general limitations of the study. In particular, it would be good if the authors could explicitly acknowledge the differences between the case and control cohorts (age, time of collection) as well as differences in detection sensitivities between sample types that could impact the findings.

We feel that the HLA limitation does need to be discussed so have retained it. We have discussed the differences between cases and controls quite carefully – e.g. some controls were sampled earlier than cases (we included contemporaneous controls for this reason) – and this is in the limitations section. We have also discussed detection sensitivities – and have clarified this further elsewhere in the text, as mentioned above.

In general the figures are still not of high quality. Layout and coloring of figure 1 could be modified to improve readability. It would be helpful to see HHV6 on figure 2A.

We have improved the quality of the figures – and if needed, can work with the journal editing team to improve this further if required. We are providing improved pdf figures as requested but can also provide tiff files if better. Unfortunately, there is no room to add HHV6 to figure 2a as suggested – but it is available as supplementary information/extended data.

In figure 3, it is hard to see how 3A and 3C relate to each other – could they be put alongside each other with the same ordering? The coloring is also confusing: I was under the impression the colors reflected whether a sample was part of this study but the first sample from this study (about 7 from the top) appears to have some green and some blue mutations. Lastly, p-values on figure 3C would be helpful.

Figure 3 has now been updated as suggested but has been moved to Extended Data, as recommended above.

There are several grammatical and formatting errors throughout in both text and figures. Please read very carefully prior to resubmission (e.g. ensuring all acronyms are defined, all figure axes have labels).

Corrections have been made, as suggested.