



Isha Kaul, MD¹, Vincent S. Staggs, PhD^{2,3}, Amber Bagherian, MS⁴, Amna Ali, BS⁵, Robert J. Shulman, MD^{1,6}, Jennifer M. Colombo, MD^{3,4}, Jennifer V. Schurman, PhD^{3,4}, Bruno P. Chumpitazi, MD, MPH, FACG^{1,6}, Craig A. Friesen, MD^{3,4}

¹Section of Pediatric Gastroenterology, Hepatology & Nutrition, Department of Pediatrics, Baylor College of Medicine, Houston, Texas. ²Biostatistics & Epidemiology, Health Services & Outcomes Research, Children's Mercy Kansas City, Kansas City, Missouri. ³Department of Pediatrics, University of Missouri – Kansas City School of Medicine, Kansas City, Missouri. ⁴Division of Pediatric Gastroenterology, Children's Mercy Kansas City, Kansas City, Missouri. ⁵The Wiess School of Natural Sciences, Rice University, Houston, Texas. ⁶Children's Nutrition Research Center, Agriculture Research Service, United States Department of Agriculture, Houston, Texas.



Background

- Studies are needed to determine the utility of the recently published pediatric Rome IV criteria for diagnosing functional gastrointestinal disorders (FGIDs).
- OBJECTIVES:**
 - (1) to determine the inter- and intra-rater agreement among pediatric gastroenterologists and gastroenterologists in training in the choice of pediatric child/adolescent Rome IV diagnosis(es).
 - (2) to evaluate agreement and variability in diagnostic testing choices among these same physicians in the process of diagnosing a pediatric FGID.

Methods

- Pediatric gastroenterologists and fellows at two medical centers completed a survey containing clinical FGID vignettes.
- For each vignette, raters identified the most likely diagnosis(es) and selected which diagnostic test(s) (if any) they typically would obtain.
- The survey was re-administered within 3 months.
- Inter-rater and intra-rater weighted percent agreement was determined.
- Linear mixed modeling identified sources of variability in diagnostic testing.

Results

- Thirty-four (21 faculty, 13 fellows) raters completed the initial survey and thirty-one (18 faculty, 13 fellows) raters completed the repeat survey.

Table 1: Number of Diagnostic Tests Selected by Physicians (n = 31) For Each Clinical Vignette During an Initial and Repeat Survey

Initial Survey											
Vignette	1	2	3	4	5	6	7	8	9	10	Overall
Median	1	0	5	4	4	3	3	2	1	5	2
Inter-quartile range	0-5	0-0	1.5-6.5	0.5-6.5	1-5	0-6	1-7	0-5	0-4.5	2-6.5	0-6
Repeat Survey											
Vignette	1	2	3	4	5	6	7	8	9	10	Overall
Median	1	0	4	2	2	1	3	1	1	2	1
Inter-quartile range	0-4	0-0	1-5	0-5	1-5.5	0-5.5	1-5.5	0-5	0-3.5	1-5.5	0-5

Results

Figure 1: Rome IV Diagnosis Inter-rater Agreement Percentage at Two Time Points

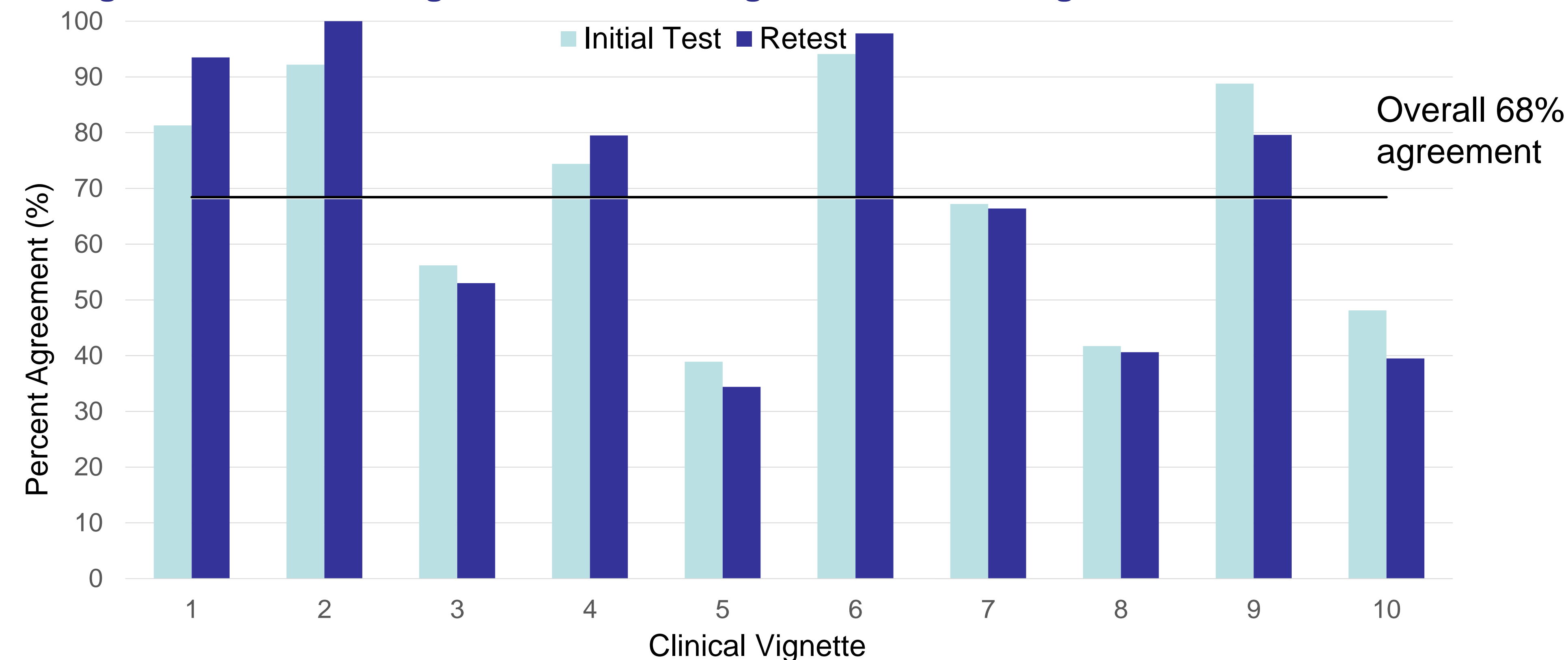


Figure 2: Diagnostic Evaluation Testing Inter-rater Agreement Percentage at Two Time Points

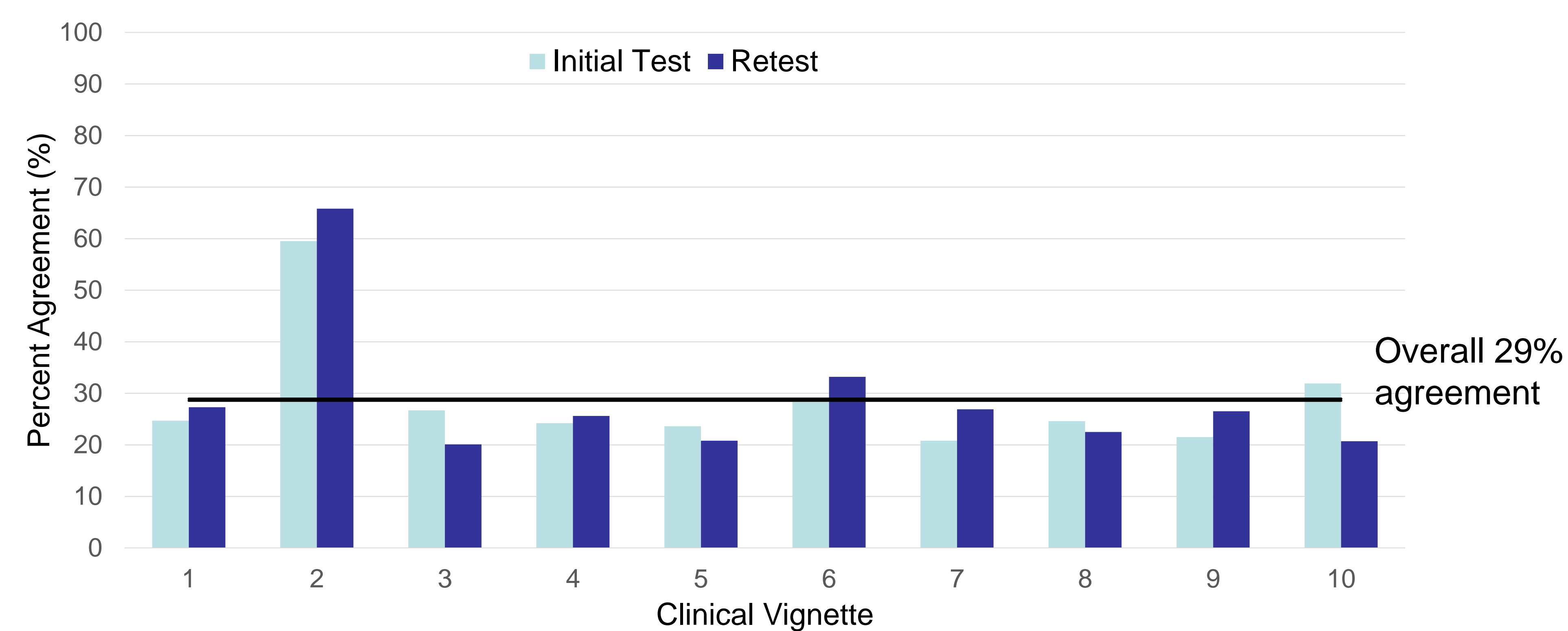


Table 2: Rome IV Diagnosis and Diagnostic Evaluation Testing Intra-Rater Agreement

Rome IV Diagnosis Intra-Rater Agreement											
Vignette	1	2	3	4	5	6	7	8	9	10	Overall
Faculty (n=18)	88.9	69.4	92.6	73.1	81.5	56.5	98.1	78.7	72.2	92.6	80.4
Fellows (n=13)	82.1	56.4	100	61.5	71.8	47.4	100	66.7	32.1	82.1	70.0
Combined	86.0	64.0	95.7	68.3	77.4	52.7	98.9	75.8	55.4	88.2	76.2
Diagnostic Evaluation Testing Intra-Rater Agreement											
Vignette	1	2	3	4	5	6	7	8	9	10	Overall
Faculty (n=18)	59.4	68.5	68.9	59.2	58.6	72.2	55.3	60.9	79.6	69.1	65.2
Fellows (n=13)	51.9	69.2	56.3	39.6	38.9	36.9	42.1	47.1	39.5	38.5	46.0
Combined	56.3	68.6	63.6	51.0	50.4	57.4	49.8	55.1	62.8	56.3	57.1

Results (continued)

Table 3: Linear Mixed Modeling Estimates with Sources of Variability in Evaluation Testing for Ten Clinical Vignettes

Variable	Level	Estimate	95% CI	P value
Survey Timing	Repeat	-0.5	-0.8, -0.1	0.014
	Initial	Referent		
Center	A	1.8	0.2, 3.4	0.027
	B	Referent		
Physician type	Fellow, 1st year	-0.1	-2.4, 2.1	0.913
	Fellow, 2nd year	-0.2	-2.3, 1.8	0.813
	Fellow, 3rd year	-0.5	-2.7, 1.8	0.677
	Faculty	Referent		
Familiar with Rome	Yes	0.2	-0.9, 1.3	0.742
	No	Referent		
Use criteria in practice	Yes	-1.1	-2.0, -0.2	0.015
	No	Referent		

- Between-physician differences accounted for 43% of the variability in the number of tests selected.
- Rater identified use of Rome criteria in clinical practice was associated with 1.1 fewer diagnostic tests on average.
- Higher intra-rater agreement was noted for diagnostic testing in faculty when compared to fellows ($p = 0.009$).

Conclusions

- In a multicenter evaluation amongst pediatric gastroenterology physicians, pediatric Rome IV diagnostic agreement was higher than that reported for previous Rome versions, and higher than agreement on diagnostic testing.

References

- Hyams JS, Di Lorenzo C, Saps M, et al. Functional Disorders: Children and Adolescents. *Gastroenterology* 2016; 150(6): 1456-1468.
- Koppen I, Nurko S, Saps M, et al. The pediatric Rome IV criteria: what's new? *Expert Review of Gastroenterology & Hepatology*. Mar 2017; 11(3):193-201.
- Rouster AS, Karpinski AC, Silver D, et al. Functional gastrointestinal disorders dominate pediatric gastroenterology outpatient practice. *Jun 2016; 62(6):847-51.*