Peer Review File

Manuscript Title: Adeno-associated virus type 2 in US children with acute severe hepatitis

Reviewer Comments & Author Rebuttals

Reviewer Reports on the Initial Version:

Referees' comments:

Referee #1 (Remarks to the Author):

The manuscript by Servellita et al presents an investigation of 16 patients with severe, acute hepatitis of unknown etiology from across the United States. The authors describe several different methods to identify pathogens (qPCR, unbiased and targeted sequencing) and identify at least 4 viruses present in a significantly larger proportion of the cases compared to controls. They further look at 13 AAV2 genomes and identify mutations shared by AAV2 genomes from US and UK acute hepatitis patients including some sites that may impact receptor binding. Overall, the manuscript is clear and well-written, and appropriately references previous literature throughout. The data collected is extensive, in particular with regards to the amount of sequencing performed, and all workflows for sequencing data analysis are clearly described.

My major comments concern the control samples used as a comparison to the cases analyzed here. The controls—a mix of healthy individuals and individuals hospitalized with either an inflammatory or a non-inflammatory condition—in my opinion, do not allow the authors to support some of the major conclusions drawn.

1)As the authors do not compare to other HAdV-positive, acute hepatitis negative individuals, I do not believe they can make the statement that the results suggest that “co-infection with AAV2 may cause a more severe disease than infection by an adenovirus or herpesvirus alone”. Likewise, the final statement of the abstract that “disease pathogenesis and/or severity may be related to co-infection with AAV2” is not well supported by their analysis. Indeed, the authors also find statistically significant associations of EBV and HHV-6 infection with cases compared to controls, both of which they note have previously been implicated in cases of liver failure. While a role for AAV2 does seem likely based on this and other findings, the study design here is not sufficient to independently support this conclusion and these statements should be modified to reflect that or more details should be provided as to why the authors focus on AAV2.

2)The authors also perform all comparisons treating the 45 controls as a single group, despite the fact they are a mix of healthy controls (N=6) and those hospitalized with either an inflammatory (N=8) or a non-inflammatory (N=31) condition. For the instances where a virus was detected in a control sample, in particular, it would be helpful to know if this was in the hospitalized or the healthy subgroup.

3)The authors note that only 23 of 45 controls were collected during the same time frame as the cases. Given the speculation around the importance of the timing for observing multiple infections in

cases it strikes me that the time frame is very important here and only time-matched controls should be included.

4)Given the ways in which these controls are not an optimal comparison set, I do not find it correct to call the controls “matched” in figure 1A. Though they are geographically close to some of the cases they are not more matched than this. To avoid misleading the reader please either remove this word or precise “geographically matched”.

The authors to a good job of synthesizing their findings in the cases from a very complex array of different assays run. While the diversity of assays in some ways supports the findings by serving as replication, the heavily simplified presentation can make it hard to understand how well each of these assays performed relative to each other and in some cases leads to confusion in the final result. More specifically;

1.The analysis of the viral loads for HAdV vs AAV2 is too simplistic and should be improved. It would be important to normalize by genome length. In addition, is there good evidence that this information is reflective of the distribution in the original sample following targeted amplification or enrichment?

2.The reasoning for which analyses (e.g., targeted sequencing vs enrichment by capture) were performed on which sets of samples seems arbitrary and is not well defended. This is made more challenging as figure 1B is very hard to read. I’m wondering if this could be more clearly conveyed through a binary heatmap format or similar with sample sets as the x and analyses as the y, where boxes are colored if an analysis was performed.

3.Figure 1B suggests that of the cases only the 8 AL cases were tested for HHV / CMV and EBV by PCR. How then is it possible that these were detected in 11 and 8 of 13 cases, respectively (page 7 first paragraph)?

4.HHV-6 and EBV were almost entirely missed by metagenomics and viral enrichment (2 samples positive for EBV and 0 for HHV-6). This is surprising and raises questions about what else may be being missed by these methods. Can the authors elaborate on this discrepancy? Were the viral loads especially low for these viruses? What does this mean for interpretation of the relevance of these infections for disease in positive individuals?

5.I may have misunderstood something, but there appeared to be some inconsistency in the presentation of HAdV results. Specifically, in the first paragraph of the section “Detection of viruses in acute, severe hepatitis cases” it is noted that of the 13 cases were HAdV was detected, one of these was HAdV-2. Then in the next paragraph it is stated that “reads from additional viruses were identified including […] HAdV-2”. Are these the same sample? Or an additional sample where this virus was found only by metagenomic sequencing? Please clarify.

6.Figure 3 would be better presented as barplots of the proportion of samples positive for each virus in cases and controls (with sample N above the bar) so that all viruses could be viewed together on the same scale. The current format takes up a lot of space and it is hard to look across all viruses at

once.

7.In the results (page 6, line 7) the authors state “no reads from AAVs other than AAV2 were detected”. This seems inconsistent with their description in methods on page 22 where it is stated “for all AAV genomes, the assembly with the highest breadth of coverage corresponded to the AAV2 reference genome”. Please clarify.

I also had some questions regarding selection of the cases:

1.Why were 12/16 samples selected to be those positive for adenovirus? The authors clearly note this caveat of their dataset on several occasions. However, it would be helpful if numbers could be added of what proportion of all PUIs at these centers were adenovirus positive to understand how strongly this may bias the findings.

2.Can the authors clarify if any of the cases were immunocompromised or had other conditions that may have predisposed to such high numbers of coinfections?

3.Also, please check the reporting of these numbers in the discussion to ensure it is consistent with the results. In the first paragraph of the discussion, it is noted that of 10/13 cases were known to be HAdV-positive and in the second paragraph this is given as 9/13 cases.

Additional miscellaneous comments:

1.The p-values in the abstract do not make sense as no comparison is given, I suggest either removing or rephrasing to state results in cases vs controls and then note significance.

2.In the second paragraph of the introduction (page 3) there is a missing parenthesis.

3.In the extended data table it would be more clear to spell out “tiling” in full.

Referee #2 (Remarks to the Author):

The manuscript by Servellita, Golzalez et al. deals with the current wave of pediatric cases of acute severe hepatitis of unknown cause. The authors performed direct PCR testing, viral enrichment based sequencing and agnostic metagenomic sequencing thus employing sophisticated next-generation tools to unravel a potential viral cause of unexplained acute hepatitis in children. The authors make a novel and striking observation – namely the high rate of 92% of cases with detectable AAV2. Together with two preprints (Refs 12, 33), this study clearly provides an intriguing avenue to better apprehend the unexplained cases of acute hepatitis.

Two findings are particularly interesting; i.) the authors find a distinct AAV2 strain in nearly all cases of acute severe hepatitis, which could be relevant since AAV2, while known to have liver tropism, is not commonly known to cause hepatitis. ii.) in all 13 cases the authors find co-infection with a

human herpesvirus, either EBV or HHV6, as an AAV2 helper virus indicating triple infection in most cases (AAvV, Adenovirus and either EBV or HHV6).

While the results from this study are certainly interesting and potentially relevant, the study disappoints by merely scratching the surface of many aspects. Thus, in the present form, I worry that the authors overstate the relevancy of their results and robust experimental and clinical data are missing to draw meaningful conclusions.

Major points:

1.) The most obvious flaw is the assigned control group (or lack thereof). While certainly important, the regional similarity is second to more obvious aspects that a sufficient control group for acute severe hepatitis would need: an actual liver pathology. Instead, the authors chose an assembly of children that are healthy blood donors, have any kind of infection (osteomyelitis, sepsis) or "noninflammatory condition" (without further details). This is inadequate leaving the results from the control group uninterpretable. What is the positivity rate of AAV2, Adenovirus, EBV, HHV6 with the respective genomics methods in pediatric patients with known (other) cause of hepatitis? With such an inadequate control group, any subsequent statistically comparison is difficult to interpret.

2.) Clinical data provided is either missing or too superficial to understand if the results reported here are laboratory artefacts or true infections. It would be striking of course if indeed all 13 cases with acute severe hepatitis had coinfection with 2-3 viruses at presentation. Is this really the case? Can the authors provide clinical correlation with the genomic data such as clinical symptoms, imaging or ultrasound. The authors need to provide some evidence of actual clinical infection otherwise it seems much more likely that we are dealing with a laboratory artefact.

3.) It is not uncommon in patients with acute severe hepatitis to get a reactive PCR to CMV or EBV e.g. without ever detecting signs of these viruses in liver tissue on subsequent workup. Do the authors have any evidence that indeed the detection of AAV2, HAdV, EBV or HHV6 in blood, nasal swabs or stool is relevant for causing hepatitis?

4.) While liver biopsy is an invasive procedure and especially in children needs to be considered carefully, it is unfortunate that the authors provide 13 cases and only 1 biopsy is available. This is another critical weakness of the study and add to the danger of overinterpreting the results. In addition, no mechanistic data are provided from in vivo, ex vivo or in vitro model systems to understand the potential of AAV2 or co-infection of AAV2 with helper viruses to cause hepatocyte injury or hepatic immune activation.

5.) In the discussion, the authors make a bold claim that social distancing measurement created a vulnerable population young children failing to develop broad immunity to common viral pathogens (P 17). Do we have any data regarding the stay-at-home history of the 13 cases affected? Is this plausible in these kids?

6.) Were any of the 13 cases or controls immunocompromised? Do the authors have sufficient clinical data and patients' history to rule this out?

7.) What is the working hypothesis of the findings in blood / nasal swabs in relation to the severe liver pathology? Direct (triple) liver infection? Immune-mediated liver injury triggered by viral infection? Can the authors provide any further data to have at least a working hypothesis? To investigate this further, a proper immunologic workup, including TCR sequencing, immunofluorescence, immune phenotyping etc. need to be done.

8.) What is the history of SARS-CoV2 in cases and controls? The authors need to provide at least the PCR data (which they have according to fig. 1). Do we have seroprevalence data?

Minor:

9.) It would be interesting to compare the genome of AAV2 in this study with the other recently as preprints presented studies (Ref. 12 and 33). Are the identified coding mutations shared between studies? Could that help explain the newly proposed liver pathology?

Referee #3 (Remarks to the Author):

In this retrospective observational cohort study, Servellita et al. describe an association of Adeno-associated virus type 2 (AAV2) and human Adenovirus (AdV-41) in children with acute severe hepatitis of unknown etiology. They investigated a cohort of 16 pediatric hepatitis cases from the United States, and 45 controls. Whole blood samples were available for 13 cases.
They detected AAV2 in 92% (12/13 samples) of the cases for which whole blood samples were available. Metagenomic sequencing revealed 35 coding mutations in AAV2 genomes recovered from hepatitis cases compared to published AAV2 reference genomes.

These are interesting and potentially important results that corroborate previous findings and might contribute to establish a link between AAV2 (co-)infections and unexplained pediatric hepatitis. The main limitation of the study is its retrospective design, the limited sample size, and the heterogeneity of the cohorts. The inclusion and exclusion criteria for the control cohort are not provided and there is thus potential for bias.

Specific comments:

1. Inclusion and exclusion criteria for both cohorts need to be provided. How were the cohorts matched?

2. It is unclear why n=8 cases (from Alabama) were selected for positive tests for adenovirus. This obviously introduces a selection bias and statistical analyses for associations between cases and HAdV are not valid.

3. A table with baseline demographic characteristics and pre-existing conditions, co-morbidities etc. of both cohorts needs to be provided

4. Figure 1B is rather confusing. Maybe the authors could instead display it as a simple x/y-matrix?

5. The authors claim that the 13 AAV2 genomes from hepatitis cases cluster together in “a distinct AAV2 subgroup”. Judging from Figure 4 the phylogenetic releationship does not seem to be very close. Could the authors comment on the results in more detail?

Author Rebuttals to Initial Comments:

RESPONSE TO REVIEWERS

We thank the reviewers and editor for their constructive comments on our manuscript. Our point-by-point responses to these comments are as follows:

The biggest concern shared by all the referees is the lack of an appropriate control group that would allow you to draw conclusions about the significance of the observed AAV2 infections. Unless you can add additional controls, we think it would be advisable to temper your conclusions about significance and explicitly state (including in the abstract) that the data do not allow you to draw firm conclusions about the role of AAV2.

As suggested by the reviewers, we generated additional data from controls consisting of the following groups: (a) 12 children with adenovirus-positive gastroenteritis (b) 11 children with adenovirus-negative gastroenteritis, (c) 30 hospitalized pediatric patients with acute hepatitis of defined etiology including 2 children with autoimmune hepatitis, (d) 27 hospitalized non-hepatitis patients with non-inflammatory conditions, (e) 15 hospitalized non-hepatitis patients with inflammatory conditions, and (f)18 donor controls.

AAV2 was detected by targeted sequencing in only 4 out of 113 control samples. Of these 4 controls, 2 were adenovirus positive children with acute gastroenteritis, 1 child was adenovirus negative with acute gastroenteritis but had a hospital discharge code for hepatic failure and was also CMV+ by PCR, and 1 child was a donor control who also had HAdV-2 detected by metagenomic sequencing. Among children with acute hepatitis of defined etiology and children hospitalized with either a non-inflammatory and inflammatory condition, no AAV2 was detected. These results further strengthen our findings regarding the observed association of AAV2 with cases of severe pediatric hepatitis during a limited outbreak period. Notably, 69 (61%) of the 113 control samples had been collected during the time frame of the outbreak.

However, we agree with the reviewer that it would be advisable to temper our conclusions and we have done so in the revised abstract.

Referees' comments:

Referee #1 (Remarks to the Author):

The manuscript by Servellita et al presents an investigation of 16 patients with severe, acute hepatitis of unknown etiology from across the United States. The authors describe several different methods to identify pathogens (qPCR, unbiased and targeted sequencing) and identify at least 4 viruses present in a significantly larger proportion of the cases compared to controls. They further look at 13 AAV2 genomes and identify mutations shared by AAV2 genomes from US and UK acute hepatitis patients including some sites that may impact receptor binding. Overall, the manuscript is clear and well-written, and appropriately references previous literature throughout. The data collected is extensive, in particular with regards to the amount of

sequencing performed, and all workflows for sequencing data analysis are clearly described.

My major comments concern the control samples used as a comparison to the cases analyzed here. The controls—a mix of healthy individuals and individuals hospitalized with either an inflammatory or a non-inflammatory condition—in my opinion, do not allow the authors to support some of the major conclusions drawn.

1)As the authors do not compare to other HAdV-positive, acute hepatitis negative individuals, I do not believe they can make the statement that the results suggest that "co-infection with AAV2 may cause a more severe disease than infection by an adenovirus or herpesvirus alone".

For the revision, we generated additional data from a control group consisting of HAdV-positive (with adenovirus gastroenteritis), acute hepatitis-negative children. In this cohort, only 2 out of 12 controls were positive for AAV2 (*P*<0.001). This statistically significant association supports our statement that "...co-infection with AAV2 may cause a more severe disease than infection by an adenovirus or herpesvirus alone...".

Likewise, the final statement of the abstract that "disease pathogenesis and/or severity may be related to co-infection with AAV2" is not well supported by their analysis. Indeed, the authors also find statistically significant associations of EBV and HHV-6 infection with cases compared to controls, both of which they note have previously been implicated in cases of liver failure. While a role for AAV2 does seem likely based on this and other findings, the study design here is not sufficient to independently support this conclusion and these statements should be modified to reflect that or more details should be provided as to why the authors focus on AAV2.

We modified the statements in the abstract to focus on the number of co-infections, including from AAV2, in children with hepatitis as suggested by the reviewer rather than solely on AAV2.

2)The authors also perform all comparisons treating the 45 controls as a single group, despite the fact they are a mix of healthy controls (N=6) and those hospitalized with either an inflammatory (N=8) or a non-inflammatory (N=31) condition. For the instances where a virus was detected in a control sample, in particular, it would be helpful to know if this was in the hospitalized or the healthy subgroup.

We agree with the reviewer and thus, for the revised manuscript, we stratified the controls into 6 non-overlapping groups: (a) donor controls, (b) hospitalized, non-hepatitis patients with inflammatory conditions, (c) hospitalized, non-hepatitis patients with non-inflammatory conditions, (d) patients with acute gastroenteritis and adenovirus PCR positivity in stool, (e) patients with acute gastroenteritis but adenovirus PCR negative in stool, and (f) hospitalized patients with acute hepatitis of defined etiology. The last 3 groups are newly added controls to the revised manuscript. The viruses detected within each group are now clearly shown in the revised Figure 2 of the manuscript.

3)The authors note that only 23 of 45 controls were collected during the same time frame as the cases. Given the speculation around the importance of the timing for observing multiple

infections in cases it strikes me that the time frame is very important here and only time-matched controls should be included.

Of the 68 new controls added, 46 were collected within the same time frame (on or after October 1, 2021). In total, 69 (61%) controls are time-matched. However, we believe that data on all the controls collected should be included as the exact time frame of the outbreak is poorly defined. We also performed a sub-analysis including only the time-matched controls and the association with AAV2 and other viruses is still significant.

4)Given the ways in which these controls are not an optimal comparison set, I do not find it correct to call the controls "matched" in figure 1A. Though they are geographically close to some of the cases they are not more matched than this. To avoid misleading the reader please either remove this word or precise "geographically matched".

We removed the word "matched" from Figure 1A as suggested by the reviewer.

The authors to a good job of synthesizing their findings in the cases from a very complex array of different assays run. While the diversity of assays in some ways supports the findings by serving as replication, the heavily simplified presentation can make it hard to understand how well each of these assays performed relative to each other and in some cases leads to confusion in the final result. More specifically;

1.The analysis of the viral loads for HAdV vs AAV2 is too simplistic and should be improved. It would be important to normalize by genome length. In addition, is there good evidence that this information is reflective of the distribution in the original sample following targeted amplification or enrichment?

The analysis for HAdV vs. AAV2 that the reviewer is referring to is based on number of viral reads detected by metagenomic sequencing and is strictly speaking not a quantitative viral load measurement. However, we agree with the reviewer that normalization is important and have thus normalized the metagenomic data by both genome length (as suggested by the reviewer) and sequencing depth (reads per million) in the revised manuscript.

Note that this information is <u>not</u> necessarily reflective of the distribution in the original sample following targeted amplification or enrichment; thus, for this study we performed agnostic metagenomic sequencing of these samples in parallel. Both methods show that the estimated viral loads (based on viral reads) are likely higher for AAV2 than for HAdV, even after normalization.

2.The reasoning for which analyses (e.g., targeted sequencing vs enrichment by capture) were performed on which sets of samples seems arbitrary and is not well defended. This is made more challenging as figure 1B is very hard to read. I'm wondering if this could be more clearly conveyed through a binary heatmap format or similar with sample sets as the x and analyses as the y, where boxes are colored if an analysis was performed.

We agree with the reviewer and have remade Figure 1B in a binary heatmap format.

3.Figure 1B suggests that of the cases only the 8 AL cases were tested for HHV / CMV and EBV by PCR. How then is it possible that these were detected in 11 and 8 of 13 cases, respectively (page 7 first paragraph)?

Sorry, this was a typo in Figure 1B, and has been corrected. We thank the reviewer for pointing this out.

4.HHV-6 and EBV were almost entirely missed by metagenomics and viral enrichment (2 samples positive for EBV and 0 for HHV-6). This is surprising and raises questions about what else may be being missed by these methods. Can the authors elaborate on this discrepancy? Were the viral loads especially low for these viruses? What does this mean for interpretation of the relevance of these infections for disease in positive individuals?

We do not find largely missing HHV-6 and EBV to be surprising. First, no viral enrichment was performed for HHV-6 and EBV; viral enrichment was only done for AAV2 and Ad41. Second, agnostic metagenomic sequencing is generally less sensitive than targeted PCR for specific viral pathogens, especially in the setting of the high human host background that is typically found in whole blood samples (whole blood samples, which are cellular, yield higher background than cell-free plasma or serum samples, which are the preferred sample matrices for metagenomic sequencing). Third, the viral loads were found to be in the low- to moderate range for these viruses b PCR. Detection of low-level herpes viral DNA in blood can be due to detection of integrated episomal provirus in white blood cells and not actively replicating viral capsids per se. Thus, the relevance of these low-level infections for disease is unclear, suggesting perhaps a most significant role of other viruses (adenoviruses and/or AAV2) in the pathogenesis of the hepatitis.

5.I may have misunderstood something, but there appeared to be some inconsistency in the presentation of HAdV results. Specifically, in the first paragraph of the section "Detection of viruses in acute, severe hepatitis cases" it is noted that of the 13 cases were HAdV was detected, one of these was HAdV-2. Then in the next paragraph it is stated that "reads from additional viruses were identified including […] HAdV-2". Are these the same sample? Or an additional sample where this virus was found only by metagenomic sequencing? Please clarify.

Yes, this is the same sample, and we clarified this in the revised text.

6.Figure 3 would be better presented as barplots of the proportion of samples positive for each virus in cases and controls (with sample N above the bar) so that all viruses could be viewed together on the same scale. The current format takes up a lot of space and is hard to look across all viruses at once.

We agree with the reviewer and now present Figure 3 as bar plots as suggested by the reviewer.

7.In the results (page 6, line 7) the authors state "no reads from AAVs other than AAV2 were

detected". This seems inconsistent with their description in methods on page 22 where it is stated "for all AAV genomes, the assembly with the highest breadth of coverage corresponded to the AAV2 reference genome". Please clarify.

We clarify in the revision that "no reads mapped specifically to AAVs other than AAV2", as some reads from conserved regions mapped to multiple AAV genomes.

I also had some questions regarding selection of the cases:

1.Why were 12/16 samples selected to be those positive for adenovirus? The authors clearly note this caveat of their dataset on several occasions. However, it would be helpful if numbers could be added of what proportion of all PUIs at these centers were adenovirus positive to understand how strongly this may bias the findings.

Unfortunately, for many samples (12 of 16), only residual samples from patients positive for adenovirus were available, as these 12 cases had been investigated by the CDC and the early focus of the investigation was on the association of pediatric hepatitis cases with adenovirus (see Baker, et al., 2022, *MMWR* and Gutierrez-Sanchez, et al., 2022, *NEJM*). The remaining 4 samples were those meeting case criteria and not necessarily positive for adenovirus and were from cases referred to the Californi aDeparmtent of Public Health. The proportion of PUIs that were adenovirus-positive at University of Alabama was 100% (n=8 of the 12 samples selected to be positive for adenovirus), and data from the UK on a similar cohort shows the proportion to be ~90% (Karpen, et al., 2022, NEJM). This information has been added to the manuscript.

2.Can the authors clarify if any of the cases were immunocompromised or had other conditions that may have predisposed to such high numbers of coinfections?

Of the 16 cases, none were immunocompromised. Among the comorbidities identified, asthma is the only condition which may have increased susceptibility to viral respiratory infections such as adenovirus, and only 2 of 16 children had this as a comorbidity. This information is now provided in revised Table 1 and Extended Data Table 1.

3.Also, please check the reporting of these numbers in the discussion to ensure it is consistent with the results. In the first paragraph of the discussion, it is noted that of 10/13 cases were known to be HAdV-positive and in the second paragraph this is given as 9/13 cases.

We corrected this in the revised text.

Additional miscellaneous comments:

1.The p-values in the abstract do not make sense as no comparison is given, I suggest either removing or rephrasing to state results in cases vs controls and then note significance.

We rephrased the abstract to state results in cases vs. controls as suggested by the reviewer.

2.In the second paragraph of the introduction (page 3) there is a missing parenthesis.

We corrected this typo.

3.In the extended data table it would be more clear to spell out"tiling" in full.

We spelled out "tiling" in full as suggested by the reviewer.

Referee #2 (Remarks to the Author):

The manuscript by Servellita, Golzalez et al. deals with the current wave of pediatric cases of acute severe hepatitis of unknown cause. The authors performed direct PCR testing, viral enrichment based sequencing and agnostic metagenomic sequencing thus employing sophisticated next-generation tools to unravel a potential viral cause of unexplained acute hepatitis in children. The authors make a novel and striking observation – namely the high rate of 92% of cases with detectable AAV2. Together with two preprints (Refs 12, 33), this study clearly provides an intriguing avenue to better apprehend the unexplained cases of acute hepatitis.

Two findings are particularly interesting; i.) the authors find a distinct AAV2 strain in nearly all cases of acute severe hepatitis, which could be relevant since AAV2, while known to have liver tropism, is not commonly known to cause hepatitis. ii.) in all 13 cases the authors find co-infection with a human herpesvirus, either EBV or HHV6, as an AAV2 helper virus indicating triple infection in most cases (AAvV, Adenovirus and either EBV or HHV6).

While the results from this study are certainly interesting and potentially relevant, the study disappoints by merely scratching the surface of many aspects. Thus, in the present form, I worry that the authors overstate the relevancy of their results and robust experimental and clinical data are missing to draw meaningful conclusions.

Major points:

1.) The most obvious flaw is the assigned control group (or lack thereof). While certainly important, the regional similarity is second to more obvious aspects that a sufficient control group for acute severe hepatitis would need: an actual liver pathology. Instead, the authors chose an assembly of children that are healthy blood donors, have any kind of infection (osteomyelitis, sepsis) or "noninflammatory condition" (without further details). This is inadequate leaving the results from the control group uninterpretable. What is the positivity rate of AAV2, Adenovirus, EBV, HHV6 with the respective genomics methods in pediatric patients with known (other) cause of hepatitis? With such an inadequate control group, any subsequent statistically comparison is difficult to interpret.

We agree with the reviewer that desirable control groups would be children with adenovirus viremia but no hepatitis (as also suggested by reviewer #1) and children with

severe hepatitis of known (other) cause of hepatitis (as suggested by reviewer #2). In the interim, we have obtained additional control samples from children in these categories. The results of these additional analyses bolster our findings of an association between multiple viral co-infections (including AAV2) and the severe pediatric hepatitis cases. As the total numbers remain low, we have chosen to temper our conclusions in the abstract and manuscript.

2.) Clinical data provided is either missing or too superficial to understand if the results reported here are laboratory artefacts or true infections. It would be striking of course if indeed all 13 cases with acute severe hepatitis had coinfection with 2-3 viruses at presentation. Is this really the case? Can the authors provide clinical correlation with the genomic data such as clinical symptoms, imaging or ultrasound. The authors need to provide some evidence of actual clinical infection otherwise it seems much more likely that we are dealing with a laboratory artefact.

We provide clinical metadata for the patients as suggested by the reviewer. They show that many of the patients had signs and symptoms of viral infection, including fever, gastrointestinal symptoms, elevated WBC (in a subset of patients), and overt signs / symptoms of hepatitis, including jaundice, nausea / vomiting, and anorexia.

3.) It is not uncommon patients with acute severe hepatitis to get a reactive PCR to CMV or EBV e.g. without ever detecting signs of these viruses in liver tissue on subsequent workup. Do the authors have any evidence that indeed the detection of AAV2, HAdV, EBV or HHV6 in blood, nasal swabs or stool is relevant for causing hepatitis?

We agree that herpesvirus detection from CMV, EBV, HHV-6 can be due to reactivation and not causative. However, this is not the case for HAdV, which is a well-known cause of hepatitis. We also note that all 4 viruses were detected in liver biopsy tissue from 1 of 6 patients and adenovirus was detected in liver biopsy tissue from 2 of 6 patients.

4.) While liver biopsy is an invasive procedure and especially in children needs to be considered carefully, it is unfortunate that the authors provide 13 cases and only 1 biopsy is available. This is another critical weakness of the study and add to the danger of overinterpreting the results. In addition, no mechanistic data are provided from in vivo, ex vivo or in vitro model systems to understand the potential of AAV2 or co-infection of AAV2 with helper viruses to cause hepatocyte injury or hepatic immune activation.

We unfortunately did not have access to additional biopsy samples for AAV2 analysis. We do present the results from clinical testing for other viruses (AdV, enterovirus, CMV, HHV-6, EBV, etc.) for 6 biopsy samples in the revised manuscript. In the current study, we focused on the strong association of AAV2 with hepatitis, and the presence of multiple co-infections identified using molecular and sequencing based approaches. Further studies regarding the mechanism of pathogenesis or immune inactivation would be the topic of a future investigation.

5.) In the discussion, the authors make a bold claim that social distancing measurement created

a vulnerable population young children failing to develop broad immunity to common viral pathogens (P 17). Do we have any data regarding the stay-at-home history of the 13 cases affected? Is this plausible in these kids?

We have incomplete data regarding the stay-at-home history for the 16 cases affected. Among the 9 cases for whom history was available, five children never attended school or day care. The other 5 did return to school / day care once these institutions reopened. This information has been added to the discussion.

6.) Were any of the 13 cases or controls immunocompromised? Do the authors have sufficient clinical data and patients' history to rule this out?

None of the 16 cases were immunocompromised. Among the 113 controls, 15 were noted to have an immunocompromised status.

7.) What is the working hypothesis of the findings in blood / nasal swabs in relation to the severe liver pathology? Direct (triple) liver infection? Immune-mediated liver injury triggered by viral infection? Can the authors provide any further data to have at least a working hypothesis? To investigate this further, a proper immunologic workup, including TCR sequencing, immunofluorescence, immune phenotyping etc. need to be done.

We agree that an immunologic workup would be helpful in further understanding possible mechanisms of AAV2 pathogenesis. This is beyond the scope of the current study and would be the topic of a future investigation.

8.) What is the history of SARS-CoV2 in cases and controls? The authors need to provide at least the PCR data (which they have according to fig. 1). Do we have seroprevalence data?

Yes, we have PCR data for SARS-CoV-2 and provide it in the revised manuscript. We do not seroprevalence data.

Minor:

9.) It would be interesting to compare the genome of AAV2 in this study with the other recently as preprints presented studies (Ref. 12 and 33). Are the identified coding mutations shared between studies? Could that help explain the newly proposed liver pathology?

We did compare the genome of AAV2 and identified shared coding mutations among the studies (see Extended Data Table 2). Further investigation would be needed to link the specific mutations with liver pathology such as hepatitis.

Referee #3 (Remarks to the Author):

In this retrospective observational cohort study, Servellita et al. describe an association of Adeno-associated virus type 2 (AAV2) and human Adenovirus (AdV-41) in children with acute severe hepatitis of unknown etiology. They investigated a cohort of 16 pediatric hepatitis cases

from the United States, and 45 controls. Whole blood samples were available for 13 cases. They detected AAV2 in 92% (12/13 samples) of the cases for which whole blood samples were available. Metagenomic sequencing revealed 35 coding mutations in AAV2 genomes recovered from hepatitis cases compared to published AAV2 reference genomes.

These are interesting and potentially important results that corroborate previous findings and might contribute to establish a link between AAV2 (co-)infections and unexplained pediatric hepatitis. The main limitation of the study is its retrospective design, the limited sample size, and the heterogeneity of the cohorts. The inclusion and exclusion criteria for the control cohort are not provided and there is thus potential for bias.

Specific comments:

1. Inclusion and exclusion criteria for both cohorts need to be provided. How were the cohorts matched?

We provide the inclusion and exclusion criteria for the control cohort in the revised manuscript. The hepatitis and control cohorts were not matched but the majority of samples were collected within the time frame of the outbreak and were geographically similar.

2. It is unclear why n=8 cases (from Alabama) were selected for positive tests for adenovirus. This obviously introduces a selection bias and statistical analyses for associations between cases and HAdV are not valid.

We agree and have removed the statistical analyses for association between cases and HAdV from the manuscript.

3. A table with baseline demographic characteristics and pre-existing conditions, co-morbidities etc. of both cohorts needs to be provided

We provide a table with baseline demographic characteristics and pre-existing conditions, co-morbidities, including immunocompromised state.

4. Figure 1B is rather confusing. Maybe the authors could instead display it as a simple x/y-matrix?

We display Figure 1B as a binary heat map, as also suggested by another reviewer.

5. The authors claim that the 13 AAV2 genomes from hepatitis cases cluster together in “a distinct AAV2 subgroup”. Judging from Figure 4 the phylogenetic releationship does not seem to be very close. Could the authors comment on the results in more detail?

We clarify in the revised manuscript that the genomes are not clustering together exclusively but are positioned within a distinct AAV2 subgroup that includes other AAV2 genomes.

Reviewer Reports on the First Revision:

Referees' comments:

Referee #1 (Remarks to the Author):

The reviewers have responded to all of my comments and the manuscript is improved by the revisions. In particular, the addition of many more controls in relevant comparisons greatly strengthens the analysis of presence of AAV2 in cases of severe, acute hepatitis of unknown etiology.

My one remaining concern is the presentation of the infections detected in addition to AAV2 and HAdV in many cases. In the abstract the authors make a fairly strong statement about the role of "multiple viral pathogens in disease severity". However, in the discussion this is presented more cautiously noting the low viral loads for these viruses and the potential for reactivation. While I agree the number of viruses in these cases is striking and intriguing, the allusion to their role in disease severity is not statistically supported by the manuscript and the interpretation of these findings should be consistent.

Besides this, I have just a small number of minor comments / clarifications.

Page 6: is the additional case of HAdV-40 identified by PCR from the "untypeable" sample described on the previous page? If so it would be helpful to specify.

The figures are visually appealing and much improved. Please just confirm all colors are readable by those with various types of color vision deficiencies.

There is no discussion of limitations of the study. This should be added to the discussion.

When the authors state that no viral enrichment was performed for HHV-6 and EBV, does this exclude the probe capture enrichment for the cases from Florida, Illinois, North Carolina and South Dakota? I assume both of these viruses are on the TWIST comprehensive viral research panel.

Referee #2 (Remarks to the Author):

The manuscript has been extensively revised. In light of two other articles, it substantiates the finding that AAV2 is associated to the wave of pediatric acute hepatitis cases. The manuscript extends the relevance of AAV2 to the US, which is an important finding per se. The lack of liver biopsies is and remains a major drawback to understand the nature of the hepato-pathology (e.g. is virus detected in the liver, is AAV2 hepatotoxic in these cases or inducing a cytopathic immune response). Thus, the underlying mechanisms remain completely unclear.

Referee #3 (Remarks to the Author):

The manuscript has been significantly improved. Particularly, the incorporation of additional control cohorts further solidifies the findings. The missing clinical and demographic information has been added and the figures have been substantially improved.
Overall, the authors have adequately addressed the reviewers'comments and revised their manuscript accordingly.
The findings of this study are in agreement with other studies demonstrating an association of pediatric severe hepatitis of unknown etiology and AAV2 infection.
The retrospective design and missing analysis of liver tissue remain a weakness of the study, but the topic is of high interest.

Author Rebuttals to First Revision:

Referee #1 (Remarks to the Author):

The reviewers have responded to all of my comments and the manuscript is improved by the revisions. In particular, the addition of many more controls in relevant comparisons greatly strengthens the analysis of presence of AAV2 in cases of severe, acute hepatitis of unknown etiology.

We thank the reviewer for the positive comments on our revised manuscript.

My one remaining concern is the presentation of the infections detected in addition to AAV2 and HAdV in many cases. In the abstract the authors make a fairly strong statement about the role of “multiple viral pathogens in disease severity”. However, in the discussion this is presented more cautiously noting the low viral loads for these viruses and the potential for reactivation. While I agree the number of viruses in these cases is striking and intriguing, the allusion to their role in disease severity is not statistically supported by the manuscript and the interpretation of these findings should be consistent.

We agree with the reviewer and tempered the statement in the abstract to focus on AAV and “one or more helper viruses” instead of “multiple viral pathogens”.

Besides this, I have just a small number of minor comments / clarifications.

Page 6: is the additional case of HAdV-40 identified by PCR from the “untypeable” sample described on the previous page? If so it would be helpful to specify.

We thank the reviewer for pointing this out. This was actually not an “additional case” of HAdV-40 but was the same case that was described on the previous page. To clarify, we move all of the HAdV and AAV2 PCR information to the earlier paragraph and specify that the “untypeable” sample is the plasma sample from NC_14 in the revised manuscript.

The figures are visually appealing and much improved. Please just confirm all colors are readable by those with various types of color vision deficiencies.

Yes, we were careful to ensure that readers with red-green color blindness would not be affected by the coloring in the figures.

There is no discussion of limitations of the study. This should be added to the discussion.

We added a paragraph discussing limitations of the study.

When the authors state that no viral enrichment was performed for HHV-6 and EBV,

does this exclude the probe capture enrichment for the cases from Florida, Illinois, North Carolina and South Dakota? I assume both of these viruses are on the TWIST comprehensive viral research panel.

Yes, both these viruses are on the TWIST comprehensive viral research panel, and EBV was detected in one of the 4 cases (13_FL). We added a sentence to the manuscript describing this observation: "...Among the 4 cases analyzed using metagenomic sequencing with probe capture viral enrichment (13_FL, 14_NC, 15_IL, and 16_SD), EBV was also detected in the blood sample from 13_FL..."

Referee #2 (Remarks to the Author):

The manuscript has been extensively revised. In light of two other articles, it substantiates the finding that AAV2 is associated to the wave of pediatric acute hepatitis cases. The manuscript extends the relevance of AAV2 to the US, which is an important finding per se. The lack of liver biopsies is and remains a major drawback to understand the nature of the hepato-pathology (e.g. is virus detected in the liver, is AAV2 hepatotoxic in these cases or inducing a cytopathic immune response). Thus, the underlying mechanisms remain completely unclear.

We agree with the reviewer and have added the lack of liver biopsies as a limitation to the study.

Referee #3 (Remarks to the Author):

The manuscript has been significantly improved. Particularly, the incorporation of additional control cohorts further solidifies the findings. The missing clinical and demographic information has been added and the figures have been substantially improved.
Overall, the authors have adequately addressed the reviewers'comments and revised their manuscript accordingly.
The findings of this study are in agreement with other studies demonstrating an association of pediatric severe hepatitis of unknown etiology and AAV2 infection.
The retrospective design and missing analysis of liver tissue remain a weakness of the study, but the topic is of high interest.

We agree with the reviewer and have added the lack of liver biopsies and retrospective study design as limitations to the study.