## Aalborg Universitet



#### **Technical report**

Inter- and intra-rater reliability of regional gastrointestinal transit times measured using the 3D-Transit electromagnet tracking system Kalsi, G K; Grønlund, D; Martin, J; Drewes, A M; Scott, S M; Birch, M J

Published in: Neurogastroenterology and Motility

DOI (link to publication from Publisher): 10.1111/nmo.13396

Publication date: 2018

Document Version Accepted author manuscript, peer reviewed version

Link to publication from Aalborg University

*Citation for published version (APA):* Kalsi, G. K., Grønlund, D., Martin, J., Drewes, A. M., Scott, S. M., & Birch, M. J. (2018). Technical report: Inter-and intra-rater reliability of regional gastrointestinal transit times measured using the 3D-Transit electromagnet tracking system. Neurogastroenterology and Motility, 30(11), [e13396]. https://doi.org/10.1111/nmo.13396

#### **General rights**

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
  You may not further distribute the material or use it for any profit-making activity or commercial gain
  You may freely distribute the URL identifying the publication in the public portal -

#### Take down policy

If you believe that this document breaches copyright please contact us at vbn@aub.aau.dk providing details, and we will remove access to the work immediately and investigate your claim.

This is the peer reviewed version of the following article: Kalsi, GK, Grønlund, D, Martin, J, Drewes, AM, Scott, SM, Birch, MJ. Technical report: Inter- and intra-rater reliability of regional gastrointestinal transit times measured using the 3D-Transit electromagnet tracking system. Neurogastroenterol Motil. 2018; 30:e13396, which has been published in final form at https://doi.org/10.1111/nmo.13396. This article may be used for non-commercial purposes in accordance with Wiley Terms and Conditions for Use of Self-Archived Versions.

# 1 Title

- 2 Technical report: Inter- and intra-rater reliability of regional gastrointestinal transit
- 3 times measured using the 3D-Transit electromagnet tracking system

# 4 Authors and affiliations

5 6	G.K. Kalsi <sup>1,2</sup> , D. Grønlund <sup>3,4</sup> , J. Martin <sup>2</sup> , A. M. Drewes <sup>3,4</sup> , S. M. Scott <sup>1</sup> , M. J. Birch <sup>1,2</sup>
7	<sup>1</sup> GI Physiology Unit, The Blizard Institute, Barts and the London School of Medicine and Dentistry,
8	Queen Mary University of London, London, UK
9	
10 11	<sup>2</sup> Clinical Physics, Barts Health NHS Trust, The Royal London Hospital, London, UK
12	<sup>3</sup> Mech-Sense, Department of Gastroenterology and Hepatology, Aalborg University Hospital,
13	Aalborg, Denmark
14	
15	<sup>4</sup> Department of Clinical Medicine, Aalborg University, Aalborg, Denmark
16	
17	
18	
19	Corresponding Author
20	
21	Gursharan Kaur Kalsi
22	Clinical Physics, Barts Health NHS Trust, The Royal London Hospital, 56-76 Ashfield Street,
23	Whitechapel, London. E1 1BB
24	Email: <u>g.k.kalsi@qmul.ac.uk</u>

- 25
- 26

27 Abstract

29	Background: The 3D-Transit electromagnet tracking system is an emerging tool for the ambulatory
30	assessment of gastrointestinal (GI) transit times and motility patterns, based on the anatomical
31	localization of ingestible electromagnetic capsules. Currently, 3D-Transit recordings are manually
32	analyzed to extract GI transit times. As this is a subjective method, there is some inherent
33	variability in the measurements, which may be experience-dependent. We therefore assessed
34	inter- and intra-rater reliability of GI transit times from 3D-Transit recordings.
35	Methods: Thirty-six 3D-Transit recordings (17 female; median age: 34 years (range: 21–80)) were
36	analyzed twice by 3 raters with varying experience. Each rater manually identified the timestamps
37	when a capsule progressed from antrum to duodenum, and from ileum to right colon. These
38	timestamps, along with the ingestion and expulsion times were used to determine whole gut
39	(WGTT), gastric emptying (GET), small intestinal (SITT) and colonic (CTT) transit times. Reliability
40	was determined using interclass correlation coefficient (ICC).
41	Key Results: For capsule progression timestamps, the most and mid-experienced raters had fair to
42	good inter- and excellent intra-rater reliability (ICC <sub>min-max</sub> =0.61-1.00), whereas the inexperienced
43	rater had poor to fair inter- and poor intra-rater reliability (ICC <sub>min-max</sub> =0.28-0.55). GET and SITT
44	reliability between the most and mid-experienced raters was fair (ICC $_{min-max}$ =0.61-0.73), while
45	reliability between these raters and the inexperienced rater was poor to fair (ICC $_{min-max}$ =0.28-0.55).
46	CTT reliability was excellent between and within all raters (ICC <sub>min-max</sub> =0.92-0.99).
47	Conclusions & Inferences: Inexperienced raters provide the least reliable measurements from 3D-
48	Transit recordings, which confirms requirement for adequate training. Automation may improve
49	reliability of measurements.

# 50 Key Points

51	• The 3D-Transit System can aid the diagnostic evaluation of gastrointestinal disorders. We
52	assessed the reliability of regional GI transit times measured by experienced and inexperienced
53	raters.
54	• Reliability of gastric emptying and small intestinal transit time was fair between the most and
55	mid-experienced raters but poor for the inexperienced rater. Whole gut and colonic transit
56	time reliability was excellent across all raters.
57	Inexperienced raters require adequate training to provide reliable measurements of GI transit
58	times from the 3D-Transit System.
59 60	Key Words
61	3D-Transit system, electromagnetic capsule, gastrointestinal, reliability, transit time
62 63	Abbreviations
64	GI: gastrointestinal; cpm: contractions per minute; WGTT: whole gut transit time; GET: gastric
65	emptying; SITT: small intestinal transit time; CTT: colonic transit time; ICC: intraclass correlation

66 coefficient; CI: confidence interval.

### 67 Introduction

68

The 3D-Transit electromagnet tracking system (Motilis Medica, SA, Lausanne, Switzerland) is a novel and minimally invasive tool for the ambulatory evaluation of total and regional gastrointestinal (GI) transit times and motility patterns. It accurately tracks and measures the position and orientation of up to three ingestible electromagnetic capsules from ingestion to expulsion using an external detector plate positioned over the abdomen.<sup>1-4</sup>

74

Total GI transit time is easily extracted from a 3D-Transit recording, as the signal start and end points indicate capsule ingestion and expulsion times. For regional GI transit times however, the timestamps when a capsule progresses from the stomach into the duodenum, and from the ileum into the right colon are manually identified by visually observing changes in the capsule's orientation angles, which reflect GI contractile activity, along with shifts in its position as it progresses from one GI region to the next.<sup>1,5</sup>

81

82 The system was originally developed using a stationary detector matrix which required subjects to 83 stay relatively immobile for long periods of time in a controlled laboratory environment, thus reducing the effects of external movement artefacts.<sup>6-9</sup> Accordingly, inter-rater variability in capsule 84 progression timestamps, and thereby GI transit times, has been reported as low.<sup>6,8</sup> The principle 85 86 advantage of the ambulatory system is that it enables continuous monitoring of GI motility under 87 physiological conditions; however, subject ambulation renders it susceptible to external magnetic 88 fields and motion artifacts, making it more difficult to identify capsule progression timestamps. 89 Hence, the accuracy in identifying these timestamps is not only dependent on the quality of the

90 recording but also on the ability of the rater to distinguish artifacts from real movements of the 91 capsule.

92

93 Recently, the inter-variability of GI transit time measurements was assessed by two experienced raters who analyzed 20 3D-Transit recordings.<sup>1</sup> Differences in regional GI transit times were 94 95 reported in 8 of the 20 recordings (40%); however, these differences were considered acceptable 96 by the authors, as the overall median difference was zero minutes.<sup>1</sup> Nevertheless, there is a need 97 to determine the level of reliability of measurements, particularly when raters are blinded to their 98 own and each other's results. Therefore, the primary aim of this study was to assess inter- and intra-99 rater reliability of capsule progression timestamps, and hence regional GI transit times. A secondary 100 aim was to assess how the experience of the rater influences the identification of these timestamps.

101

#### 102 Materials & Methods

103

#### 104 **3D-Transit recording selection**

3D-Transit recordings were selected from a database of healthy volunteer studies conducted at the Neurogastroenterology Unit at Aarhus University Hospital (Aarhus, Denmark), Department of Gastroenterology and Hepatology at Aalborg University Hospital (Aalborg, Denmark) and the GI Physiology Unit at Queen Mary University (London, UK) between March 2012 and February 2016. In these studies, healthy volunteers swallowed up to three capsules, each taken a day apart after an overnight fast. Recordings were selected if they were complete with clear ingestion and expulsion points. For studies where volunteers ingested more than one capsule, only one recording was selected irrespective of capsule number. Poor quality recordings or recordings with more than 2
hours of missing data were excluded from the study. From this, a sample of 36 3D-Transit recordings
were randomly selected (17 female; median age: 34 years (range: 21–80)), 12 from each research
center.

116

#### 117 Data Collection

Three independent raters with varying experience of analyzing 3D-Transit recordings participated in the study. Rater experience was based on the number of previously analyzed recordings as follows: ≥100 recordings: most experienced (rater 1); approximately 40 recordings: midexperienced (rater 2); <5 recordings: least experienced (rater 3). All raters were prescribed written instructions on analyzing 3D-Transit recordings (dated May 2017) and the 3D-Transit System Instructions for Use (dated September 2014).

124

Recordings were analyzed using the 3D-Transit software, version 0.4 (Motilis Medica, SA, Lausanne, Switzerland). This involved identifying four timestamps as described by Haase et al. (2014)<sup>1</sup>: (i) ingestion: start of recording; (ii) duodenum: capsule's progression from the stomach into the duodenum; (iii) right colon: capsule's progression from the distal ileum to the caecum; (iv) expulsion: end of recording indicated by a loss of signal. For intra-rater reliability, each rater analyzed the 36 recordings twice with a minimum period of two weeks between repeat analyses.

131

#### 132 Data Analysis

133 The timestamps were used to determine WGTT (whole gut transit time; time between capsule 134 ingestion and expulsion), GET (gastric emptying; time between ingestion and passage into the duodenum), SITT (small intestinal transit time; time between the duodenum and right colon timestamps) and CTT (colonic transit time; time between the right colon timestamp and capsule expulsion). Transit times were automatically extracted from the 3D-Transit software and exported as text files for inter- and intra-rater comparison.

139

#### 140 Statistical Analysis

141 To calculate inter- and intra-rater reliability of the duodenum and right colon timestamps and 142 regional transit times, the ICCs (intraclass correlation coefficients) and their 95% confidence 143 intervals (CIs) were calculated based on a single rating, absolute agreement, 2-way random-effects 144 model. ICC values range between 0 and 1 with a higher value indicating better reliability (<0.5, poor; 0.5-0.75, fair; 0.75-0.9, good; >0.9, excellent).<sup>10</sup> The timestamps were subtracted from the ingestion 145 146 timestamp to convert the data into hours for the ICCs to be determined. Scatterplots, means and 147 95% CI were used to illustrate and compare GI transit times within and between raters. All statistical 148 analyses were performed using SPSS Statistics Version 25 (IBM, New York, USA).

149

#### 150 **Results**

151

#### 152 Inter-rater reliability of duodenum and right colon timestamps

Between raters, the degree of inter-rater reliability of both the duodenum and right colon timestamps was poor, with the ICC ranging between 0.42 and 0.47 (95% CI = 0.24-0.63). However, when comparing the most and mid-experienced raters, the reliability of both timestamps was fair to good. Reliability between raters 1 (most-experienced) and 2 (mid-experienced) against rater 3 (least-experienced) was poor to fair for the duodenum timestamp and very poor for the right colontimestamp (Table 1).

159

#### 160 Intra-rater reliability of duodenum and right colon timestamps

- 161 Intra-rater reliability of both timestamps was good to excellent for raters 1 and 2 with the ICC
- 162 ranging between 0.89 and 1.00 (95% CI = 0.79-1.00). However, reliability of these timestamps was
- 163 poor for rater 3 (Table 1).
- 164

#### 165 Inter-rater reliability of regional GI transit times

Scatterplots for inter-rater reliability of whole gut and regional GI transit times are presented in Figure 1. GET and SITT reliability between all raters was low, supported by poor ICCs ranging between 0.41 and 0.47 (95% CI = 0.25-0.63), while reliability of CTT was excellent (Table 2). ICC values for GET and SITT were consistently fair between raters 1 and 2, while reliability between

- 170 these raters and rater 3 was poor. WGTT reliability was excellent across all raters.
- 171

#### 172 Intra-rater reliability of regional GI transit times

- 173 Scatterplots for intra-rater reliability are presented in Figure 2. For raters 1 and 2, good to excellent
- 174 intra-rater reliability was seen for GET and SITT, with ICC values ranging between 0.84 and 1.00 (95%
- 175 CI = 0.71-1.00), while reliability was poor for rater 3 (ICC = 0.20-0.48, 95% CI = -0.14-0.71) (Table 2).
- 176 CTT and WGTT reliability was excellent for all raters.

178 **Discussion** 

179

180 We assessed the inter- and intra-rater reliability of regional GI transit times based on the manual 181 identification of the duodenum and right colon capsule progression timestamps in 3D-Transit 182 recordings. Our results showed that the inter- and intra-rater reliability of both timestamps is 183 generally fair to excellent amongst the most and mid-experienced raters and as expected, poor in 184 an inexperienced rater. This explains the fair inter-rater, and good to excellent intra-rater reliability 185 of GET and SITT seen amongst the more and mid experienced raters. However, reliability of these 186 transit times was poor in the inexperienced rater, indicating a need for an adequate period of 187 training.

188

189 Surprisingly, the inter- and intra-rater reliability of CTT was excellent amongst all raters. This was 190 unexpected, as the CTT is dependent on the right colon timestamp, the reliability of which was poor 191 in the inexperienced rater. Furthermore, general consensus amongst all raters was that the right 192 colon timestamp was subjectively more difficult to identify than the duodenum. However, this may 193 be explained by examining the magnitudes of the measurements. CTT is approximately eight times 194 longer than GET, and four times longer than SITT; therefore, the CTT measurement is less sensitive 195 to the uncertainty in the right colon timestamp due to its large magnitude and a fixed capsule 196 expulsion timestamp.

197

198 Nevertheless, it is apparent that the manual method of analyzing 3D-Transit recordings is not 199 optimal, even amongst experienced raters who only showed fair inter-rater reliability for GET and 200 SITT. Furthermore, the reliability of GI transit times was assessed using good quality recordings. Poorer quality recordings, which are difficult to interpret due to the increased presence of artifacts may produce less reliable measurements. Therefore, there is a need to improve the current methodology to obtain better estimates of GI transit times. This may be achieved through automation by using artifact rejection algorithms and pattern-recognition techniques to better detect the various gut contraction frequencies and hence, the capsule progression timestamps.

206

207 In conclusion, we assessed the inter- and intra-rater reliability of GI transit times as measured using 208 the 3D-Transit system. Reliability was generally fair between experienced raters. An inexperienced 209 rater provided the least reliable results, indicating a need for adequate training. Automation may 210 improve reliability of the method.

211

# 212 Acknowledgements

- We acknowledge the assistance of Dr. Jonathan Reeves for the selection of the 3D-Transitrecordings.
- 215

216	Fun	ding
-	-	- 0

- 217 Researchers were indirectly funded from their research and training positions.
- 218
- 219

```
220 Disclosure
```

- 221
- 222 The authors have no competing interests.
- 223
- 224
- 225

226

# 227 Author Contribution

228 229

230 GKK and MJB designed the study. GKK, DG and JM collected the data. GKK and DG analyzed tl				
2.30 GNN and IVIJB designed the study. GNN. DG and JIVI collected the data. GNN and DG analyzed th	220	CKK and MID designed the study	CKK DC and IM collected the data	CKK and DC analyzed the
	230	GKK and IVIJB designed the study.	GKK, DG and JIVI Collected the data.	GKK and DG analyzed the

- 231 data and drafted the initial manuscript. GKK completed and finalized the manuscript. AMD, SMS
- and MJB reviewed and approved the final manuscript as submitted. All authors have approved this
- 233 version of the manuscript.

234

# 236 **References**

240

243

247

252

255

259

262

265

- 2371.Haase AM, Gregersen T, Schlageter V, et al. Pilot study trialling a new ambulatory method238for the clinical assessment of regional gastrointestinal transit using multiple239electromagnetic capsules. Neurogastroenterol Motil. 2014;26(12):1783-1791.
- 2412.Haase AM, Gregersen T, Christensen LA, et al. Regional gastrointestinal transit times in242severe ulcerative colitis. Neurogastroenterol Motil. 2016;28(2):217-224.
- 2443.Poulsen JL, Nilsson M, Brock C, Sandberg TH, Krogh K, Drewes AM. The Impact of Opioid245Treatment on Regional Gastrointestinal Transit. Journal of neurogastroenterology and246motility. 2016;22(2):282-291.
- 4. Mark EB, Poulsen JL, Haase AM, et al. Assessment of colorectal length using the
  electromagnetic capsule tracking system: a comparative validation study in healthy
  subjects. *Colorectal disease : the official journal of the Association of Coloproctology of Great Britain and Ireland.* 2017;19(9):O350-o357.
- 2535.Gronlund D, Poulsen JL, Sandberg TH, et al. Established and emerging methods for254assessment of small and large intestinal motility. Neurogastroenterol Motil. 2017;29(7).
- Stathopoulos E, Schlageter V, Meyrat B, Ribaupierre Y, Kucera P. Magnetic pill tracking: a
   novel non-invasive tool for investigation of human digestive motility. *Neurogastroenterol Motil.* 2005;17(1):148-154.
- 260 7. Hiroz P, Schlageter V, Givel JC, Kucera P. Colonic movements in healthy subjects as
  261 monitored by a Magnet Tracking System. *Neurogastroenterol Motil.* 2009;21(8):838-e857.
- 2638.Worsoe J, Fynne L, Gregersen T, et al. Gastric transit and small intestinal transit time and264motility assessed by a magnet tracking system. BMC Gastroenterol. 2011;11:145.
- Hedsund C, Joensson IM, Gregersen T, Fynne L, Schlageter V, Krogh K. Magnet tracking
   allows assessment of regional gastrointestinal transit times in children. *Clin Exp Gastroenterol.* 2013;6:201-208.
- 27010.Koo TK, Li MY. A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for271Reliability Research. J Chiropr Med. 2016;15(2):155-163.

272

## **Figure Legends**

Figure 1: Inter-rater reliability of total and regional GI transit times compared across raters where rater 1 is most experienced, rater 2 is mid-experienced and rater 3 is least experienced. GET, gastric emptying; SITT, small intestine transit time; CTT, colonic transit time; WGTT, whole gut transit time.

277 All transit times are in hours.

278

Figure 2: Comparison of first and repeat analyses to assess intra-rater reliability of total and regional GI transit times within raters, where rater 1 is most experienced, rater 2 is mid-experienced and rater 3 is least experienced. GET, gastric emptying; SITT, small intestine transit time; CTT, colonic transit time; WGTT, whole gut transit time. All transit times are in hours.

283

# **Table Captions**

Table 1: Inter- and intra-rater reliability of duodenum and right colon timestamps between and within raters of varying levels of experience where rater 1 (R1) is most experienced, rater 2 (R2) is mid-experienced and rater 3 (R3) is least experienced.

288

289

Table 2: Inter- and intra-rater reliability of regional GI transit times between and within raters of varying levels of experience where rater 1 (R1) is most experienced, rater 2 (R2) is mid-experienced and rater 3 (R3) is least experienced. GET, gastric emptying; SITT, small intestine transit time; CTT, colonic transit time; WGTT, whole gut transit time.

TEST	ICC (95% CI)	
INTER-OBSERVER ANALYSIS		
Duodenum Timestamp		
R1-R2-R3"	0.47 (0.32 – 0.63)	
R1-R2†	0.61 (0.45 – 0.75)	
R1-R3†	0.55 (0.38 – 0.71)	
R2-R3†	0.47 (0.27 – 0.65)	
Right Colon Timestamp		
R1-R2-R3"	0.42 (0.24 – 0.60)	
R1-R2†	0.82 (0.72 – 0.89)	
R1-R3†	0.28 (0.10 – 0.48)	
R2-R3†	0.30 (0.11 – 0.50)	
INTRA-OBSERVER ANALYSIS		
Duodenum Timestamp		
R1 <sup>§</sup>	0.96 (0.92 – 0.98)	
R2 <sup>§</sup>	1.00 (0.99 – 1.00)	
R3 <sup>§</sup>	0.48 (0.16 – 0.71)	
Right Colon Timestamp		
R1 <sup>§</sup>	0.89 (0.79 – 0.94)	
R2 <sup>§</sup>	0.93 (0.87 – 0.96)	
R3 <sup>§</sup>	0.28 (-0.34 – 0.55)	
"Pooled values from 6 measurem † Pooled values from 4 measuren §Pooled values from 2 measurer	nents (first and repeat analyse	

# 311 Table 2

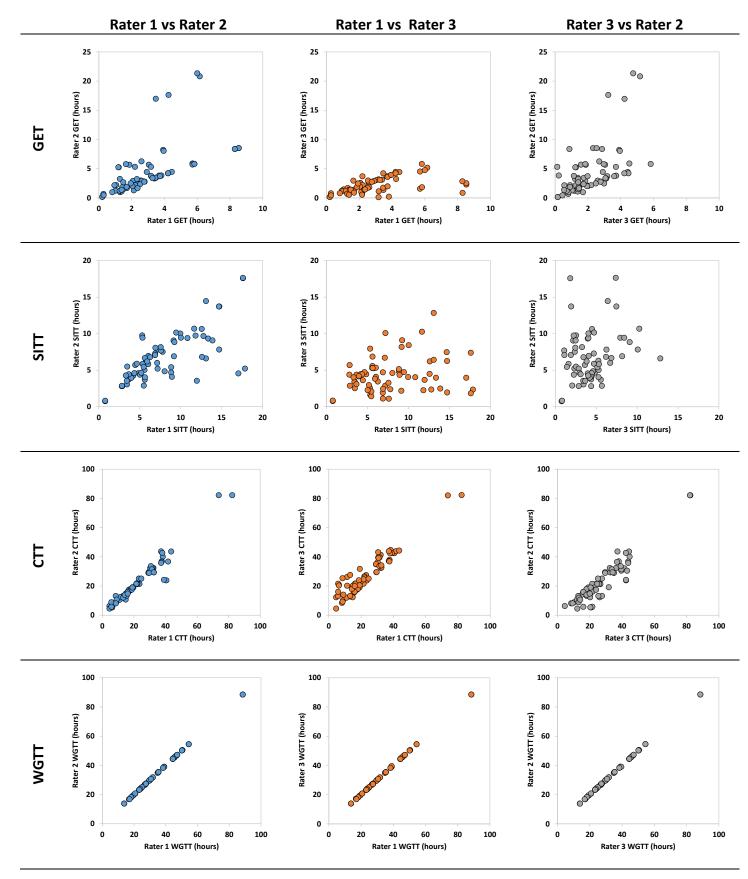
TEST	MEAN (95% CI)*	ICC (95% CI)
INTER-RATER ANALYSIS		
Gastric Emptying Time (GET)		
R1-R2-R3"	3.2 (2.8-3.6)	0.47 (0.32-0.63)
R1-R2 <sup>+</sup>	3.7 (3.1-4.2)	0.61 (0.45-0.75)
R1-R3 <sup>+</sup>	2.6 (2.3-2.9)	0.55 (0.38-0.71)
R2-R3†	3.3 (2.8-3.8)	0.47 (0.27-0.65)
Small intestine transit time (SITT)		
R1-R2-R3"	6.3 (5.8-6.8)	0.41 (0.25-0.58)
R1-R2†	7.3 (6.7-8.0)	0.73 (0.61-0.84)
R1-R3†	6.1 (5.5-6.7)	0.28 (0.11-0.48)
R2-R3†	5.6 (5.1-6.1)	0.32 (0.15-0.51)
Colonic transit time (CTT)		
R1-R2-R3"	24.3 (22.4-26.2)	0.94 (0.88-0.97)
R1-R2†	22.8 (20.5-25.2)	0.98 (0.97-0.99)
R1-R3†	25.2 (22.8-27.5)	0.93 (0.84-0.97)
R2-R3†	25.0 (22.6-27.3)	0.92 (0.82-0.96)
INTRA-RATER ANALYSIS		
Gastric Emptying Time (GET)		
R1 <sup>§</sup>	3.0 (2.5-3.4)	0.96 (0.92-0.98)
R2 <sup>§</sup>	4.4 (3.4-5.4)	1.00 (0.99-1.00)
R3 <sup>§</sup>	2.2 (1.9-2.5)	0.48 (0.16-0.71)
Small intestine transit time (SITT)		
R1 <sup>§</sup>	7.9 (6.9-8.9)	0.86 (0.75-0.93)
R2 <sup>§</sup>	6.8 (6.0-7.6)	0.84 (0.71-0.92)
R3 <sup>§</sup>	4.4 (3.8-4.9)	0.20 (-0.14-0.50
Colonic transit time (CTT)		
R1 <sup>§</sup>	23.0 (19.7-26.3)	0.99 (0.98-0.99)
R2 <sup>§</sup>	22.7 (19.3-26.0)	0.99 (0.98-1.00)
R3 <sup>§</sup>	27.3 (24.0-30.6)	0.97 (0.94-0.99)

\* Values expressed in hours

"Pooled values from 6 measurements (first and repeat analyses) + Pooled values from 4 measurements (first and repeat analyses) §Pooled values from 2 measurements (first and repeat analyses)

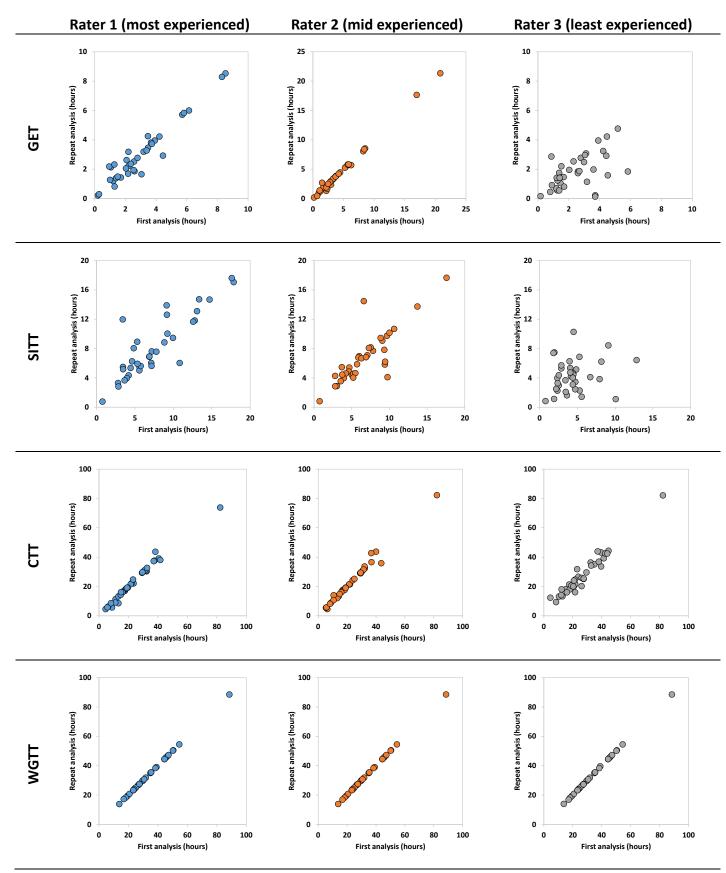
312





3D-Transit Reliability Study





3D-Transit Reliability Study