Author's response to reviews

Title: Comparison of Clinical and Biochemical Markers of Dehydration in Children: A Prospective Double-cohort Trial

Authors:

Ron K Tam (rtam@cheo.on.ca)
Hubert Wong (hwong@rougevalley.ca)
Amy Plint (plint@cheo.on.ca)
Nathalie Lepage (nlepage@cheo.on.ca)
Guido Filler (Guido.Filler@lhsc.on.ca)

Version: 2 Date: 21 April 2014

Author's response to reviews: see over
April 16th, 2014

Catherine Olino
Journal Editorial Office
BioMed Central
BMC Pediatrics

**MS: 5204700171192502**

**Comparison of Clinical and Biochemical Markers of Dehydration in Children: A Prospective Double- cohort Trial**

Authors: Ron K Tam, Hubert Wong, Amy Plint, Nathalie Lepage and Guido Filler

Dear Reviewers,

Thank you for reviewing our manuscript submission and for providing us with your comments. Please find our replies with commentary to the concerns below. For added clarity, we have structured our reply in a point-by-point format, where reviewer comments are first repeated in normal font, followed by our commentary and reply in *italics*, with changes to the original submission relating to that section indicated in *highlighted text*. Subsequent to our modifications in response to your concerns, we hope our revised manuscript now meets criteria required to pursue this study and wish to resubmit.

**Reviewer’s Comments to Author:**

**Reviewer 1**

**Major Compulsory Revisions:**

1. **Title:**
   a. Please amend to include the CDS as it seems to be the study’s primary focus.
      • *The title was changed to include the highlighted section concerning the CDS:*
        
        **Comparison of Clinical and Biochemical Markers of Dehydration with the Clinical Dehydration Scale in Children: A Prospective Double-cohort Trial**

2. **Abstract:**
   b. The methods section does not define what is meant by “dehydration.” This is an important clinical term to this study as it is used for both the inclusion criteria and the outcome measure.
      • *The following sentence was added to the methods section of the abstract:*
        
        **The clinical definition for dehydration is the loss of body water, with or without salt, at a rate greater than the body can replace it; it is diagnosed through laboratory testing and clinical assessment (according to the Canadian Paediatric Society, Figure 3).**
c. The conclusion is too much of a jump. Bicarbonate is not the gold standard for dehydration even though this study found the highest correlation with dehydration. The conclusion also should not present new data not included in the results.

- The conclusion section of the abstract was re-written to reflect this comment. The phrase was re-written from the following:
  This study validates the use of the CDS as a dehydration marker in children.
  To the following:
  Although serum bicarbonate is not the gold standard for dehydration, this study provides further evidence for the usefulness of the CDS as a dehydration marker in children.

- No new data was included in the conclusion of the abstract—this data can be found in the second last paragraph of the results section.

3. Background:

a. Please provide more information about the three prospective studies which have validated the CDS including number of patients, age groups, year of publication, etc. This is essential to proving the need for this current study.

- The following text was added to provide information about the 3 preceding studies, highlighting the need for additional studies to validate the CDS scale:
  Following the development of the scale in 2004, Goldman et al., the originators of the scale, were first to attempt to validate the scale in a paper published in 2008.[7] Their prospective observational study consisted of 205 children between 1 month and 5 years of age with suspected acute gastroenteritis. Since the original scale was developed using children 1-36 months of age, the aim of this study was to test this scale in a new cohort of children. Although the investigators found the dehydration categories of the scale to have a statistically significant correlation with length of stay (LOS) from time of arrival in triage and intravenous (i.v.) fluid rehydration, this study had numerous limitations: (i) it was only conducted in one center; (ii) it had a small number of children with moderate/severe dehydration; (iii) using LOS as an endpoint is questionable because LOS is multifactorial; (iv) staff may have changed their practices because of the study (Hawthorne effect), and, most importantly; (v) only a small number of the study population had blood tests performed, so the team could not validate their hypothesis that the dehydration categories positively correlated with abnormal serum pH values or bicarbonate levels (a primary outcome of the study). They indicated that future research is needed to provide information on this hypothesis.

A second study attempting to validate the CDS in a different emergency department was published in June of 2010.[9] With 150 patients from 1 month to 5 years of age diagnosed with gastroenteritis, enteritis, or gastritis, the primary outcome of this study was LOS after being seen by the attending physician and the perceived need for IV fluid administration. Although serum bicarbonate and CO2 were measured, this was one of several secondary outcomes. This study found the CDS to be statistically significantly correlated with LOS from seeing the physician, perceived need for IV rehydration, and utilization of laboratory blood tests. Measured serum bicarbonate and CO2 were not found to significantly vary between the categories. Once again, this study had multiple limitations, the most important being that LOS may be influenced by multiple factors, and although this was measured from the time the patient saw the physician, confounding factors may have still played a role.
Last, Gravel et al. [10] performed a multicenter validation of the CDS, published a few months later in October 2010. 264 children between the ages of 1 month and 5 years were recruited at 3 Canadian centers, presenting for acute vomiting and/or diarrhea. The primary outcome of this study was the percentage of dehydration (difference in weight), while secondary outcomes included proportion of blood test measurements, IV use, hospitalization, and inter-rater agreement. This study found a statistically significant correlation between the CDS and percent dehydration (by weight), number of blood test measurements, IV rehydration use, hospitalization, and abnormal plasma bicarbonate. This study was limited in that it did not exclusively include patients with a gastroenteritis diagnosis, though a subgroup analysis was performed producing similar results, and the primary outcome could not be measured in 45 (17%) of patients. Finally, the use of percent dehydration is limited by certain confounders.

- The following highlighted sections were also added to the paragraph that follows: Although these studies have further validated this measure of dehydration, the primary outcome has differed in each study and all possess limitations (particularly LOS), none have employed the use of a reliable control group, nor have they included a wide array of surrogate markers. The limitations of the preceding studies suggests the need for additional tests of validity for the CDS using other clinical markers.

b. What is meant by “reliable” control group? This should be expounded upon as it provides the central argument for the importance of this study.

- The following text (highlighted) was added to explain what is meant by “reliable control group”:
  
  ...none have employed the use of a reliable control group (all 3 studies used a CDS score of 0 – “no dehydration” for baseline measurements rather than a separate, non-dehydrated group)

c. The primary study hypothesis does not clearly state why there was a control group. It is unclear why a control group without dehydration was used if the study is primarily designed to investigate the hypothesis “that the CDS would correlate with standard biochemical markers for dehydration.”

  i. It is implied the control group is being used to validate the biomarkers and then the biomarkers are being used to validate the CDS. Is this the study’s design?

  - Yes, this is the reason why a control group was used – to first provide baseline values for the established and novel markers of dehydration. Comparing the biomarkers in both a control group and a dehydration group allows for a comparator and further ensures the accuracy of the results through statistical comparison.

  The following text was added to clarify this point in the background section:

  Measuring the biomarkers in a control group provided baseline values and allowed us to validate the biomarkers in a healthy population prior to validating them in the dehydration cohort.

4. Materials and Methods:

a. The choice of cases and controls needs more explanation:

  i. The choice of a trauma “control” group could affect a few of the dehydration markers being studied, specifically the heart rate and some urinary markers especially if the patient had any unidentified renal or head trauma which could impair sodium handling, acid base balance, cause tubular damage or lessen
GFR. Were there any additional exclusion or inclusion criteria for the controls? We completely understand this concern. We agree with this, and apologize for failing to delineate it more clearly that patients with abdominal trauma or head trauma were excluded in the control group for precisely the very reason you cite. We essentially only used patients with arm or leg injuries.

ii. Cases were identified as those who were found to have a gastroenteritis and "dehydration." Dehydration has to be more clearly defined. How exactly was this determined by the screening research nurse?

We used the Clinical Dehydration Scale for Children with Acute Gastroenteritis. There are 4 characteristics that will get a score of 0, 1 or 2. This is a robust and clear scale. Please see below:

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Score of 0</th>
<th>Score of 1</th>
<th>Score of 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>General appearance</td>
<td>Normal</td>
<td>Thirsty, restless, or lethargic, but irritable when touched</td>
<td>Drowsy, limp, cold, or sweaty; comatose or not</td>
</tr>
<tr>
<td>Eyes</td>
<td>Normal</td>
<td>Slightly sunken</td>
<td>Very sunken</td>
</tr>
<tr>
<td>Mucous membranes</td>
<td>Moist</td>
<td>Sticky</td>
<td>Dry</td>
</tr>
<tr>
<td>(tongue)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tears</td>
<td>Tears</td>
<td>Decreased tears</td>
<td>Absent tears</td>
</tr>
</tbody>
</table>

We inserted the scale as table 3. The CDS consists of four clinical characteristics (general appearance, eyes, mucous membranes, and tears), each of which are scored 0, 1, or 2 for a total score of 0 to 8, with 0 representing no dehydration; 1 to 4, some dehydration; and 5 to 8, moderate/severe dehydration. This score has been validated externally and is robust. We used exactly the same criteria as Benoit Bailey et al., Academic Emergency Medicine, 2010:17(6):583-88.

iii. Was there validation of the screening nurse’s determination that the cases were indeed dehydrated? What criteria did she use? This is essential to ensure there was no case misclassification.

The MRPs responsible for the scoring attended a training session on the CDS before beginning the study and was well trained in the emergency department with long-term experience, including training in triage, and a formal evaluation by a specialized nurse educator. The MRP only had the score sheet, and as outlined in our text, the MRP subsequently was asked to assess the degree of dehydration based on her or his own experience. These assessments were then obtained by the dedicated research nurse, and independently recorded by a dedicated research nurse. This process avoided any manipulation of the assessment. We feel we clearly documented that in the text and did not make any additional changes.

b. Need to discuss why inter and intra-observer error was not assessed in the CDS scoring in a study about a subjective rating scale.

In some instances (approximately 10%), the study nurse also assessed the MRP ratings, and there were no discrepancies, not even in a single case, which made us confident that we did not have to formally assess the inter and intra-observer error. In retrospect, we should have documented this, as it would have sufficed to address this concern.

We changed the text as follows:
Inter- and intra-observer error was not assessed as there were no discrepancies between the rater assessments and independent assessments of the dedicated research nurse. We also added a sentence to the limitations section: We also did not formally assess the inter- and intra-observer error for the CDS score.

c. Please outline all biomarkers used as surrogate markers of dehydration with references supporting their use as “established biochemical surrogate markers of dehydration.”

We measured serum and urine electrolytes (sodium, potassium, chloride), blood urea, serum bicarbonate and creatinine, osmolality, urine alpha-1 microglobulin and urine microalbumin, albumin, cystatin C. With these, we calculated all of the parameters in table 2. Of these, fractional sodium excretion (FeNa) is the most commonly used marker of acute kidney injury to differentiate prerenal from intrinsic renal failure. Fractional excretion of urea was one of the study questions. The revised table 2 now reads:

Table 2. Student’s t-test of established and potential markers of dehydration and acute kidney injury with key references listed.

<table>
<thead>
<tr>
<th>Markers</th>
<th>Dehydrated</th>
<th>Control</th>
<th>AUC (SE)</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n=73</td>
<td>n=143</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Urine Na/K ratio [26]</strong></td>
<td>0.69 (0.64)</td>
<td>2.3 (0.56)</td>
<td>0.789(0.03)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td><strong>Serum Osmolality [mOsm] [27]</strong></td>
<td>286 (231-684)</td>
<td>289 (199-391)</td>
<td>0.11</td>
<td></td>
</tr>
<tr>
<td><strong>Urine Na [mmol/L] [28]</strong></td>
<td>66.0 ± 55.6</td>
<td>144 ± 74</td>
<td>0.798(0.03)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td><strong>FeNa [29]</strong></td>
<td>0.19 (0.089)</td>
<td>0.52 (0.104)</td>
<td>0.753(0.03)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td><strong>FeUrea [30]</strong></td>
<td>0.398 (0.678-0.96)</td>
<td>0.377 (0.055-4.504)</td>
<td>0.66</td>
<td></td>
</tr>
<tr>
<td><strong>Serum bicarbonate [mmol/L]</strong></td>
<td>20 (10-27)</td>
<td>24 (18-30)</td>
<td>0.821(0.03)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td><strong>Blood urea [g/L]</strong></td>
<td>5.7 (1.2-41)</td>
<td>5.2 (2.9-2.9)</td>
<td>0.613(0.05)</td>
<td>0.01</td>
</tr>
<tr>
<td><strong>Serum creatinine [µmol/L]</strong></td>
<td>38 (16-408)</td>
<td>45 (3.8-9.7)</td>
<td>0.666(0.04)</td>
<td>0</td>
</tr>
<tr>
<td><strong>Schwartz eGFR [mL/min/1.73 m²]</strong></td>
<td>131.6±32.3</td>
<td>136.9±23.2</td>
<td>0.18</td>
<td></td>
</tr>
<tr>
<td><strong>Cystatin C [mg/L]</strong></td>
<td>0.67 (0.43-2.89)</td>
<td>0.66 (0.44-1.08)</td>
<td>0.58</td>
<td></td>
</tr>
<tr>
<td><strong>Cystatin C eGFR [mL/min/1.73 m²]</strong></td>
<td>144.0±36.9</td>
<td>147.6±5.1</td>
<td>0.41</td>
<td></td>
</tr>
<tr>
<td><strong>Urine Osmolality [mOsm]</strong></td>
<td>805 ± 306</td>
<td>876 ± 402</td>
<td>0.2</td>
<td></td>
</tr>
<tr>
<td><strong>Urine microalbumin/creatinine ratio [mg/mmol] [31]</strong></td>
<td>4.4 (0.4-61.1)</td>
<td>2.3 (0.3-9.4)</td>
<td>0.69 (0.04)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td><strong>Urinary α1-microglobulin/creatinine ratio [mg/mmol] [31]</strong></td>
<td>1.75 (0.30-14.70)</td>
<td>0.70 (0.20-11.30)</td>
<td>0.809 (0.04)</td>
<td>P&lt;0.001</td>
</tr>
</tbody>
</table>

d. When performing ROC analysis, one should be specific about the binary outcome of interest. I assume the outcome being tested is the presence of absence of
dehydration per initial screening but this should be explicitly stated. We thank the reviewer for pointing out our omission. The sentence now reads: All statistically significant markers were then compared with receiver operator curves (ROC) to determine the marker with the highest sensitivity and specificity for the binary outcome of the presence or absence of dehydration as per the initial screening.

e. Please specify if you performed spearman or Pearson correlation analyses. We did not include a correlation analysis as the scores from the dehydration scale are integers and cannot assume any numerical value. However, we can perform this should the reviewer feel it is important.

5. Results:
   a. Please be clearer about what the ROC analysis was testing predictability for. It is standard to also include the 95% CI and the p-value for every ROC curve. If you are comparing the groups, then this analysis is complicated by the fact that there was not a difference in percent dehydration between those with dehydration and your controls. The text now reads: We performed ROC analysis to compare sensitivity and 1-specificity between both groups. The binary outcome of interest for the ROC analysis was the presence of absence of dehydration per initial screening.

6. Why was no correlation analysis performed on the CDS and percent dehydration, which is what the scale was designed to predict? Please comment on the urine-production status of the patient with HUS as some readers may assume that patients needing dialysis may be oliguric or anuric and, therefore, not clinically dehydrated at all, but more fluid overloaded. We already addressed this in the methods section. The CDS scores can be used as categorical scores, but they are integers and cannot assume any numerical value. If the reviewer insists, we can perform ANOVA, or apply a correlation analysis with the integers. To satisfy the reviewer, we did perform a non-linear regression analysis using the CDS scores as the x-axis for urinary alpha1 microglobulin/creatinine ratio (exponential growth was best model). The new figures was added below. We elected against including this in the manuscript. We are unsure what the reviewer means with regards to percent dehydration? The estimated fluid loss versus the CDS score? That seems unfruitful in view
of the limitations discussed.

To satisfy the reviewer, we also performed a linear regression analysis between the subjective dehydration score and the CDS score. This was statistically significant (R square 0.3130, p<0.0001). For the reasons discussed, we elected against the inclusion of this information in the manuscript, although we realize that this would strengthen the paper. We thank the reviewer for the suggestion.
7. Discussion:
   a. There was no statistical difference between percent dehydration in the dehydrated patients when compared to controls. This raises significant questions as to the validity of this study, i.e. were the cases and controls truly different enough to validate a scale of dehydration?
   b. The difference in blood urea which is statistically significant is not clinically significant and should not warrant much comment.
   c. If the author’s statement that the CDS was validated against serum bicarbonate in its inception, then the authors cannot use serum bicarbonate to indirectly validate the CDS as seen in the second paragraph on page 12.
   d. The paragraph discussing the potential for selection bias is important but muddled. In this manuscript, there is insufficient support for the statement: “First, our selection criteria biased our dehydration group to children with more severe disease.” This is also contradicted by the authors’ later comments that parental hydration is started more often than oral despite oral hydration’s superiority. Please clarify if the authors argue that patients with dehydration who receive intravenous fluids are more severely ill or not.
   e. If the CDS is not validated for children over 5 years of age as outlined in the top paragraph on page 13, why were older children included? The control group’s median age was 8 years old. The dehydrated patients were within the appropriate age group. If CDS was used to prove the controls were not dehydrated; however, the scale was not intended for children over 5, how can the authors reliably ensure the controls were indeed hydrated?
   f. Please comment on the age difference as a potential confounder for the difference in blood urea and serum creatinine concentrations seen between the cases and controls. Heart rate could also be influenced by other factors in younger rather than older children who are more anxious in a health care setting.
We agree with the reviewer. Creatinine is clearly influenced by height which differs by age. This is why we calculated age-independent Schwartz eGFR. Urea should not be affected very much by age, but heart rate clearly does. To include this possible confounder, the following text was added to the limitations section of the discussion:

The current study has several limitations, including the difference in age between the dehydrated group and the control group. This may have influenced the difference in serum creatinine concentrations seen between the two groups, although we corrected for this by using the Schwartz formula for the estimation of eGFR per body surface area. Heart rate may also differ by age.

g. The use of percent dehydration in this study meets many challenges:

i. The authors have to address the fact that post-hydration weights were only available on 60% of the dehydrated group and this weight was needed to calculate the true percent dehydration which would be the only non-subjective outcome for hydration used in this study.

ii. Why didn’t the pre and post-hydration weights change in the patients with dehydration if the HR, systolic and diastolic z-scores declined? In fact, weight z-scores declined in both groups implying patients lost weight, as commented on in the discussion. These things are incongruous with physiology: why would the systolic blood pressure and weight go down after hydration? One plausible reason may be that the patients did not have intravascular dehydration upon presentation.

iii. Why would trauma patients lose weight after treatment? This calls into question the validity of any of the weights done in this study.

iv. The discussion on page 13 which concludes that percent dehydration is a “subjective parameter” is incorrect. The parameter is objective as it relies only on actual measured data. It may not reflect hydration status as the rest of the paragraph describes, but the parameter itself not subjective. The other variables used in this study such as the CDS score and MRPs assessment of hydration are subjective and therefore need validation with inter-rater reliability testing.

v. This discussion has to include the possibility that the dehydrated patients were not all truly dehydrated as no gold standard was used to classify the cases and the controls. Or that the controls were not all hydrated. The urinary osmolality of the controls speaks against being well hydrated.

Not included in the Discussion but needs to be:

h. Please explain the incongruity inherent in the concept that a CDS score of 0-2 was considered “mild” dehydration for the cases however the argument was made that a CDS of 0 for the controls proves they were not dehydrated. Even if every patient in the control group scored a 0 on the CDS, there were patients in the dehydration group whom also scored a 0 and were considered “mildly” dehydrated.

Minor Essential Revisions:
1. Abstract:
   a. The background implies that the CDS is a well-known tool. Needs more explanation to those who are not familiar with it.

Comment [MK1]: I highlighted this section in the manuscript (end of page 13). The sentence the reviewer is referring to is as follows: "As discussed in their report, both length of stay and the need for intravenous rehydration are subjective parameters influenced by a number of factors including the severity of the patient’s illness." I may be misunderstanding something, but it seems that the manuscript is referring to LOS and the need for IV rehydration, not % dehydration.
• The following text was added to the beginning of the background section in the abstract to serve as an introduction to the CDS:

The clinical dehydration scale (CDS) is a quick, easy-to-use tool with 4 clinical items and a score of 1-8 that serves to classify dehydration in children with gastroenteritis as no, some or moderate/severe dehydration.

b. The statistics are incorrectly described and presented in the methods and results sections:
   i. Is that the median with the total range or interquartile 1-3 range?
   ii. P-values are necessary to describe differences between the groups in terms of age and CDS.
   iii. The sentence starting with “Difference in heart rate” needs to be more specific about what you are comparing. It is assumed that it’s the control and the dehydrated groups, but it has to be expressed. This sentence is also grammatically incorrect (Difference… were).

• The sentence was changed to reflect this comment:

The following parameters were statistically significant (p<0.05) between the control group and the dehydrated group: difference in heart rate, diastolic blood pressure, urine sodium/potassium ratio, urine sodium, fractional sodium excretion, serum bicarbonate, and urea and creatinine measurements.

iv. When presenting a correlation, the correlation coefficient not just the p value should be shown.

2. Background:
   a. The sentence starting with “It is also the most common cause” needs to state in which population this is true.

• The following text was added to elucidate the population in which this is true:

It is also the most common cause of electrolyte abnormalities in children presenting at the ED.

b. Need a reference for the second half of the sentence starting with “Although previous studies…”

3. Materials and Methods:
   a. Technically this is a case control study since patients were chosen based on their disease status.

• As per the Editor’s comments, in the opening sentence of the materials and methods section, the phrase “prospective double cohort control trial” was changed to “case comparison trial” (this change was also made to the title).

b. Please specify what ages were included.

• All eligible pediatric patients (>18 years old) were included in the study – as such, this text was added to the materials and methods section.

c. Why did you not exclude patients with a sodium derangement?

d. Please specify when the attending physician conducted the CDS scoring, was it always before any fluid was administered?

• Yes. The following text was added to reflect this change:

The attending MRP conducted scoring for the CDS prior to the administration of any fluid.

4. Results:
   a. Please remove “experimental” from the sentence staring with “seventy-three” as this implies an intervention was performed.

• The word “experimental” was removed from this sentence.

b. Please insert p-values after each comparison offered in the text.
All p-values were added within the text of the results, as follows:

- Urine Na/K ratio (p<0.0001), urine Na (p<0.0001), FeNa (p=0.0001), blood urea (p=0.01), and serum bicarbonate (p<0.0001) and creatinine (p=0) were all significantly different between both groups (Table 2). Serum cystatin C (p=0.58), % dehydration by body weight (p=0.61), FeUrea (p=0.66), urine osmolality (p=0.2), and serum osmolality (p=0.11) did not reach statistical significance. Both the urinary microalbumin (p<0.0001) and urinary alpha-1 microglobulin (p<0.001) reached a high statistical significance level.

5. Discussion:
   a. Cannot conclude “Friedman’s CDS scoring system is a useful clinical tool.” This is anecdotal and was the entire point of the study. Would remove this sentence or rephrase.
      - This sentence was removed.
   a. Please remove “anecdotal” from the 4th sentence of the second paragraph on page 12. It is more appropriate to simply state if an association existed or not.
      - The word “anecdotal” was removed from this sentence.
   b. Also, the authors should discuss what is the clinical utility of a tool to determine dehydration if it does not lead to important morbid outcomes like LOS, admission, etc.
      - Previous studies validating the CDS tool (discussed within the Background section) have shown that the scale is valuable in predicting LOS, the need for intravenous fluid rehydration, weight gain, the occurrence of blood tests, hospitalization, and abnormal plasma bicarbonate (a significant correlation with LOS was found in 2 of these studies). Although these are valuable indicators of the clinical utility of this tool, clinical endpoints such as LOS are limited by the possibility of numerous confounders. Our study sought to use measurable physiological parameters, which may be less subjective and can be a good indicator of dehydration, to evaluate the accuracy of the scale.
      - The following text was added to the Discussion section of the manuscript to explain the clinical utility of this tool:

Previous studies validating the CDS tool [7, 9, 10] have shown that the scale is valuable in predicting LOS, the need for intravenous fluid rehydration, weight gain, the occurrence of blood tests, hospitalization, and abnormal plasma bicarbonate (a significant correlation with LOS was found in 2 of these studies). These results imply that this tool is clinically useful, and this notion is strengthened by the results of the current study which used endpoints that are not limited by the same confounders and provide measurable physiological parameters.

c. What is meant by “the likelihood of complications” as there is neither intervention from the study nor any specific therapy which would cause complications which would impact results?
   - The following text was added to clarify this point:

we anticipated greater differences between the dehydration and control groups and an increase in the likelihood of complications associated with dehydration
Please explain what the clinical significance is of the urinary microalbumin/creatinine ratio and the urinary α1-microglobulin/creatinine ratio.

6. Tables/Figures:
   a. Figure 1: Initial number in the top box, 29, must be a typo. The text lists that as “229” patients.
      • This typo was corrected.
   b. Table 1:
      i. Please include the n for each group in the header.
      • This change was made in both tables.
      ii. Are you presenting the total range or interquartile ranges with the median values?
      iii. Statistical analysis of the weight, systolic, diastolic and HR changes pre and post-treatment inside of the groups is important as well to prove that the dehydrated patients were adequately hydrated and the control patients were not dehydrated. The results presented in the text should be somehow included on this table.
   c. Table 2:
      i. Please include the normal values for these markers, the control group cannot be assumed to be “normal.”
      ii. The sentence starting with “Data from the control group” is not essential to that table nor is it phrased in a way that implies understanding of ROC curves and should be removed.
      • This sentence was removed.

Discretionary Revisions:
1. Abstract:
   a. Would mention which direction the differences are in heart rate, diastolic BP, etc. between groups.
      • The following (highlighted) text was added to the abstract to reflect these changes:
      The following parameters were statistically significant (p<0.05) between the comparison group and the dehydrated group: difference in heart rate and diastolic blood pressure (both were significantly higher in the dehydration group pre-treatment), urine sodium/potassium ratio, urine sodium, fractional sodium excretion, serum bicarbonate (all four were significantly higher in the comparator group) and urea and creatinine (both were significantly higher in the dehydrated group) measurements.

2. Background:
   a. The sentence beginning with “When assessing a patient” is too wordy and awkward. It should be reworded.
      • This sentence was re-worded as follows:
      Clinicians must determine whether patients only need to be rehydrated or whether they face more substantial morbidity, which can be challenging.
   b. Table to illustrate how to calculate the CDS would be extremely helpful for others to use this tool if they are convinced by this article.
      • The CDS was inserted as Table 3.

3. Materials and Methods:
   a. The sentence starting with “Results in this current article” is worded poorly and should be rewritten for clarity.
This sentence was re-phrased as follows:

Data used for this study was originally collected during a trial devised to examine the role of cystatin C as a biomarker of renal dysfunction in children with dehydration. Results were obtained from a secondary analysis of this data.

b. Would include how many MRPs were included in the study to give the readers a sense of how many different opinions would be forming the subjective hydration status and be performing the CDS.

4. Results:
a. Presenting the height data does not add anything to the results.
   • The height data was removed.
b. Giving actual heart rates and blood pressure values would be helpful for clinical utility and translation to physician readers. The Z-score is excellent but not as helpful to the reader immediately.
c. Remove “Apart from comparing cohorts” from the first sentence of the second paragraph on page 10. Too anecdotal.
   • This part of the sentence was removed.
d. Performing the ROC curve analysis to determine the predictability of serum bicarbonate and CDS for each strata of dehydration would strengthen the study’s findings.
e. The amount of fluid resuscitation provided to the cases and controls would be helpful information. In fact, looking at the IVF boluses given to each strata of dehydration would strengthen the argument that these patients were indeed more dehydrated.
f. It could also potentially strengthen the results to re-analyze the data excluding any patients who are missing important data as mentioned at the bottom of page 6.

5. Discussion:
a. The last few sentences in the final paragraph of the discussion feel “tacked on” and do not add to the strength of the manuscript.
   • The last paragraph of the discussion was removed.

Reviewer 2

Major Compulsory Revisions:
1. In Materials and Methods, paragraph 1: the study is called a Prospective Double Cohort Control Trial. In fact, this is actually a Prospective Double Cohort Trial. There are two separate cohorts, one exposed to dehydration and one not exposed to dehydration. A true control for the gastroenteritis group with dehydration would be a gastroenteritis group without dehydration (same population without the exposure).
   • As per the Editor’s comments, this is a case comparison study, and all terminology throughout the paper has been modified to reflect this. For example, rather than a “control group”, the hydrated patients are now referred to as a “comparison group”.

2. Materials and Methods, paragraph 4: In what time frame were the lab data collected (on presentation prior to IV rehydration)?

3. Materials and Methods, under Calculations and Statistical Analyses, paragraph 1: the percent dehydration calculation is actually the fraction dehydration calculation. To be a percent, it should say “x 100”
“x 100” was added to the formula so that it is truly percent dehydration.

4. Results, paragraph 1: this says there are 73 children in the experimental group and 143 children in the "control" group (see above) from a total of 215 patients. However, 73+143 = 216. This error in calculation is also in Figure 1.
   - The correct total number of patients is 216. This error in calculation was corrected throughout the manuscript.

5. Discussion, paragraph 3: “we included a control group to strengthen our analysis” should be modified per #1 above to say that "we included a hydrated cohort to strengthen our analysis"
   - This change was made.

6. Discussion, paragraph 3: This should include a discussion of how the results could be affected by the differing age and gender of the two cohorts.
   - The influence of age was mentioned by Reviewer 1 and the manuscript was modified accordingly (Major Compulsory Revisions, 6f).
   - A section was added to the discussion regarding the differing gender proportions of the two groups and the ways this could have affected the results.
   - This change was made.

7. Discussion, paragraph 5: “…we believe employing the use of an independent control group…” per #1 above should read “…use of a hydrated cohort as a control group…”
   - This change was made.

8. Discussion, paragraph 7: Are the children with dehydration but not gastroenteritis included in the musculoskeletal injury cohort? This further supports this study as a Double Cohort Study, not a Double Cohort Control Study. It is not controlled if both cohorts have potential to be exposed to the exposure of interest, dehydration. If not, this paragraph needs to be clarified. Also, it would help to explain why this is a "clinically relevant population." This paragraph is probably not necessary.

Minor Essential Revisions:
1. Materials and Methods, under Calculations and Statistical Analyses, paragraph 1:
   - “Comparisons were first made between the dehydration and the control group” - because this is a Double Cohort study without a true control, this should read “comparisons were first made between cohorts”
   - This change was made.

2. Discussion, paragraph 1: reads “Vega et al have demonstrated that in addition to serum urea, serum bicarbonate declines...”. Should this read, “Vega et al have demonstrated that in addition to serum urea increasing, serum bicarbonate declines...”?  
   - This correction was made. The sentence now reads:  
     Vega et al. have demonstrated that in addition to serum urea increasing, serum bicarbonate declines with increasing percentage of lost body weight.

3. Discussion, paragraph 5: “Findings based on data from these 77 patients were than...” should say “Findings based on data from these 77 patients were then...”
   - This correction was made.

Editor’s Comments

Comment [MK2]: Age was addressed above but the difference in gender between the 2 cohorts was not.
Overall, there is content here which would be useful to readers, given some major revisions. Please address all the recommendations of the two reviewers. In particular:

- There is no reason to describe this study as a double cohort study. In truth, the period of longitudinal follow-up during the treatment phase is fraught with biases that do not help. I would call this a case-comparison study (not case-control). Essentially, as Reviewer 1 notes, you have two groups of patients: dehydrated and Not dehydrated, and you are comparing them for lab values. (you are not taking two groups of patients with the same disease and providing alternative exposures/treatments over time). I would not call the second group a "control" as both reviewers point out this is not a true "Control" group -- some of the patients in that group may be dehydrated. Since you have no gold standard definition of dehydration, then you can only refer to that group as a legitimate comparator. We agree, and this resolves the discrepancies between reviewer 1 and reviewer 2. Consequently, the study has now been labeled as a Case Comparison Trial. Also, the term case comparison trial was used consistently throughout the manuscript.

- Please use confidence intervals instead of standard errors when reporting the AUC data. You need to decide what your gold standard is. It seems that bicarbonate would be a logical choice given the previous literature. As reviewer 1 points out, you cannot set out to validate the CDS score by comparing it to bicarb, and then perform AUC curves to test the sensitivity of the bicarb value for picking up a CDS score consistent with dehydration.

- I do not see any value in using the % dehydration calculation based on weight gained at end of treatment. Too much can impact how the weight changes. (one can overhydrate, for example). In practice, % dehydration is estimated BEFORE treatment, based on previously documented weight. Also, you are missing post treatment weights on some children, so again, this measure is not particularly useful.

- Similarly, I am not clear on the reason that the physicians rated the level of dehydration subjectively. I may have missed the results. If you are using this information to state that the CDS score is more sensitive and specific than the clinicians’ subjective rating of dehydration, this would be important to demonstrate in the results. In the end, I am not convinced that a CDS of 2 or more is better than my subjective rating of dehydration. Please address in the discussion.

- In the introduction, the dehydration group is described as having gastroenteritis; however later you describe children with other disorders, including hemolytic uremic syndrome. Since other disorders can have more profound effects on bicarbonate, I would exclude those subjects from the analyses.

- Please set up the hypothesis a priori that you are intending to validate the dehydration score by testing whether it correlates with certain factors, including bicarb, sodium, etc. Then in your results, you simply say that as you hypothesized, the score did correlate with these factors.

- If the CDS has been validated previously with interobserver reliability scores, please describe these. If not, please note as a limitation in the study. Please address Reviewer 1's concerns in the limitations section of the discussion (validity of using the CDS in older kids,
that the comparison group may be dehydrated, etc.). Please make the corrections to the sample size pointed out by reviewer 2.

Sincerely,

Guido Filler, MD, PhD, FRCPC
Chair/Chief, Department of Paediatrics
Professor of Paediatrics