

PEDIATRIC ROME IV DIAGNOSIS AGREEMENT IS GREATER THAN AGREEMENT ON DIAGNOSTIC TESTING

Isha Kaul¹, Vincent S Staggs², Amber Bagherian³, Amna Ali⁴, Robert J Shulman⁵, Jennifer M Colombo⁶, Jennifer V Schurman⁶, Bruno P Chumpitazi⁵, Craig A Friesen⁶

¹ Baylor College of Medicine, Department of Pediatrics, Gastroenterology, Hepatology & Nutrition

² Children's Mercy Kansas City, Biostatistics & Epidemiology, Health Services & Outcomes Research, Pediatrics

³ Children's Mercy Kansas City, Division of Pediatric Gastroenterology, Pediatric Gastroenterology

⁴ Rice University, The Wiess School of Natural Sciences, Natural Sciences

⁵ Baylor College of Medicine, Section of Pediatric Gastroenterology, Hepatology & Nutrition, Pediatrics

⁶ Children's Mercy Kansas City, Division of Pediatric Gastroenterology, Pediatrics

Keywords: "child", "abdominal pain", "diagnosis", "irritable bowel syndrome", "dyspepsia"

Background: Pediatric Rome IV criteria are used to diagnose childhood functional gastrointestinal disorders (FGIDs). This study of pediatric gastroenterology physicians measured their agreement in: 1) Making a pediatric Rome IV FGID diagnosis; and 2) Diagnostic testing for patients with FGIDs.

Materials/Methods: Pediatric gastroenterologists and pediatric gastroenterology fellows at two medical centers completed a survey containing clinical FGID vignettes. For each vignette, raters identified the most likely Rome IV 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 raters completed the initial survey of whom thirty-one (91%) completed the repeat survey. Overall inter-rater agreement on Rome IV diagnoses was 68% for initial and repeat surveys whereas intra-rater agreement was 76%. In contrast, overall inter-rater agreement on diagnostic testing was <30% for both initial and repeat surveys and intra-rater agreement was only 57%. 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 (95% CI 0.2-2.0, $p = 0.015$). 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.

Images / Graph / Table: No image uploaded