Office of Evidence Based Practice (EBP) – Critically Appraised Topic: Clinical Assessment Tools for Dehydration in Acute Gastroenteritis

Specific Care Question

For patients who present to the emergency department (ED) with acute gastroenteritis (AGE) do either the Clinical Dehydration Scale (CDS) or the Gorelick 10-item scale have the sensitivity and specificity to assess the degree of dehydration present compared to the percent weight loss assessed by the gold standard: Percent weight loss due to dehydration defined as: 100 – [(Weight at presentation / Hydrated weight) X 100].

Recommendations from the AGE Team

A conditional recommendation is made for selecting the CDS based on review of current literature or the Age CPG Team, provided by the Department of EBP. As shown in Tables 1 and 2, the overall certainty in the evidence is very low. Sensitivity of both scales is low. While specificity of both tests is low when patients are showing less signs of dehydration, but higher when dehydration is pronounced.

For the CDS scale, the major factor that decreases the certainty of the evidence is the risk of bias in the included studies. Two studies, the first was published in French and the second has not been published but included in a systematic review that was published in English. Both studies are included in this analysis. In general, studies in a language other than English and those that have not yet been published are excluded from evidence at CM. However, since there is so little research on this comparison it was decided to include them.

For the Gorelick 10-item scale, the major factors that decreases certainty in the evidence are a) only two studies are included, and b) each used a different reference test. Neither the CDS, nor the Gorelick 10-item scale are sensitive tests to rule out dehydration, and both tests get more specific in ruling in dehydration as more signs and symptoms are present (See Figures 3 and 4).

As far as reliability and validity are concerned, the CDS has more studies that assess these factors, although various statistical methods have been used (See Table 3). Only two papers report on reliability and validity of the Gorelick 10- item scale. One paper is the initial Gorelick (Gorelick, Shaw, & Murphy, 1997).

Literature Summary

Background. National Health and Nutrition Examination Survey data from 2005 to 2014 reveals the prevalence of AGE in the US population aged 0-9 years (*n* = 9366) years was 14.2% and in 10-19-year olds (*n* = 8703) prevalence was 14.5% (Kim et al., 2017). Dehydration accompanies gastroenteritis and is a major component of morbidity and mortality (King et al., 2003). The standard assessment of dehydration is the percent difference in body weight at presentation to the ED and bodyweight after rehydration (Guarino et al., 2014). Since post-illness weight is not available at ED presentation, percent dehydration cannot be assessed. Clinicians assess severity of dehydration in patients with acute gastroenteritis by reviewing specific the signs and symptoms, such as general appearance, weight loss, capillary refill time, skin turgor, etc. (Geurts, Steyerberg, Moll, & Oostenbrink, 2017). Developed economies have incorporated dehydration scales to increase diagnostic accuracy when caring for this patient population. Two scales, the CDS and the Gorelick Scale 10-item Scale, are commonly used (Friedman, Goldman, Srivastava, & Parkin, 2004; Gorelick, et.al., 1997). This review will summarize identified literature to answer the specific care question regarding the sensitivity, specificity, reliability, and validity of the CDS and Gorelick 10-item scale

***The gold standard*.** The reference test, or gold standard, employed by most of the studies was the percent difference in weight from presentation to weight after rehydration (Guarino et al., 2014). However, there is no consensus on when the rehydrated weight should be obtained. Friedman et al. (2004) obtained the rehydrated weight when the treating physician considered the patient's fluid status is replete, while Gorelick et al. (1997) assessed the difference of a pre-illness weight to the presentation to the ED weight.

Prevalence. The prevalence of dehydration was determined from data collected at Children's Mercy Hospital (CMH) EDs, from the months 11/1/2018 to 10/31/2019 (Allen, 2019). The AGE Team defined levels of dehydration as label the prevalence at CMH. "No Dehydration" (<3%) was defined as patients with ICD10 codes of A02.0, A04.3, A04.4, A04.6, A04.9, A05.9, A07.0, A07.1, A08.0, A08.4, A09, K52.9, R11.10, and R19.7

who did not receive ondansetron. "Some Dehydration" (≥3% to <6%) was defined as patients who received ondansetron, and no intravenous fluid, and "Moderate to Severe Dehydration" (≥6%) was defined as requiring the administration of intravenous fluid.

	Prevalence at CMH *N* = 3444
No Dehydration	45%
Some Dehydration	47%
Moderate to Severe Dehydration	8%

Study characteristics. The search for suitable studies was completed on August 20, 2019 (PubMed) and August 27. 2019 (CINAHL). The PubMed search was repeated December 16, 2019. J. Michael DO reviewed the 75 titles and/or abstracts found in the search using Rayyan[a] and identified 24 single studies believed to answer the question. After an in-depth review of the remaining articles, six studies (Falszewska, Dziechciarz, & Szajewska, 2017; Gorelick et al., 1997; Gravel et al., 2010; Kanjanaphan & Amornchaicharoensuk, 2018; Parkin, Macarthur, Khambalia, Goldman, & Friedman, 2010; Pomorska, Dziechciarz, & Szajewska, n.d.) assessed the diagnostic accuracy of the scales, and reported sensitivity and specificity on one or both of the scales. Six cohort studies (Bailey, Gravel, Goldman, Friedman, & Parkin, 2010; Friedman et al., 2004; Goldman, Friedman, & Parkin, 2008; Gorelick et al., 1997; Gravel et al., 2010; Jauregui et al., 2014; Kinlin & Freedman, 2012) assessed either the validity or reliability or both (see Figure 1).

Summary by Outcome

Diagnostic test accuracy of tools to assess dehydration in patients with AGE.
Five studies (*n* = 755) assessed the diagnostic test accuracy of the CDS (Falszewska et al., 2017; Gravel et al., 2010; Kanjanaphan & Amornchaicharoensuk, 2018; Parkin et al., 2010; Pomorska et al., n.d.):

- Four studies (*n* = 559) included No Dehydration or < 3% data (Falszewska et al., 2017; Gravel et al., 2010; Parkin et al., 2010; Pomorska et al., n.d.).
- Four studies (*n* = 634) provided Some Dehydration or 3% to 6% data (Falszewska et al., 2017; Gravel et al., 2010; Parkin et al., 2010; Pomorska et al., n.d.).
- Five studies (*n* = 755) provided Moderate/Severe Dehydration, > 6% dehydration data (Falszewska et al., 2017; Gravel et al., 2010; Kanjanaphan & Amornchaicharoensuk, 2018; Parkin et al., 2010; Pomorska et al., n.d.).

Three studies (*n* = 563) assessed the Gorelick 10-item scale (Falszewska et al., 2017; Gorelick et al., 1997; Kanjanaphan & Amornchaicharoensuk, 2018).

- All three studies (*n* = 563) provided data on dehydration ≥ 5% or <10% (Falszewska et al., 2017, Gorelick et al., 1997; Kanjanaphan & Amornchaicharoensuk, 2018).
- Two studies (*n* = 338) provided data on dehydration ≥ 10% (Falszewska et al., 2017; Gorelick et al., 1997).

Bias assessment of the six included studies are in Figure 2.

***Sensitivity and Specificity*.** Sensitivity is a measure of the proportion of subjects who actually have the condition and test positive for it, and specificity is a measure of the proportion of subjects who do not have the condition who are correctly classified (Nordenstrom, 2007). See Figures 3 and 4 for forest plots of sensitivity and specificity. The following table shows the ranges of sensitivity/specificity of

Test/Cut-off	Number of Studies	Number of subjects	Sensitivity range	Specificity range
CDS < 3%	4	559	20% to 71%	37% to 100%

CDS 3% to 6%	4	634	63% to 93%	38% to 67%
CDS > 6%	5	775	22% to 67%	38% to 97%
Gorelick 10-item ≥ 5% & <10%	3	5563	9% to 82%	58% to 90%
Gorelick 10-item ≥10%	2	330	82% to 100 %	75% to 91%

Note: CDS - (Falszewska et al., 2017; Gorelick et al., 1997; Gravel et al., 2010; Parkin et al., 2010; Pomorska et al., n.d.). Gorelick 10-Item scale - (Falszewska et al., 2017; Gorelick et al., 1997; Kanjanaphan & Amornchaicharoensuk, 2018)

Certainty of the evidence for the diagnostic test accuracy of tools to assess dehydration. The certainty of the body of evidence for both the CDS and the Gorelick 10-item scale was very low based on four factors: *within-study risk of bias, consistency among studies, directness of evidence, and precision of effect estimates*. The specifics for each test are below.

CDS. For the CDS studies, *risk of bias* was very serious primarily due to patient selection. The method to select subjects was not reported in one study (Kanjanaphan & Amornchaicharoensuk, 2018). Gravel et al. (2010) was completed in Canada and published in French. It met inclusion criteria for this analysis as it was included in a SR completed by Falszewska, Szajewska, and Dziechciarz (2018). Pomorska et al. (n.d.) has not yet been published; however, the pertinent data was included in Falszewska, Szajewska, and Dziechciarz (2018) SR and subsequently included in this analysis. Falszewska et al. (2017) only enrolled subjects when study personnel were available, and Parkin et al. (2010) selected a sub-group of subjects from a dataset of a previous study. The included studies were *inconsistent* in how the reference standard was obtained. One study used a pre-dehydration (from a previous visit) body weight, while others used a post-rehydration weight. The post-hydration weight was obtained either when rehydration was considered complete or at a specific time after the intervention visit, such as one week or two weeks. The sensitivities for the CDS in the included studies vary greatly while the specificities show greater consistency. For all measures, the findings are *imprecise* as there are few studies, with few subjects to answer this question. See Table 1).

Gorelick 10-item scale. For the studies that evaluated the DTA of the Gorelick 10-item scale, the *risk of bias* was serious. Gorelick et al. (1997) only enrolled when study personnel were available, and Kanjanaphan and Amornchaicharoensuk (2018) did not report the enrollment methods utilized. *Inconsistency* was serious as the reference standard in Gorelick was pre-illness weight compared to study admission weight, while Falszewska et al. (2017) obtained a weight specifically for the research study after the subject was discharged. *Imprecision* was very serious, only three studies are included in the analysis for a 5-10% dehydration range, and two studies are included for >10% dehydration. Precision in DTA increases when tools are assessed in multiple locations (Price et al., 2015), See Table 2.

Reliability and validity of tools used to assess dehydration. Three studies (*n* = 517) assessed reliability and/or validity of the CDS (Bailey et al., 2010; Friedman et al., 2004; Kinlin & Freedman, 2012). One study (*n* = 225) assessed the reliability and/or validity of the Gorelick 10-item Scale (Gorelick et al.1997). For validity testing five studies reported on the CDS (Bailey et al., 2010; Friedman et al., 2004; Goldman et al., 2008; Jauregui et al. 2014; Kinlin & Freedman, 2012). The studies employed various statistical techniques to measure reliability and validity. Reliability and validity are not static measures, multiple studies are needed to establish the ability to measure reliability and validity with confidence (Price et al. 2015).

Test	Range	Interpretation
*Cohen's kappa (K)	Values between -1 to +1	*K* = 0 to .20; None *K* = .21 to .39; Minimal *K* = .40 to .59; Weak

		K = .60 to .79; Moderate *K* = .80 to .90; Strong *K* = > .90; Almost perfect
** Interclass Correlation Coefficient *(ICC)*	Values between- 0 to 1	*ICC* < .5; poor reliability *ICC* = .5 to .75; moderate reliability *ICC* = .75 to .9; good reliability *ICC* > .9; excellent reliability

Note: * McHugh (2012); **Koo and Li, (2017)

Reliability.

CDS. Bailey et al. (2010) stated there was excellent agreement between the CDS score and LOS, meaning those with higher CDS scores had longer LOS. Friedman et al. (2004) reported the *Interclass Correlation Coefficient (ICC)* > .6 (moderate reliability) for all items on the CDS except "general appearance" which rated lower. Finally, Kinlin and Freedman (2012) reported interobserver reliability with the *weighted K* = .52, 95% CI [0.41, .63], or weak reliability.

Gorelick 10-item scale. Gorelick et al. (1997) reported the *weighted K* statistic of individual items on the scale and for agreement of any two observers on the presence of any three or more findings. All items in the scale, except "abnormal respirations" had a *weighted K* ≥ 0.5, meaning the assessment of abnormal respirations was the item that varied between observers more than other items. The weighted *K* of agreement between observers of any three or more findings, *weighted K* = .68, or moderate reliability

Validity.

CDS. Criterion validity is the extent in which the assessment tool correlates with other variables (Price, Jhangiani, & Chiang, 2015). In this instance it would be how well either the CDS or the 10-item Gorelick Scale correlates with capillary refill, serum bicarbonate (HCO3), or heart rate. Friedman (2004) reported the final validity of the scale as with a Pearson's correlation coefficient, *r* = .36, 95% CI [.17,.53], or a weak relationship. Goldman, Friedman, and Parkin (2008) reported no agreement for pH < 7.2 or serum bicarbonate level using ANOVA. Construct validity is the amount of correlation between the measure and the construct of interest (Price et al., 2015) in this instance it would be how well either scale correlated with length of stay (LOS) or need for hospitalization. The CDS was significantly associated with LOS and need for intravenous (IV) fluid (*p* < .01), but not associated with successful rehydration. Bailey et al. (2010) reported construct validity for nurses and physicians in their study. For nurses, *r* = .51, 95% CI [.7, .63] and for physicians *r* = .57, 95% CI [.44, .68]. Goldman et al. (2008) reported significant agreement of the CDS score with LOS and Need for IV hydration (*p* < .001).

*Gorelick 10-item scale**.*** Receiver operator curves, and area under the curve (*AUC*) are reported by both Gorelick et al. (1997) and Jauregui et al (2014) for the Gorelick 10-item scale. The *AUC* = 91% when the scale was developed by Gorelick et al. (1997). When

retested by Jauregui et al (2014) for external validity, the *AUC* = 71%. External validity or testing outside of the original research setting increases confidence in the findings original study (Price et al. 2015).

Certainty of the evidence for reliability and validity. The certainty of the body of evidence was very low. based on four factors: *within-study risk of bias, consistency among studies, directness of evidence, and precision of effect estimates*. The body of evidence was assessed to have not *within-study risk of bias, consistency among studies, directness of evidence, and precision of effect estimates*. The findings are inconsistent because the included studies used different statistical tests to report their findings. Precision is high when many studies report findings in a small range, and confidence intervals are narrow; however, in the included studies, findings vary widely, and may not be comparable. See Tables 3 and 4.

Identification of Studies

Search Strategy and Results (see Figure 1)

PubMed:

Search: (("Dehydration/diagnosis"[Mesh] OR "assessing dehydration" OR "dehydration assessment" OR "Clinical Dehydration Scale" OR "WHO scale" OR "World Health Organization scale" OR Gorelick[tiab]) AND "Gastroenteritis"[Mesh]) AND (child OR children OR pediatr* OR paediatr* OR infant) Searched 8/20/2019 and 12/16 2019 *n* = 31

CINAHL

#	Query	Results
S8	S5 AND S6 Limiters - Age Groups: Infant, Newborn: birth-1 month, Infant: 1-23 months, Child, Preschool: 2-5 years, Child: 6-12 years, Adolescent: 13-18 years, All Infant, All Child	37
S7	S5 AND S6	38
S6	(MH "Gastroenteritis+")	26,280
S5	S1 OR S4	425
S4	S2 AND S3	274
S3	(MH "Scales") OR (MH "Clinical Assessment Tools+") OR "Clinical Dehydration Scale" OR "WHO scale" OR "Gorelick" OR "World Health Organization scale" OR "dehydration assessment" OR "assessing dehydration" OR "Clinical Signs of Dehydration"	205,720
S2	(MH "Dehydration")	3,586
S1	(MM "Dehydration/DI")	202

Records identified through database searching *n* = 105
Additional records identified through other sources *n* = 2

Studies Included in this Review

Citation	Study Type
Bailey et al. (2010)	Cohort
Falszewska et al. (2017)	Diagnostic Test Accuracy
Friedman et al. (2004)	Cohort
Goldman et al. (2008)	Cohort
Gorelick et al. (1997)	Cohort
Gravel et al. (2010)	Diagnostic Test Accuracy
Jauregui et al. (2014)	Cohort
Kanjanaphan and Amornchaicharoensuk (2018)	Diagnostic Test Accuracy
Kinlin and Freedman (2012)	Cohort
Parkin et al. (2010)	Diagnostic Test Accuracy
Pomorska et al. (n.d.)	Diagnostic Test Accuracy

Studies Not Included in this Review with Exclusion Rationale

Citation	Reason for exclusion
Colletti et al. (2010)	Does not answer the question; proposes a different dehydration scale based on change in body weight
Falszewska, Dziechciarz, and Szajewska (2014)	Updated by Falszewska, Szajewska, and Dziechciarz (2018)
Falszewska et al. (2018)	Added data from Kanjanaphan and Amornchaicharoensuk (2018)
Freedman, Adler, Seshadri, and Powell (2006)	Does not answer the question; an RCT ondansetron vs. placebo
Freedman, DeGroot, and Parkin (2014)	Does not answer the question; does not include any dehydration scale
Freedman et al. (2015)	Assess ultrasound, and urinalysis
Geurts et al. (2017)	Does not answer the question
Hayajneh, Jdaitawi, Al Shurman, and Hayajneh (2010)	Performed in a developing country[a]
T. F. Hoxha et al. (2014)	Does not answer the question; also performed in countries with Economies in Transition (Kosovo and Serbia)
T. Hoxha et al. (2015)	Does not answer the question; also performed in countries with Economies in Transition (Kosovo and Serbia)
Kuge, Morikawa, and Hasegawa (2017)	Does not answer the question; index test was uric acid
Levine et al. (2010)	Performed in a Developing/Low income country[a] (Rwanda)
Levine et al. (2013)	Performed in a Developing/Low income country[a] (Rwanda)
Powell, Priestley, Young, and Heine (2011)	Used the Gorelick score; did not test the Gorelick score
Pringle et al. (2011)	Performed in a Developing/Low income country[a] (Rwanda)
Shavit, Brant, Nijssen-Jordan, Galbraith, and Johnson (2006)	Index test was digitally measured capillary refill time

Steiner, Nager, and Wang (2007)	Index test was urine specific gravity and urine ketones
Tam, Wong, Plint, Lepage, and Filler (2014)	Use case comparison method
Vega and Bhimji (2018)	Index test was physician assessment, only

Methods Used for Appraisal and Synthesis

[a]The United Nations report on the world economic situation was used to delineate economically developed countries from non-developed countries.
[b]Rayyan is a web-based software used for the initial screening of titles and / or abstracts for this analysis (Ouzzani, Hammady, Fedorowicz & Elmagarmid, 2017).
[c]The GRADEpro Guideline Development Tool (GDT) is the tool used to create the Summary of Findings table(s) for this analysis (see Tables 1 and 2).
[d]The Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) flow diagram depicts the process in which literature is searched, screened, and eligibility criteria is applied (Moher, Liberati, Tetzlaff, & Altman, 2009).
[e]The Quality Assessment of Diagnostic Accuracy Studies-2 (QUADAS-2) (Whiting et al., 2011) is was used to assess the sources of bias and variation in the diagnostic studies found in this analysis.
[f]Review Manager (Higgins & Green, 2011) is a Cochrane Collaborative computer program used to assess the study characteristics as well as the risk of bias and create the forest plots found in this analysis.

[af]United Nations Department of Economic and Social Affairs (2019). World Economic Situation and Prospects. Retrieved from https://www.un.org/development/desa/dpad/wp-content/uploads/sites/45/WESP2019_BOOK-web.pdf
[b]Ouzzani, M., Hammady, H., Fedorowicz, Z., & Elmagarmid, A. (2016). Rayyan-a web and mobile app for systematic reviews. *Systematic Reviews, 5*(1), 210. doi:10.1186/s13643-016-0384-4
[c]GRADEpro GDT: GRADEpro Guideline Development Tool (2015). McMaster University, (developed by Evidence Prime, Inc.). [Software]. Available from gradepro.org.
[d]Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group (2009). *P*referred *R*eporting *I*tems for *S*ystematic Reviews and *M*eta-*A*nalyses: The PRISMA Statement. PLoS Med 6(7): e1000097. doi:10.1371/journal.pmed1000097 **For more information, visit www.prisma-statement.org.**
[e]Whiting, P. F., Rutjes, A. W., Westwood, M. E., Mallett, S., Deeks, J. J., Reitsma, J. B., ... & Bossuyt, P. M. (2011). QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Annals of internal medicine*, *155*(8), 529-536.
[f]Higgins, J. P. T., & Green, S. e. (2011). *Cochrane Handbook for Systematic Reviews of Interventions [updated March 2011]* (Version 5.1.0 ed.): The Cochrane Collaboration, 2011.

Question Originator:
The Acute Gastroenteritis CPG Team

Medical Librarian Responsible for the Search Strategy
Keri Swaggart, MLIS, AHIP

EBP Scholar's Responsible for Analyzing the Literature
Teresa Bontrager, MSN, RN, CPEN
Justine Edwards, RN, MSN, CPEN
Rebecca Frederick, PharmD
Kori Hess, PharmD
David Kemper, BHS, RRT, RRT-NPS, C-NPT
Linda Martin, RN, BSN, CPAN
Robyn McCracken, RRT, NPS
Marla Michaels, MSN, RN, CPN
Helen Murphy, BHS RRT AE-C
Lucy Pappas, MS RD CSP LD

Anthony Randall, MHA, RRT, RRT-ACCS, RRT-NPS, C-NPT, CPPS
Kim Robertson, MBA, MT-BC
Giselle Scott, LMSW
Audrey Snell, MS, RD, LD
Britney Snodgrass, MSN, RN, CPN
Rhonda Sullivan, MS, RD, CSP, LD
Rochelle Vossman, MSN, RN-BC
Jami Wierson, RN, BSN, MBA, CCRC
Ashley Wilson, BSN, RN, CPN

EBP Team Member Responsible for Reviewing, Synthesizing, and Developing this Document

Nancy H. Allen, MS, MLS, RD, LD, CPHQ

Acronyms Used in this Document

Acronym	Explanation
AGE	Acute gastroenteritis
ANOVA	Analysis of variance
AUC	Area under the curve
CDS	Clinical Dehydration Score
CoE	Confidence in evidence
EBP	Evidence Based Practice
ED	Emergency department
HCO3	Bicarbonate
ICC	Interclass correlation coefficient
ICD10	International classification of disease 10 ed.
IQR	Interquartile range
IV	Intravenous
LOS	Length of stay
ORT	Oral rehydration solution
PRISMA	Receiver operator curve
ROC	Receiver operating characteristic
WHO	World Health Organization

Date Developed

February 2020

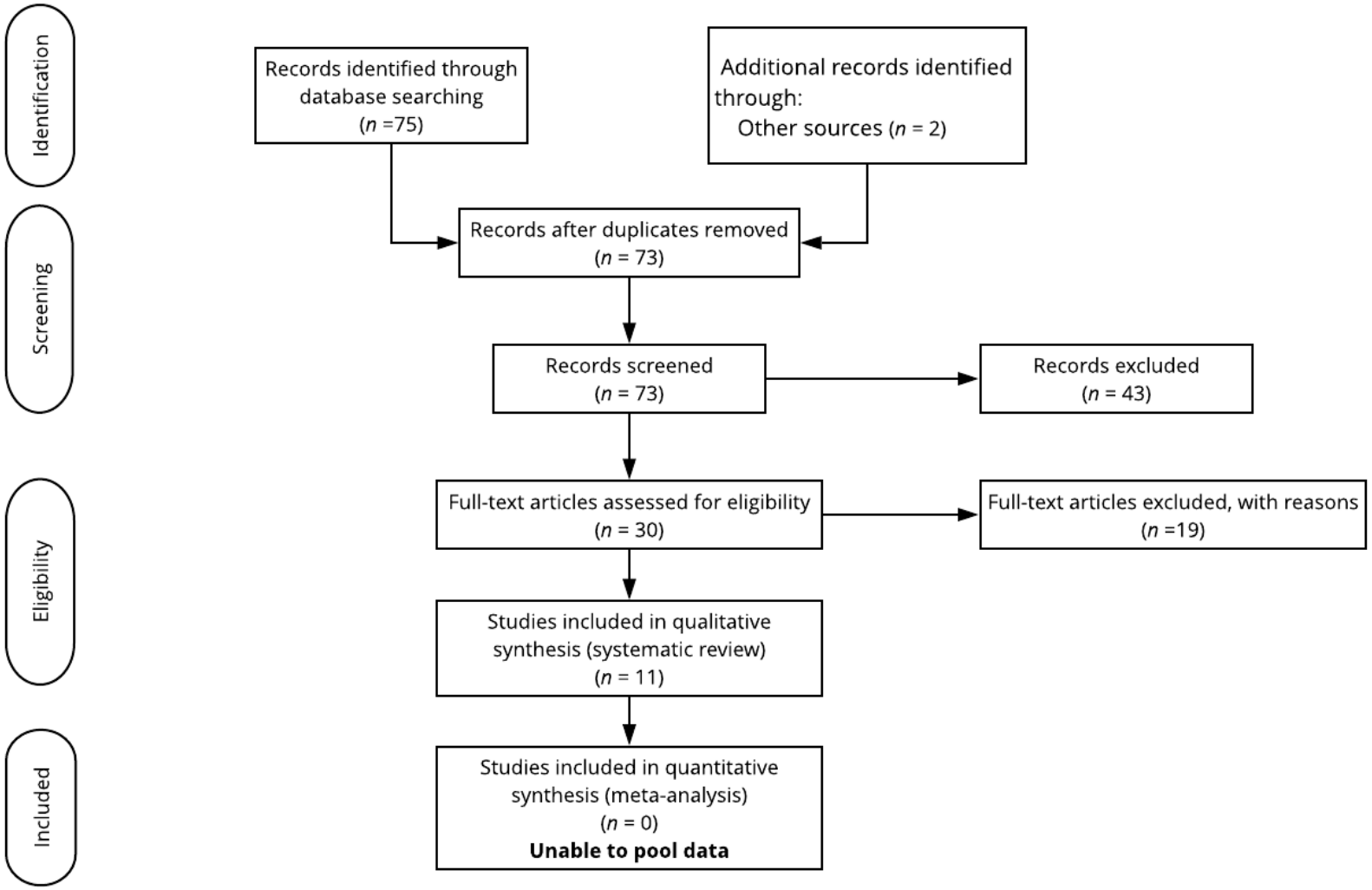


Figure 1. Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRIMSA)[d]

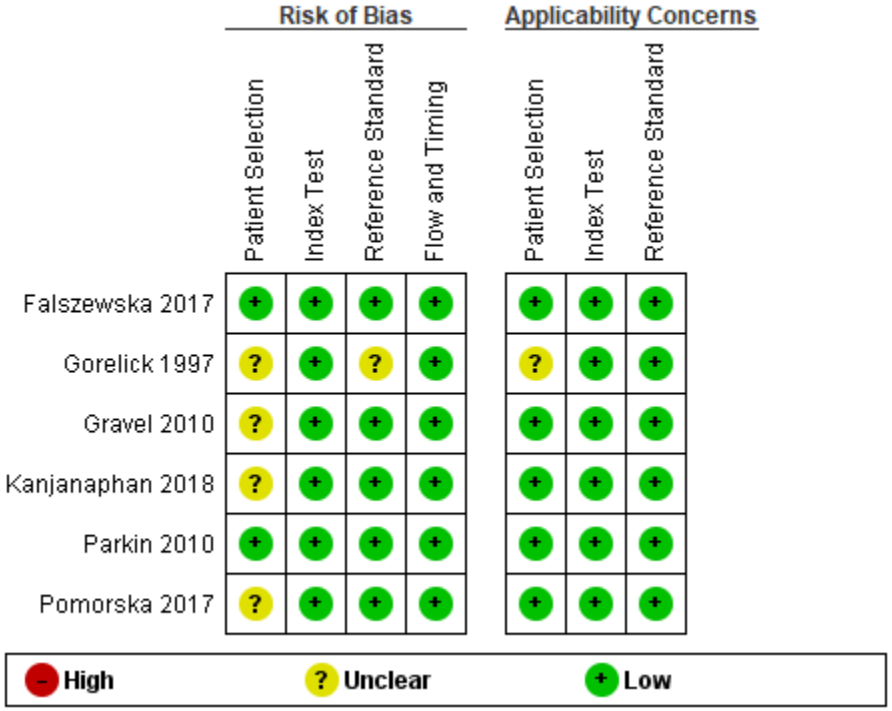


Figure 2. Risk of Bias Summary for DTAs

Summary of Findings Tables[c]

Table 1

Question: Should CDS be used to assess the severity of dehydration in acute gastroenteritis?

	CDS = 0 No Dehydration (<3%)	CDS = 1-4 Some Dehydration (3% to 6%)	CDS = 4-8 Moderate to Severe Dehydration (>6%)
Prevalence at CMH	45%	47%	8%
Range of Sensitivity	20% to 71%	63% to 93%	22% to 67%
Range of Specificity	37% to 100%	38% to 67%	38% to 97%

Outcome	Nº studies Nº patients	Study design	Factors that may decrease certainty of evidence					Effects per 1,00 patients tested	Test accuracy CoE
			Risk of bias	Indirectness	Inconsistency	Imprecision	Publication Bias	Pre-test probability based on prevalence	
CDS = 0, No Dehydration (<3%)									
True positives (patients with no dehydration)	4 studies 373 patients	cohort & case-control type studies	very serious[a]	not serious	serious[b,c]	serious [d]	none	18 to 54	⊕◯◯◯ VERY LOW
False negatives (patients incorrectly classified as not having no dehydration)								26 to 62	
True negatives (patients without no dehydration)	4 studies 186 patients	cohort & case-control type studies	very serious[a]	not serious	serious[b,c]	serious[d]	none	305 to 892	⊕◯◯◯ VERY LOW
False positives (patients incorrectly classified as having no dehydration)								28 to 570	
CDS 1-4, Some Dehydration (3-6%)									
True positives (patients with	4 studies 229 patients		very serious[a]	not serious	serious[b,c]	serious[f]	none	296 to 437	

some dehydration)									
False negatives (patients incorrectly classified as not having some dehydration)		cohort & case-control type studies						33 to 174	⊕◯◯◯ VERY LOW
True negatives (patients without some dehydration)								201 to 355	
False positives (patients incorrectly classified as having some dehydration)	4 studies 405 patients	cohort & case-control type studies	very seriousa	not serious	seriousf	not serious	none	175 to 329	⊕◯◯◯ VERY LOW
CDS > 4 Severe Dehydration (>6%)									
True positives (patients with severe dehydration)								18 to 54	
False negatives (patients incorrectly classified as not having severe dehydration)	5 studies 123 patients	cohort & case-control type studies	very seriousa	not serious	not serious	serious k	none	26 to 62	⊕◯◯◯ VERY LOW
True negatives (patients without severe dehydration)								350 to 892	
False positives (patients incorrectly classified as having severe dehydration)	53 studies 632 patients	cohort & case-control type studies	very seriousa	not serious	serious h	not serious	none	28 to 570	⊕◯◯◯ VERY LOW

Explanations

a. Selection practice of two studies was not determined. Gravel et al. (2010) completed in Canada but published in French is included. It met inclusion criteria for this study as it was included in a SR completed by Falszewska et al. (2018). Pomorska et al. (n.d.) has not yet been published. Falszewska et al. (2017 only enrolled subjects when study personnel were available, and Parkin (2018) selected subjects from a dataset of a previous study.

b. Confidence intervals do not overlap.

c. Ages of included subjects differed among the included studies

d Falszewska et al. (2017) the study that has not yet been published, had zero true positive tests for this cut-off

e Precision is assessed by the width of confidence intervals for the sensitivities reported by the included studies. For this comparison the confidence intervals are wide.

f. The range of specificity varied widely

Table 2
Question: Should the Gorelick 10-item Scale be used to assess the severity of dehydration in acute gastroenteritis?

	Gorelick-10 item scale < 5 No Dehydration (<5%)	Gorelick 10-item scale ≥ 5 and < 10 Some Dehydration (≥5% to 10%)	Gorelick 10-item scale ≥10 Moderate to Severe Dehydration (≥10%)
Prevalence at CMH	45%	47%	8%
Range of Sensitivity	Not reported	9% to 82%	22% to 67%
Range of Specificity	Not reported	82% to 100%	75% to 91%

Outcome	Nº studies Nº patients	Study design	Factors that may decrease certainty of evidence					Effects per 1,00 patients tested	Test accuracy CoE
			Risk of bias	Indirectness	Inconsistency	Imprecision	Publication Bias	Pre-test probability based on prevalence	
Gorelick 10-item scale ≥ 5% and < 10% fluid deficit									
True positives (patients with dehydration)	3 studies 258 patients	cohort & case-control type studies	serious[a]	not serious	serious[b]	serious [c]	none	42 to 385	⊕◯◯◯ VERY LOW
False negatives (patients incorrectly classified as not having dehydration)								85 to 428	
True negatives (patients without dehydration)	3 studies 305 patients	cohort & case-control type studies	serious[a]	not serious	serious[b]	Serious[c]	none	307 to 447	⊕◯◯◯ VERY LOW
False positives (patients incorrectly classified as having dehydration)								53 to 223	
Gorelick 10-item scale, ≥ 10% fluid deficit									

True positives (patients with dehydration)	2 studies 270 patients	cohort & case-control type studies	seriousd	not serious	seriousb,c	seriousf	none	66 to 80	⊕◯◯◯ VERY LOW
False negatives (patients incorrectly classified as not having no dehydration)								0 to 14	⊕◯◯◯ VERY LOW
True negatives (patients without dehydration)	2 studies 68 patients	cohort & case-control type studies	seriousd	not serious	seriousb	seriousc	none	690 to 837	⊕◯◯◯ VERY LOW
False Positives (patients incorrectly classified as having dehydration)								83 to 230	⊕◯◯◯ VERY LOW

Explanations:

a. Gorelick (2017) only enrolled when study personnel were available, Falszewska (2017) reported they used convenience sampling, but did not define the method used, and Kanjanaphan (2018) did not report the method used to enroll subjects

b. Different standards were used for the reference test. Gorelick (1997) used a pre illness weight, while Falszewsksa (2017) and Kanjanaphan (2018) used a post illness weight obtained specifically for this research

c. Three studies with small numbers of subjects (*n* = 563) have been included in this analysis

d Gorelick (2017) only enrolled when study personnel were available, Falszewska (2017) reported they used convenience sampling, but did not define the method used

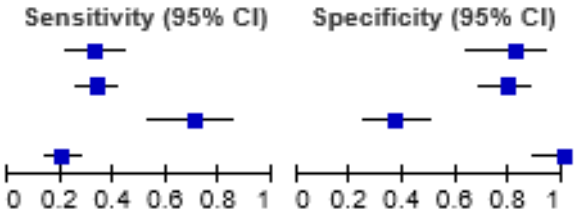
e. Different standards were used for the reference test. Gorelick (1997) used a pre illness weight, while Falszewsksa (2017) used a post illness weight obtained specifically for this research

f. Two studies with small numbers of subjects (*n* = 338) have been included in this analysis.

Forest Plots

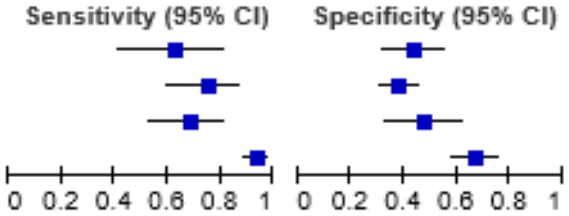
CDS < 3%

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Falszewska 2017	23	5	48	23	0.32 [0.22, 0.45]	0.82 [0.63, 0.94]
Gravel 2010	50	14	100	55	0.33 [0.26, 0.41]	0.80 [0.68, 0.88]
Parkin 2010	24	37	10	22	0.71 [0.53, 0.85]	0.37 [0.25, 0.51]
Pomorska 2017	24	0	94	30	0.20 [0.13, 0.29]	1.00 [0.88, 1.00]



CDS 3-6%

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Falszewska 2017	15	42	9	32	0.63 [0.41, 0.81]	0.43 [0.32, 0.55]
Gravel 2010	30	111	10	68	0.75 [0.59, 0.87]	0.38 [0.31, 0.46]
Parkin 2010	30	26	14	23	0.68 [0.52, 0.81]	0.47 [0.33, 0.62]
Pomorska 2017	113	34	8	69	0.93 [0.87, 0.97]	0.67 [0.57, 0.76]



CDS > 6%

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Falszewska 2017	2	12	2	82	0.50 [0.07, 0.93]	0.87 [0.79, 0.93]
Gravel 2010	9	5	20	185	0.31 [0.15, 0.51]	0.97 [0.94, 0.99]
Kanjanaphan 2018	16	21	57	127	0.22 [0.13, 0.33]	0.86 [0.79, 0.91]
Parkin 2010	4	8	4	77	0.50 [0.16, 0.84]	0.91 [0.82, 0.96]
Pomorska 2017	6	71	3	44	0.67 [0.30, 0.93]	0.38 [0.29, 0.48]

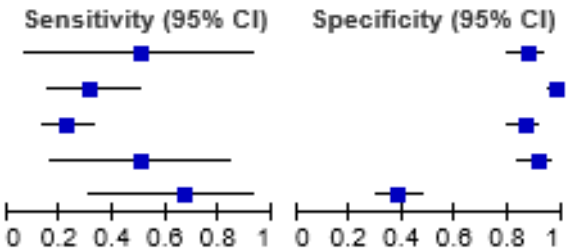


Figure 3. ***Sensitivity and Specificity of the CDS at 3 Cut-off Points***

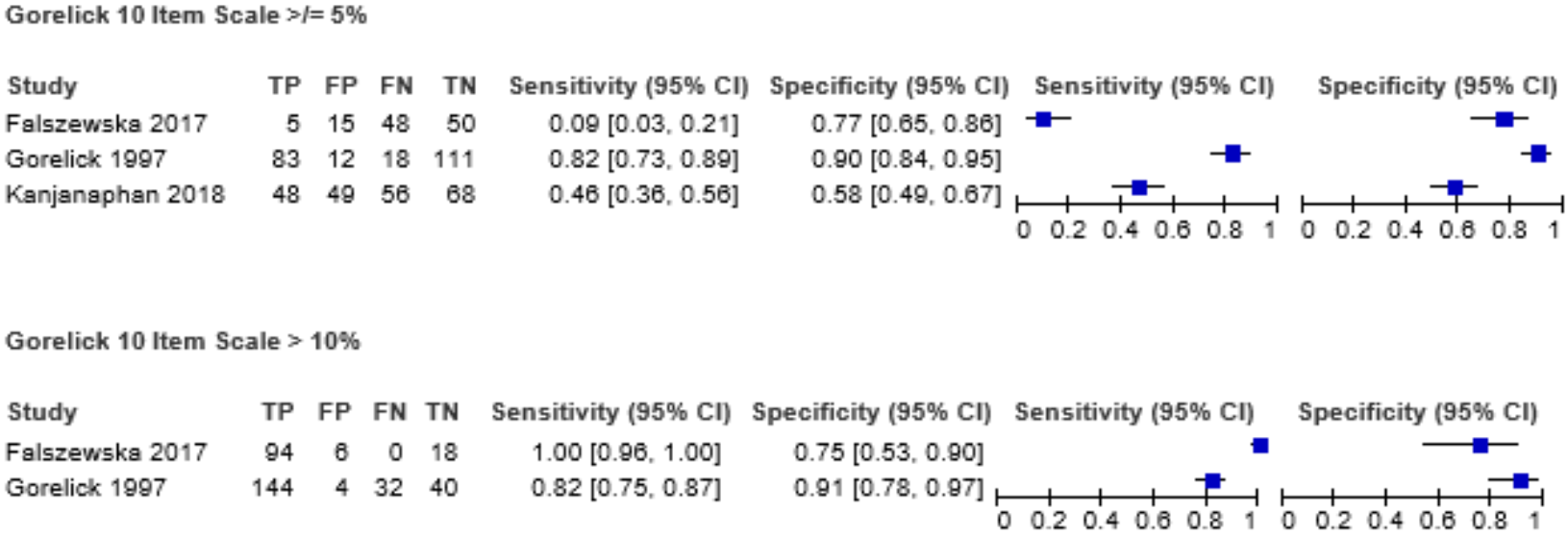


Figure 4. ***Sensitivity and Specificity of the Gorelick 10-item scale at 2 Cut-off Points***

Table 3

Reliability and Validity of the CDS

<table>
<tr><td colspan="3">Reliability</td></tr>
<tr><td>Study</td><td>Test</td><td>Finding</td></tr>
<tr><td>Bailey et al. (2010)</td><td>Bland Altman chart – measures agreement, not correlation (Rangnathan, Pramesh, Aggarwal, 2017)</td><td>Rated “excellent”. Reported Mean Bias = 0.5 minutes, 95% CI [-2, 3] for the main outcome LOS. Interpret as for any paired CDS score, LOS varied by a mean of 0.5 minutes with 95% CI as indicated above. (Rangnathan, Pramesh, Aggarwal, 2017)</td></tr>
<tr><td>Friedman et al. (2004)</td><td>Intra-class correlation coefficient (ICC) measures consistency or agreement. Range is 0 to 1. Higher is better (Pett et al, 2003)
Discriminatory Power - Ferguson’s Delta (δ) range is (0 to 1). Higher is better.
Responsiveness to Change - Wilcoxon’s signed rank test</td><td><table><tr><th>Measurement</th><th>Result</th></tr><tr><td>Interclass Correlation coefficient</td><td>ICC > .6 for all items except “general appearance”</td></tr><tr><td>Ferguson’s δ</td><td>Ferguson’s δ = 0.83</td></tr><tr><td>Wilcoxon’s signed rank test</td><td>Start of therapy = 2, range (0,8)
End of therapy = 0, range (0,2)</td></tr></table></td></tr>
<tr><td>Kinlin and Freedman (2012)</td><td>Interobserver reliability – weighted K statistic</td><td>K = .52. 95% CI [.41, .63]</td></tr>
<tr><td colspan="3">Validity</td></tr>
<tr><td>Bailey et al. (2010)</td><td>Assessed association between LOS after seen by a physician and CDS score. Mann-Whitney test was used to evaluate each pair of CDS categories when continuous measures were different. For dichotomous values Chi-square was used.</td><td><table><tr><th>Association between LOS after seen by physician and CDS score</th><th>LOS Minutes, median, (IQR) (p < .01) for all</th><th>Need for IV fluid n (%) (p < .01) for all</th><th>Successful oral rehydration (%) (p = .06) for all</th></tr><tr><td>No dehydration, n = 56</td><td>54 [26,175]</td><td>5 (12%)</td><td>17/19 (90%)</td></tr><tr><td>Some dehydration n (%) = 74</td><td>128 [25,334]</td><td>23 (32%)</td><td>22/29 (76%)</td></tr><tr><td>Moderate/Severe dehydration n (%) = 20</td><td>425 [218, 673]</td><td>13 (65%)</td><td>5/10 (50%)</td></tr></table></td></tr>
</table>

Friedman et al. (2004)

Pearson's correlation coefficient for criterion and construct validity

Measurement	N	Pearson's correlation coefficient	95% CI
Criterion validity	93	.36	[.17, .53]
Construct validity (RNs)	122	.51	[0.7, .63]
Construct validity (MDs)	120	.57	[.44, .68]

Goldman et al (2008)

Analysis of variance for continuous data and *Chi-square* for dichotomous data

	CDS <3	CDS = 3-6	CDS = >6	*p*
LOS, mean ± SD, min	245 ± 181	397 ± 302	501 ± 389	<.001
IV rehydration, *n* (%)	17 (15)	41 (49)	4 (80)	.001
pH of < 7.32, *n* (%)	2 (14)	14 (34)	1 (25)	.36
HCO3 level of < 18 mEq/L, *n* (%)	4 (29)	16 (39)	3 (75)	.22

Jauregui et al. (2014)

Receiver Operator Curves (ROC)

Measure and cut off point	*AUC* (95% CI)
CDS (2 of an 8-point scale)	.72 (.60, .84)

Kinlin and Freedman (2012)

Pearson correlation coefficient or Spearmen rank correlation, depending on distribution of the data

Measure		Result
Construct validity	Weight gain	r_s = .04, [-.14, .19],
	LOS	*r = .24 [.11, .36]*
	Serum HCO3	*r* = -.35, [-.46, -.23]
Discriminative validity	Hospitalization	*AUC* = .35 [.57, .73]
Responsiveness of CDS	Start of Therapy vs End of therapy	*p* < .01

Table *4*

Reliability and Validity of the Gorelick-10 item scale

Reliability			
Study	Test(s)	Finding(s)	
Gorelick et al. (1997)	Interobserver reliability – *Weighted K* statistic for dichotomous values *ICC* for continuous values	*K* = ≥ .5 for all but abnormal respirations *ICC* = .71 for capillary refill time *K* = .68 for agreement between observers on the presence of any three or more findings.	
Validity			
Gorelick et al. (1997)	*ROC*	Measure and cut off point	*AUC*
		Gorelick 10-item scale	.91
Jauregui et al. (2014)	*ROC*	Measure and cut off point	*AUC* [95% CI]
		Gorelick (2 of a 10-item scale)	.71 [.57, .85]

Characteristics of Diagnostic Tests of Accuracy Studies

Falszewska, et al. (2017)

Patient Sampling	Prospective observational study convenience series • Convenience series is not defined, such as only included subject when study trained staff was available, hour if specific shifts were excluded etc.
Patient characteristics and setting	**Participants:** • Children 1 month to 5 years with acute diarrhea • Greater or equal to 3 evacuations in 24 hours • Lasting no longer than 5 days **Setting:** Pediatric Inpatient wards of a University hospital **Number enrolled into the study:** *N* = **Number completed: the study:** *N* = **Gender, males**: *n* = **Race/ethnicity or nationality as defined by the researchers): Not reported,** **Age:** • Gorelick scale - 15 months **Exclusion criteria:** • Dehydration caused by other causes, such as ketoacidosis, kidney failure, heart failure, etc. **Registration:** Trial was registered at ClinicalTrials.gov NCT02249845
Index test	• Clinical Dehydration Scale (CDS) • World Health Organization (WHO) scale • Gorelick scale ○ Children < 3 years of age, all three scales were used ○ Children > 3 years, the WHO and Gorelick scales were used
Target condition and reference standard(s)	The target condition is dehydration, 3%, 3-6%, and > 6% The reference standard is percent weight change, calculated as (final weight subtracted from initial weight) divided by final weight times by 100
Flow and timing	Subjects < 3 years were scored on all three dehydration scales, while subjects > 3 were scored on the WHO and CDS scale on admission. The scores were recorded on a prespecified data sheet, but not scores were not totaled. All subjects were weighed in a standard manner, on calibrated age appropriated scales on hospital admission and at hospital discharge.
Notes	Ten subjects were withdrawn from the analysis. Four subjects had missing discharge weights; Five subjects left against medical advice and were not weighed at discharge; one subject was transferred to another hospital.

Patient Selection

A. Risk of Bias	
Was a consecutive or random sample of patients enrolled?	Unclear

Was a case-control design avoided?	Yes
Did the study avoid inappropriate exclusions?	Yes
Could the selection of patients have introduced bias?	Low risk
B. Concerns regarding applicability	
Are there concerns that the included patients and setting do not match the review question?	Low concern

All tests

A. Risk of Bias	
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes
If a threshold was used, was it pre-specified?	Yes
Could the conduct or interpretation of the index test have introduced bias?	Low risk
B. Concerns regarding applicability	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?	Low concern

Reference Standard

A. Risk of Bias	
Is the reference standards likely to correctly classify the target condition?	Yes
Were the reference standard results interpreted without knowledge of the results of the index tests?	No
Could the reference standard, its conduct, or its interpretation have introduced bias?	Low risk
B. Concerns regarding applicability	
Are there concerns that the target condition as defined by the reference standard does not match the question?	Low concern

Flow and Timing

A. Risk of Bias	
Was there an appropriate interval between index test and reference standard?	Yes
Did all patients receive the same reference standard?	Yes
Were all patients included in the analysis?	No
Could the patient flow have introduced bias?	Low risk

Gorelick et al. (1997)

Patient Sampling	Age 1 month to 5 years, urban pediatric emergency department, USA. Only enrolled when study staff was on service
Patient characteristics and setting	**Participants:** Children aged 1 month to 5 years

	Setting: Pediatric emergency department **Number enrolled into study:** *N* = 225, • *n* = 116 hospitalized for AGE • *n* = 109 followed as outpatients **Number completed:** *N* = 186 **Gender, males:** *n* = 102 (55%) **Race / ethnicity or nationality:** Not reported. The study was performed at an urban pediatric ED in the USA **Age, median:** 13 months, 89% were < 13 months of age **Exclusion criteria:** • Symptoms > 5 days • History of cardiac disease, renal disease, diabetes mellitus, malnutrition, failure to thrive or treatment in the past 12 hours at another facility • If serum electrolytes were obtained, subjects with hypo or hypernatremia were excluded • Tonsillectomy in the past 10 days and managed by the otolaryngology physicians Unable to contact family for follow-up
Index test	Gorelick 10-item scale
Target condition and reference standard(s)	Percent fluid deficit determined by pre and post illness weight. Unclear timing on pre-illness weight.
Flow and timing	Body weight taken at admit and post weight was determined in all subjects who were admitted, and 30% of those discharged from the ED
Notes	Interobserver reliability was measured in a subset of subjects. Eighty-four subjects had independent Gorelick-10 item scale assessments. For individual items assessed on the scale all had $K \geq .5$ for all but one of the findings. For the presence of any three or more findings agreement was good with $K = .68$

Patient Selection

A. Risk of Bias	
Was a consecutive or random sample of patients enrolled?	No
Was a case-control design avoided?	Yes
Did the study avoid inappropriate exclusions?	Yes
Could the selection of patients have introduced bias?	Unclear risk
B. Concerns regarding applicability	
Are there concerns that the included patients and setting do not match the review question?	Unclear concerns

All tests

A. Risk of Bias	
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes

If a threshold was used, was it pre-specified?	No
Could the conduct or interpretation of the index test have introduced bias?	Low risk
B. Concerns regarding applicability	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?	Low concern

Reference Standard

A. Risk of Bias	
Are the reference standards likely to correctly classify the target condition?	Yes
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes
Could the reference standard, its conduct, or its interpretation have introduced bias?	Unclear risk
B. Concerns regarding applicability	
Are there concerns that the target condition as defined by the reference standard does not match the question?	Low concern

Flow and Timing

A. Risk of Bias	
Was there an appropriate interval between index test and reference standard?	Yes
Did all patients receive the same reference standard?	No
Were all patients included in the analysis?	Yes
Could the patient flow have introduced bias?	Low risk

Gravel et al. (2010)

Patient Sampling	Study in French, bias assessment taken from Falszewska et al. (2018), other information was taken from the abstract only Convenience sampling
Patient characteristics and setting	**Participants:** Children aged 1 to 60 months **Setting:** Three university affiliated EDs in Canada **Number enrolled into study: *N*** = 264 **Number completed: *N*** = 264 **Gender, males: *n* =** Not reported **Race / ethnicity or nationality:** Not reported, study was performed in Canada **Age, months (median, range):** Not reported **Exclusion criteria**: Not reported in abstract
Index test	CDS

Target condition and reference standard(s)	
Flow and timing	

Patient Selection

A. Risk of Bias	
Was a consecutive or random sample of patients enrolled?	
Was a case-control design avoided?	
Did the study avoid inappropriate exclusions?	
Could the selection of patients have introduced bias?	Unclear risk
B. Concerns regarding applicability	
Are there concerns that the included patients and setting do not match the review question?	Low concern

All tests

A. Risk of Bias	
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes
If a threshold was used, was it pre-specified?	Yes
Could the conduct or interpretation of the index test have introduced bias?	Low risk
B. Concerns regarding applicability	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?	Low concern

Reference Standard

A. Risk of Bias	
Are the reference standards likely to correctly classify the target condition?	Yes
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes
Could the reference standard, its conduct, or its interpretation have introduced bias?	Low risk
B. Concerns regarding applicability	
Are there concerns that the target condition as defined by the reference standard does not match the question?	Low concern

Flow and Timing

A. Risk of Bias	
Was there an appropriate interval between index test and reference standard?	Yes
Did all patients receive the same reference standard?	Yes

Were all patients included in the analysis?	No
Could the patient flow have introduced bias?	Low risk

Jauregui et al. (2014)

Patient Sampling	Prospective enrollment of non-consecutive subjects. Secondary analysis of data from a randomized control trial
Patient characteristics and setting	**Participants:** Children, ≤ 18 years old **Setting:** Rhode Island, USA **Number enrolled into study: *N*** = 148 **Number completed: *N*** = 113 **Gender, males: *n*** = 51% **Race / ethnicity or nationality:** Not reported; the study was performed in Rhode Island, USA **Age, years (median, range):** 6 years (1 month, 17 years) **Exclusion criteria**: Positive pressure ventilation, significant traumatic injury, large volume fluid administration prior to enrolment, surgical abdomen, or known congenital cardiac disease
Index test	Physician gestalt, the Gorelick 10-item scale, the WHO scale, and the CDS.
Target condition and reference standard(s)	Percent weight loss with the illness. Weight taken at admission and at discharge from either the ED, or hospital if admitted
Flow and timing	The attending physician recorded their gestalt on level of dehydration, and then the attending physician completed each of the scores, in a blinded fashion. Interobserver reliability not assessed, nor was validity
Notes	Prospective enrollment of non-consecutive subjects. Secondary analysis of data from a randomized control trial

Patient Selection

A. Risk of Bias	
Was a consecutive or random sample of patients enrolled?	Yes
Was a case-control design avoided?	Yes
Did the study avoid inappropriate exclusions?	No
Could the selection of patients have introduced bias?	Unclear risk
B. Concerns regarding applicability	
Are there concerns that the included patients and setting do not match the review question?	Low concern

All tests

A. Risk of Bias	
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes
If a threshold was used, was it pre-specified?	No
Could the conduct or interpretation of the index test have introduced bias?	Unclear risk

B. Concerns regarding applicability	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?	Unclear concern

Reference Standard

A. Risk of Bias	
Is the reference standards likely to correctly classify the target condition?	Yes
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes
Could the reference standard, its conduct, or its interpretation have introduced bias?	Low risk
B. Concerns regarding applicability	
Are there concerns that the target condition as defined by the reference standard does not match the question?	Low concern

Flow and Timing

A. Risk of Bias	
Was there an appropriate interval between index test and reference standard?	Yes
Did all patients receive the same reference standard?	Yes
Were all patients included in the analysis?	Yes
Could the patient flow have introduced bias?	Low risk

Kanjanaphan and Amornchaicharoensuk (2018)

Patient Sampling	Not reported
Patient characteristics and setting	**Participants:** Children 1 month to 15 years **Setting:** Inpatient setting **Number enrolled into study:** *N* = 220 **Number completed:** *N* = 220 **Gender, males:** Not reported **Race / ethnicity or nationality:** Not reported; study was performed in Thailand **Age (in months) median:** 39
Index test	Gorelick 10-item scale and the CDS
Target condition and reference standard(s)	Pre and post treatment body weight. Percent dehydration from change in body weight was assessed after post body weight was obtained.
Flow and timing	The degree of dehydration was assessed by the physician, and the data was recorded in the two index scales. All patients were treated with fluid replacement as maintenance plus fluid to correct a #5, 6%, or 10% fluid deficit.

Patient Selection

A. Risk of Bias	
Was a consecutive or random sample of patients enrolled?	Unclear
Was a case-control design avoided?	Yes
Did the study avoid inappropriate exclusions?	Yes
Could the selection of patients have introduced bias?	Unclear risk
B. Concerns regarding applicability	
Are there concerns that the included patients and setting do not match the review question?	Low concern

All tests

A. Risk of Bias	
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes
If a threshold was used, was it pre-specified?	Yes
Could the conduct or interpretation of the index test have introduced bias?	Low risk
B. Concerns regarding applicability	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?	Low concern

Reference Standard

A. Risk of Bias	
Are the reference standards likely to correctly classify the target condition?	Yes
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes
Could the reference standard, its conduct, or its interpretation have introduced bias?	Low risk
B. Concerns regarding applicability	
Are there concerns that the target condition as defined by the reference standard does not match the question?	Low concern

Flow and Timing

A. Risk of Bias	
Was there an appropriate interval between index test and reference standard?	Yes
Did all patients receive the same reference standard?	Yes
Were all patients included in the analysis?	Yes
Could the patient flow have introduced bias?	Low risk

Parkin, et al. (2010)

Patient Sampling	Selected subjects from a database, not consecutive, they had to meet criteria

Patient characteristics and setting	
Index test	CDS-four item scale (range 0 to 8, lower better) • Score = 0, None • Score = 1-4, Some Score = 5-8, Moderate to Severe
Target condition and reference standard(s)	Dehydration Percent weight change • None, <3% weight gain • Some, ≥ 3% to ≤ 6% Moderate to severe, > 6%
Flow and timing	Scores and pre-weights obtained by enrolling study personnel. Post weight was obtained when attending physician deemed subject was ready for discharge. Appears flow and timing were appropriate

Patient Selection

A. Risk of Bias	
Was a consecutive or random sample of patients enrolled?	Unclear
Was a case-control design avoided?	Yes
Did the study avoid inappropriate exclusions?	Yes
Could the selection of patients have introduced bias?	Low risk
B. Concerns regarding applicability	
Are there concerns that the included patients and setting do not match the review question?	Low concern

All tests

A. Risk of Bias	
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes
If a threshold was used, was it pre-specified?	Yes
Could the conduct or interpretation of the index test have introduced bias?	Low risk
B. Concerns regarding applicability	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?	Low concern

Reference Standard

A. Risk of Bias	
Is the reference standard likely to correctly classify the target condition?	Yes

Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes
Could the reference standard, its conduct, or its interpretation have introduced bias?	Low risk
B. Concerns regarding applicability	
Are there concerns that the target condition as defined by the reference standard does not match the question?	Low concern

Flow and Timing

A. Risk of Bias	
Was there an appropriate interval between index test and reference standard?	Yes
Did all patients receive the same reference standard?	Yes
Were all patients included in the analysis?	Unclear
Could the patient flow have introduced bias?	Low risk

Pomorska, (n.d.)

Patient Sampling	Paper has not been published but is included as data was included in Falszewska et al. (2018).
Patient characteristics and setting	Not reported
Index test	CDS
Target condition and reference standard(s)	Target condition: Dehydration in AGE Reference standard: Difference between post-treatment weight and pre-treatment weight as a percent
Flow and timing	Not reported

Patient Selection

A. Risk of Bias	
Was a consecutive or random sample of patients enrolled?	Unclear
Was a case-control design avoided?	Unclear
Did the study avoid inappropriate exclusions?	Unclear
Could the selection of patients have introduced bias?	Unclear risk
B. Concerns regarding applicability	
Are there concerns that the included patients and setting do not match the review question?	Low concern

All tests

A. Risk of Bias	
Were the index test results interpreted without knowledge of the results of the reference standard?	

If a threshold was used, was it pre-specified?	Yes
Could the conduct or interpretation of the index test have introduced bias?	Low risk
B. Concerns regarding applicability	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?	Low concern

Reference Standard

A. Risk of Bias	
Is the reference standard likely to correctly classify the target condition?	Yes
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes
Could the reference standard, its conduct, or its interpretation have introduced bias?	Low concern
B. Concerns regarding applicability	
Are there concerns that the target condition as defined by the reference standard does not match the question?	Low concern

Flow and Timing

A. Risk of Bias	
Was there an appropriate interval between index test and reference standard?	
Did all patients receive the same reference standard?	
Were all patients included in the analysis?	No
Could the patient flow have introduced bias?	High risk

Characteristics of Reliability and Validity Studies
Bailey et al. (2010)

Characteristics of Study	
Methods	Prospective cohort
Participants	**Participants:** Pediatric patients aged one month to five years presenting at a pediatric ED with symptoms consistent with dehydration. **Setting:** Canada, Tertiary care pediatric Emergency Department, April 2008-March 2009 **Number enrolled into study:** *N* = 150 (Groups assigned by their CDS) • **Group 1,** No dehydration, CDS = 0**:** *n* = 56 • **Group 2,** Some dehydration, CDS = 1-4: *n* = 74 • **Group 3,** Moderate/severe dehydration, CDS = 5-8: *n* = 20 **Gender, males (as defined by researchers):** • **Group 1:** *n* = 27 (48%) • **Group 2:** *n* = 39 (53%) • **Group 3:** *n* = 12 (60%) **Race / ethnicity or nationality:** This study occurred in the Centre Hospitaller Universitaire Sainte-Justine (CHU Sainte-Justine), Montreal, Canada. The authors did not identify race or ethnicity of the participants. **Age, mean ± SD in months** • **Group 1:** 21 ± 15 • **Group 2:** 23 ± 14 • **Group 3:** 22 ± 11 **Inclusion criteria:** • Age 1 month to 5 years • Presented to the ED with vomiting and/or diarrhea • Patient assigned a CDS score at triage • Discharge ED diagnosis of gastroenteritis, enteritis, or gastritis **Exclusion criteria:** • Previous ED visit for the same illness in the 7 days prior to arrival • Diarrhea of more than 10 days • Patient left the ED without being seen by a physician • Cause of dehydration other than presumptive diagnosis of gastroenteritis • Chronic disease • Rehydrated with IV solution within previous 24 hours **Covariates identified:** Not reported
Interventions	Validation of CDS for children with gastroenteritis • Participating nurses attended training on proper use of CDS • These nurses used the CDS during triage of included patients
Outcomes	• Validation of the CDS

	• Association between CDS and LOS after being seen by primary physician • Association between CDS and total LOS • Need for IVF • Successful oral rehydration (ORT)
Notes	**Outcome 1:** Validation of the CDS-Study reports that the CDS is a good predictor of: • LOS in the ED after being seen by a physician • Perceived need for IV rehydration • Utilization of laboratory blood tests • Inter-rater agreement was determined to be "excellent" for the CDS based on Bland-Altman method: mean bias for scale validation was 0.5 minutes 95% CI [-2, 3] with upper and lower agreement limits of 11 minutes 95% CI [6, 16] and -10 minutes 95% CI [-15 to -5] respectively • CDS was perceived to have a strong association with assigned triage category • There was a trend toward failure of ORT as patients had a higher CDS **Outcome 2:** Association between CDS and LOS after being seen by primary physician, minutes [median (IQR) • Group 1: 54 (26-175) • Group 2: 128 (25-334) • Group 3: 425 (218-673) **Outcome 3:** Association between CDS and total LOS, minutes median, (IQR) • Group 1: 300 (189-456) • Group 2: 334 (182-480) • Group 3: 580 (304-860) **Outcome 4:** Need for IV fluids • Group 1: 5 (12%) • Group 2: 23 (32%) • Group 3: 13 (65%) **Outcome 5:** Successful ORT • Group 1: 17/19 (90%) • Group 2: 22/29 (76%) • Group 3: 5/10 (50%) **Notes:** • IV fluids were given to 41 patients, see Outcome 4 above • Oral rehydration was ordered by a physician in 58 patients • Oral rehydration was not counted if the nurse started ORT without a physician order

Friedman, et al. (2004)

Characteristics of Study	
Methods	Cohort
Participants	**Participants:** Children presenting to the ED with AGE **Setting:** Pediatric ED of The Hospital for Sick Children, Toronto, Canada **Number enrolled into study:** *N* = 141 **Number completed:** *N* = 102 **Gender, males:** (as defined by researchers) • *n* = 69 (50%) **Race / ethnicity or nationality** (as defined by researchers): • The study occurred in Canada. The authors did not identify race or ethnicity of the participants. **Age,** median in months, range • 18 (2-35) **Inclusion criteria:** • Subjects for whom the attending physician established the diagnosis of gastroenteritis with dehydration **Exclusion criteria:** • Chronic disease, such as renal, gastrointestinal, cystic fibrosis • Underlying malnutrition • Treatment with IV fluid within the past 24 hours **Covariates identified:** Not reported
Interventions	**Both:** • Attending physician examined the subject, and all therapy was carried out independent of the study • The attending physician were asked to record their assessment of dehydration that was present • Study nurses collected baseline data, and completed the CDS on all subjects prior to starting hydration therapy • Attending physician determined when therapy was complete, and subject was re-weighed by the study nurse
Outcomes	Primary outcome(s): • *Validity of CDS • *Reliability of CDS • *Discriminatory Power of CDS • *Responsiveness to Change of CDS *Outcomes of interest to the CMH CPG or CAT development team
Notes	**Results:** • *Validity of CDS ○ "Item-total correlation", or the correlation of the item with the total scale score, measured with Pearson's correlation coefficient, (Pett et al., 2003) was < .01 for each item • Reliability of CDS - Inter-rater reliability was calculated, and all items except "general appearance had an *ICC* > .6. Intra-class coefficient measures the consistency or agreement of values of the items within the respondents. *ICC* range is from 0-1, closer to 1 is desired (Pett et al., 2003)

	• Discriminatory Power of CDS- assesses by Ferguson's Delta was 0.83 • Responsiveness to Change of CDS- assessed with Wilcoxon' signed rank test was detected a change ($p < .01$) median score was 2 (*Range* = 0 to 8, $n = 126$) at baseline, and median score decreased to 0 (*Range* = 0 to 2, $n = 33$) following rehydration therapy

Goldman et al. (2008)

Characteristics of Study	
Methods	Prospective cohort
Participants	**Participants:** Pediatric patients aged one month to five years presenting to a pediatric Emergency Department (ED) with symptoms consistent with acute gastroenteritis Children with symptoms of acute gastroenteritis **Setting:** Tertiary care pediatric emergency department in Canada, January 2005 - May 2005. **Number enrolled into study:** *N* = 206 (Groups assigned by their Clinical Dehydration Score [CDS]) • **Group 1**, No Dehydration, CDS = 0: *n* = 117 • **Group 2**, Some Dehydration, CDS = 1-4: *n* = 84 • **Group 3**, Moderate/Severe Dehydration, CDS = 5-8: *n* = 5 **Number completed:** *N* = 206 • **Group 1:** *n* = 117 • **Group 2:** *n* = 84 • **Group 3:** *n* = 5 **Gender, males: (as defined by researchers)** • **Group 1:** *n* = 58 (50%) • **Group 2:** *n* = 43 (51%) • **Group 3:** *n* = 2 (40%) **Race / ethnicity or nationality (as defined by researchers):** • The study occurred in British Columbia Children's Hospital, Vancouver, Canada. The authors did not identify race or ethnicity of the participants. **Age, mean ± SD in months** • **Group 1:** 20.7 ± 13.6 • **Group 2:** 24.1 ± 15.9 • **Group 3:** 34.2 ± 21.2 **Inclusion criteria:** • Children 1 month to 5 years of age • Children with vomiting or diarrhea during the 24 hours before arrival to ED **Exclusion criteria:** • Children with diarrhea for >10 days • Any suspected cause of dehydration other than presumptive gastroenteritis • Chronic disease including coexisting malnutrition or failure to thrive • Recent intravenous fluids (IVF) within the previous 24 hours • ED visit for the same illness in the last 7 days **Covariates identified:** Not reported
Interventions	Validate CDS score
Outcomes	**Primary outcomes:** • Validate the CDS by assessing association with these outcomes*

	○ LOS* ○ Proportion of children receiving IV rehydration* ○ Proportion of children with abnormal serum pH values or bicarbonate levels **Safety outcome:** • Not reported *Outcomes of interest to the CMH CPG or CAT development team
Notes	**Results:** The three CDS categories were positively associated with the LOS, proportions of children who received IVF rehydration, and a trend towards positive association with the proportions of children with abnormal serum pH values or bicarbonate levels (but this was not statistically significant). See table.

Jauregui et al. (2014)

Characteristics of Study	
Methods	**Prospective, Non-consecutive Cohort**
Participants	**Participants: Children with AGE** **Setting: Emergency Department of a regional pediatric referral hospital, Rhode Island, USA** **Number enrolled into study:** *N* = 148 **Number completed:** *N* = 113 **Gender, males: (as defined by researchers)** • *n* = 51% **Race / ethnicity or nationality (as defined by researchers):** Only subjects who spoke English were enrolled **Age, median in years, range** • 6, (1 month - 18 years) **Inclusion criteria:** • ·All children ≤ 18 years of age • Chief complaint of vomiting and or diarrhea • Suspicion of dehydration by attending pediatric ED physician **Exclusion criteria:** • Positive pressure ventilation • Significant traumatic injury • Large volume fluid administration prior to enrollment • Surgical abdomen • Known congenital cardiac disease **Covariates identified:** Not reported
Interventions	**Both:** All subjects underwent physical examination, and physician gestalt estimation of level of dehydration was obtained. A standard form was then used to document the signs and symptoms observed, and the attending physician completed each the CDS, Gorelick 10-item scale, and WHO dehydration scale.
Outcomes	**Primary outcome(s):** • * Accuracy of the CDS, Gorelick and WHO scales compared to percent weight change **Secondary outcome(s)** • ·Accuracy of physician gestalt of dehydration compared to percent weight change **Safety outcome(s):** • Not reported
Notes	**Results:** • Based on weight change with rehydration: ○ Average 2.8% weight gain with rehydration

	○ Twelve patients had weight gain greater than 5% with rehydration, considered significant dehydration • *AUC* ○ CDS, *AUC* = .72, 95% CI [.6, .84], LR+ ○ Gorelick 10-item scale, *AUC* = .71, 95% CI [.57, .85] ○ WHO scale, *AUC* = .61, 95% CI [.45, .77] ○ Physician gestalt, *AUC* = .61, 95% CI [.44, 078] ▪ The CDS and Gorelick 10-item scale were significantly different from the reference line and were statistically predictors of dehydration. ▪ The cutoff for the CDS was 2 of 8, and the cutoff for the Gorelick 10-item scale was also 2 of 10. The WHO scale was not significantly different from the reference line

Kinlin and Freedman (2012)

Characteristics of Study	
Methods	**Cohort, prospective**
Participants	**Participants:** Children with acute gastroenteritis (AGE) and dehydration **Setting:** Urban Emergency Department, enrollment period December 2006 to April 2010 **Number enrolled into study:** *N* = 226 **Number completed:** *N* = 208 **Gender, males: (as defined by researchers)** • *n* = 51.8 **Race / ethnicity or nationality (as defined by researchers):** • The study occurred in Toronto, Canada. The authors did not identify race or ethnicity of the participants **Age, median, years, (IQR)** • 2.1 (1.36, 3.96) **Inclusion criteria:** • Diagnosis of AGE • Required intravenous therapy, as determined by attending physician • CDS ≥ 3 • Capillary refill time ≥ 2 seconds • Abnormal skin turgor, with prolonged retraction time and "tenting" or • Abnormal respiratory pattern for age in years **Exclusion criteria:** • Body weight < 5 kg • Significant underlying diseases (e.g. renal insufficiency, diabetes mellitus) • Suspicion of previously undiagnosed cardiac or renal disease • History of abdominal surgery • Acute surgical abdomen • History of head, chest or abdominal trauma within 7 days • Evidence of hypotension, hypo- or hyper-glycemia • Previous study enrollment **Covariates identified:** Not reported
Interventions	Both: • CDS assigned by the study nurse (trained in CDS assignment); attending physician was blinded to the CDS assignment by study nurse • CDS assigned by the attending physician (untrained in CDS assignment, but given directions) • Insertion of intravenous catheter • Laboratory tests: sodium, potassium, chloride, pH, bicarbonate, carbon dioxide, glucose, blood urea nitrogen, and creatinine • A study nurse reassessed and documented the CDS every 30 minutes for 4 hours
Outcomes	**Primary outcome(s):**

	• *Evaluate the interobserver reliability by using simultaneous, blinded assessments **Secondary outcome(s)** • *Report the association of the CDS with clinical data. ○ Construct validity: Gold standard was weight change, other comparisons were to number of diarrhea and vomiting episodes prior to presentation, respiratory rate, capillary refill time, serum bicarbonate, serum pH, attending physician's assessment of ready to discharge ○ Discriminative validity: The ability to discriminate between patients with and without sign/symptoms against the outcome of hospital admission ○ Responsiveness: how did the CDS perform as intravenous rehydration, a therapy known to treat dehydration was administered? *Outcomes of interest to the CMH CPG or CAT development team
Notes	**Results:** Interobserver reliability • Interobserver agreement between the study nurse and attending physician was moderate, κ = 0.52, 95% CI [.41, .63] ○ *K* coefficient for individual elements of the CDS ▪ Eyes, κ = .32, 95% CI [.18, .46] ▪ Mucous membranes, κ = .38, 95% CI [.26, .50] ▪ Tears, κ = .40, 95% CI [.27, .51] ▪ General appearance, κ = .49, 95% CI [.35, .60] ○ *K* coefficient by age group ▪ <36 months, κ = .51, 95% CI [.40, .65] ▪ ≥ 36 months, κ = .53, 95% CI [.27, .68] • Construct validity ○ CDS did not correlate with percent weight gain, r_s = -.04, 95% CI [-.19, .10] ○ CDS did correlate with serum bicarbonate value, r = -.35, 95% CI [-.46, -.23] ○ Analysis by age group < 36 months and ≥ 36 months did not show additional strong correlations • Discriminate validity ○ Optimal cut off point was ≥ to 5 ▪ Sensitivity 62% ▪ Specificity 66% ▪ Positive predictive value = 35%, 95% CI [25, 45], ▪ Negative predictive value = 85%, 95% CI [78, 91] ○ Discriminative ability did not differ when compared by age group • Responsiveness • CDS scores decreased over time, after therapy with intravenous fluid • Median scores at 2 and 4 hours were significantly lower than score at baseline (p < .001)

References

Allen, N.H. (2019). *Acute Gastroenteritis Data November 2018 to October 2019*. [Data file]. Retrieved from CMH internal data.

Bailey, B., Gravel, J., Goldman, R. D., Friedman, J. N., & Parkin, P. C. (2010). External validation of the clinical dehydration scale for children with acute gastroenteritis. *Acad Emerg Med, 17*(6), 583-588. doi:10.1111/j.1553-2712.2010.00767.x

Colletti, J. E., Brown, K. M., Sharieff, G. Q., Barata, I. A., Ishimine, P., & Committee, A. P. E. M. (2010). The management of children with gastroenteritis and dehydration in the emergency department. *J Emerg Med, 38*(5), 686-698. doi:10.1016/j.jemermed.2008.06.015

Falszewska, A., Dziechciarz, P., & Szajewska, H. (2014). The diagnostic accuracy of Clinical Dehydration Scale in identifying dehydration in children with acute gastroenteritis: a systematic review. *Clin Pediatr (Phila), 53*(12), 1181-1188. doi:10.1177/0009922814538493

Falszewska, A., Dziechciarz, P., & Szajewska, H. (2017). Diagnostic accuracy of clinical dehydration scales in children. *Eur J Pediatr, 176*(8), 1021-1026. doi:10.1007/s00431-017-2942-8

Falszewska, A., Szajewska, H., & Dziechciarz, P. (2018). Diagnostic accuracy of three clinical dehydration scales: a systematic review. *Arch Dis Child, 103*(4), 383-388. doi:10.1136/archdischild-2017-313762

Freedman, S. B., Adler, M., Seshadri, R., & Powell, E. C. (2006). Oral ondansetron for gastroenteritis in a pediatric emergency department. *N Engl J Med, 354*(16), 1698-1705. doi:10.1056/NEJMoa055119

Freedman, S. B., DeGroot, J. M., & Parkin, P. C. (2014). Successful discharge of children with gastroenteritis requiring intravenous rehydration. *J Emerg Med, 46*(1), 9-20. doi:10.1016/j.jemermed.2013.04.044

Freedman, S. B., Pasichnyk, D., Black, K. J., Fitzpatrick, E., Gouin, S., Milne, A., . . . Pediatric Emergency Research Canada Gastroenteritis Study, G. (2015). Gastroenteritis Therapies in Developed Countries: Systematic Review and Meta-Analysis. *PloS one, 10*(6), e0128754. doi:10.1371/journal.pone.0128754

Friedman, J. N., Goldman, R. D., Srivastava, R., & Parkin, P. C. (2004). Development of a clinical dehydration scale for use in children between 1 and 36 months of age. *J Pediatr, 145*(2), 201-207. doi:10.1016/j.jpeds.2004.05.035

Geurts, D., Steyerberg, E. W., Moll, H., & Oostenbrink, R. (2017). How to Predict Oral Rehydration Failure in Children With Gastroenteritis. *J Pediatr Gastroenterol Nutr, 65*(5), 503-508. doi:10.1097/MPG.0000000000001556

Goldman, R. D., Friedman, J. N., & Parkin, P. C. (2008). Validation of the clinical dehydration scale for children with acute gastroenteritis. *Pediatrics, 122*(3), 545-549. doi:10.1542/peds.2007-3141

Gorelick, M. H., Shaw, K. N., & Murphy, K. O. (1997). Validity and reliability of clinical signs in the diagnosis of dehydration in children. *Pediatrics, 99*(5), E6. doi:10.1542/peds.99.5.e6

Gravel, J., Manzano, S., Guimont, C., Lacroix, L., Gervaix, A., & Bailey, B. (2010). [Multicenter validation of the clinical dehydration scale for children]. *Arch Pediatr, 17*(12), 1645-1651. doi:10.1016/j.arcped.2010.09.009

Guarino, A., Ashkenazi, S., Gendrel, D., Lo Vecchio, A., Shamir, R., Szajewska, H., . . . European Society for Pediatric Infectious, D. (2014). European Society for Pediatric Gastroenterology, Hepatology, and Nutrition/European Society for Pediatric Infectious Diseases evidence-based guidelines for the management of acute gastroenteritis in children in Europe: update 2014. *J Pediatr Gastroenterol Nutr, 59*(1), 132-152. doi:10.1097/MPG.0000000000000375

Hayajneh, W. A., Jdaitawi, H., Al Shurman, A., & Hayajneh, Y. A. (2010). Comparison of clinical associations and laboratory abnormalities in children with moderate and severe dehydration. *J Pediatr Gastroenterol Nutr, 50*(3), 290-294. doi:10.1097/MPG.0b013e31819de85d

Hoxha, T., Xhelili, L., Azemi, M., Avdiu, M., Ismaili-Jaha, V., Efendija-Beqa, U., & Grajcevci-Uka, V. (2015). Performance of clinical signs in the diagnosis of dehydration in children with acute gastroenteritis. *Med Arch, 69*(1), 10-12. doi:10.5455/medarh.2015.69.10-12

Hoxha, T. F., Azemi, M., Avdiu, M., Ismaili-Jaha, V., Grajqevci, V., & Petrela, E. (2014). The usefulness of clinical and laboratory parameters for predicting severity of dehydration in children with acute gastroenteritis. *Med Arch, 68*(5), 304-307. doi:10.5455/medarh.2014.68.304-307

Jauregui, J., Nelson, D., Choo, E., Stearns, B., Levine, A. C., Liebmann, O., & Shah, S. P. (2014). External validation and comparison of three pediatric clinical dehydration scales. *PloS one, 9*(5), e95739. doi:10.1371/journal.pone.0095739

Kanjanaphan, T., & Amornchaicharoensuk, Y. (2018). Comparison of the accuracy of the Clinical Dehydration Scale and Gorelick 10-point scale versus pre- and post- hydration body weight among children with acute gastroenteritis. *Southeast Asian J Trop Med Publid Health, 49*(4).

Kim, H. S., Rotundo, L., Nasereddin, T., Ike, A., Song, D., Babar, A., . . . Ahlawat, S. K. (2017). Time trends and predictors of acute gastroenteritis in the United States. *J Clin Gastroenterol, 51*. doi:10.1097/MCG.0000000000000907

King, C. K., Glass, R., Bresee, J. S., Duggan, C., Centers for Disease, C., & Prevention. (2003). Managing acute gastroenteritis among children: oral rehydration, maintenance, and nutritional therapy. *MMWR Recomm Rep, 52*(RR-16), 1-16. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/14627948

Kinlin, L. M., & Freedman, S. B. (2012). Evaluation of a clinical dehydration scale in children requiring intravenous rehydration. *Pediatrics, 129*(5), e1211-1219. doi:10.1542/peds.2011-2985

Koo, T. K., & Li, M. Y. (2017). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J Chiropr Med, 16*(4).

Kuge, R., Morikawa, Y., & Hasegawa, Y. (2017). Uric acid and dehydration in children with gastroenteritis. *Pediatr Int, 59*(11), 1151-1156. doi:10.1111/ped.13366

Levine, A. C., Munyaneza, R. M., Glavis-Bloom, J., Redditt, V., Cockrell, H. C., Kalimba, B., . . . Drobac, P. C. (2013). Prediction of severe disease in children with diarrhea in a resource-limited setting. *PloS one, 8*(12), e82386. doi:10.1371/journal.pone.0082386

Levine, A. C., Shah, S. P., Umulisa, I., Munyaneza, R. B., Dushimiyimana, J. M., Stegmann, K., . . . Noble, V. E. (2010). Ultrasound assessment of severe dehydration in children with diarrhea and vomiting. *Acad Emerg Med, 17*(10), 1035-1041. doi:10.1111/j.1553-2712.2010.00830.x

McHugh, M. L. (2012). Interrater reliability: the kappa statistic. *Biochem Med (Zagreb), 22*(3), 276-282. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/23092060

Nordenstrom, J. (2007). *Evidence-Based Medicine in Sherlock Holmes' Footsteps*. Malden, Massachusetts Blackwell Publishing.

Parkin, P. C., Macarthur, C., Khambalia, A., Goldman, R. D., & Friedman, J. N. (2010). Clinical and laboratory assessment of dehydration severity in children with acute gastroenteritis. *Clin Pediatr (Phila), 49*(3), 235-239. doi:10.1177/0009922809336670

Pett, M. A., Lackey, N. R., & J.J., S. (2003). *Making Sense of Factor Analysis: the Use of Factor Analysis for Instrument Development in Health Care Research*. Thousand Oaks, CA USA: Sage Publications, Inc.

Pomorska, D., Dziechciarz, P., & Szajewska, H. (n.d.). Assessment of the accuracy of the Clnicial Dehydration Scale (CDS).

Powell, C. V., Priestley, S. J., Young, S., & Heine, R. G. (2011). Randomized clinical trial of rapid versus 24-hour rehydration for children with acute gastroenteritis. *Pediatrics, 128*(4), e771-778. doi:10.1542/peds.2010-2483

Price, P., Jhangiani, R., & Chiang, I. (2015). Research Methods in Psychology -- 2nd Canadian Edition. In. Victoria, B.C.: BCcampus.

Pringle, K., Shah, S. P., Umulisa, I., Mark Munyaneza, R. B., Dushimiyimana, J. M., Stegmann, K., . . . Levine, A. C. (2011). Comparing the accuracy of the three popular clinical dehydration scales in children with diarrhea. *Int J Emerg Med, 4*, 58. doi:10.1186/1865-1380-4-58

Shavit, I., Brant, R., Nijssen-Jordan, C., Galbraith, R., & Johnson, D. W. (2006). A novel imaging technique to measure capillary-refill time: improving diagnostic accuracy for dehydration in young children with gastroenteritis. *Pediatrics, 118*(6), 2402-2408. doi:10.1542/peds.2006-1108

Steiner, M. J., Nager, A. L., & Wang, V. J. (2007). Urine specific gravity and other urinary indices: inaccurate tests for dehydration. *Pediatr Emerg Care, 23*(5), 298-303. doi:10.1097/01.pec.0000270162.76453.fa

Tam, R. K., Wong, H., Plint, A., Lepage, N., & Filler, G. (2014). Comparison of clinical and biochemical markers of dehydration with the clinical dehydration scale in children: a case comparison trial. *BMC Pediatr, 14*, 149. doi:10.1186/1471-2431-14-149

Vega, R. M., & Bhimji, S. S. (2018). Dehydration, Pediatric. In *StatPearls*. Treasure Island (FL).