Neurogastroenterol Motil (2005) **17,** 810–820 doi: 10.1111/j.1365-2982.2005.00709.x

Barostat testing of rectal sensation and compliance in humans: comparison of results across two centres and overall reproducibility

F. CREMONINI,*[1] L. A. HOUGHTON,†[1] M. CAMILLERI,* I. FERBER,* C. FELL,† V. COX,† E. J. CASTILLO,* D. H. ALPERS,‡ O. E. DEWIT,§ E. GRAY,§ R. LEA,† A. R. ZINSMEISTER¶ & P. J. WHORWELL†

*Clinical Enteric Neuroscience Translational and Epidemiological Research (C.E.N.T.E.R.) Program, Mayo Clinic College of Medicine, Rochester, MN, USA

†Neurogastroenterology Unit, Academic Division of Medicine and Surgery, University Hospital of South Manchester, Manchester, UK

‡Washington University School of Medicine, St Louis, MO, USA

§Neurology and GI Centre of Excellence of Drug Discovery, GlaxoSmithKline, Harlow, UK

¶Department of Health Sciences Research, Division of Biostatistics, Mayo Clinic College of Medicine, Rochester, MN, USA

Abstract *We assessed reproducibility of measurements of rectal compliance and sensation in health in studies conducted at two centres. We estimated samples size necessary to show clinically meaningful changes in future studies. We performed rectal barostat tests three times (day 1, day 1 after 4 h and 14–17 days later) in 34 healthy participants. We measured compliance and pressure thresholds for first sensation, urgency, discomfort and pain using ascending method of limits and symptom ratings for gas, urgency, discomfort and pain during four phasic distensions (12, 24, 36 and 48 mmHg) in random order. Results obtained at the two centres differed minimally. Reproducibility of sensory end points varies with type of sensation, pressure level and method of distension. Pressure threshold for pain and sensory ratings for non-painful sensations at 36 and 48 mmHg distension were most reproducible in the two centres. Sample size calculations suggested that crossover design is preferable in therapeutic trials: for each dose of medication tested, a sample of 21 should be sufficient to demonstrate 30% changes in all sensory thresholds and almost all sensory ratings. We conclude that reproducibility varies with sensation type, pressure level and distension method, but in a two-centre study, differences in observed results of sensation are minimal and pressure threshold for pain and sensory ratings at 36–48 mmHg of distension are reproducible.*

Keywords *ascending method of limits, perception, sensory ratings, visceral.*

Address for correspondence
Dr Michael Camilleri, Mayo Clinic, Charlton 8-110, 200 First St S.W., Rochester, MN 55905, USA.
Tel: 507-266-2305; fax: 507-255-5720;
e-mail: camilleri.michael@mayo.edu
Received: 6 April 2005
Accepted for publication: 20 June 2005
[1]FC and LAH equally contributed to this study.

INTRODUCTION

Visceral hypersensitivity is defined as an excessive perception of visceral stimuli that are not usually painful. Visceral hypersensitivity is currently considered a determinant, and possibly a biological measure, of functional bowel disorders such as the irritable bowel syndrome (IBS)[1,2] and functional dyspepsia.[3] The presence of visceral hypersensitivity in IBS provides the rationale for the development of novel therapies modulating afferent neurotransmission.[4] Predefined and standardized distensions of the bowel wall using a barostat device are the current standard for the assessment of sensorimotor function in experimental trials in health and disease.[5]

Visceral sensitivity can be quantified operationally by measuring the threshold (expressed as the volume or pressure at which a sensation occurs), or the sensory rating or intensity in response to a stimulus, such as distension or electrical stimulation.[5] Intensity is rated while applying a known stimulus, and measuring the perception, typically using visual analogue scales (VAS).[5] However, there is also a fundamental difference in the significance of these sensory end points. Threshold implies a single end point (a point separating

conditions that will produce a given effect from conditions of a higher or lower degree that will not produce the effect[6]). In contrast, sensory ratings allow the definition of pain or other sensation across a broad range from subnoxious to noxious levels.

The methods adopted for these measurements, and the sensory events used for declaring thresholds are not standard across different centres and studies. For example, some groups declare separate thresholds for evoking a first sensation, urgency and pain while distending the balloon in the rectum.[5] Other authors use a 0–6 sensation rating scale, with 0 signifying no sensation and 6 pain; in these studies, the term 'threshold' is used to summarize the pressure or volume needed to evoke a sensation score higher than 3 of a (maximum) score of 6.[7] Still other authors define a 'mean threshold' as the average pressure at which the participant reports a particular symptom during a standard sequence of distensions.[8] With the latter approach, differences in the number of distensions or in the scaling of pressures in the sequence of distensions will influence the mean value.

There are incomplete validation data available for any of these end points or tests for rectal sensation to plan experimental therapeutic studies. The absence of such validation with rectal sensation studies contrasts with the recent report that tested the reproducibility of the barostat balloon in tests of gastric sensitivity.[9]

Thus, there is no agreement on the most reliable protocol to measure visceral hypersensitivity. For example, it is unclear whether parallel group or crossover design, and thresholds or sensory ratings provide the optimal experimental designs and end points. Our current lack of understanding of the intraindividual and interindividual variability of these measurements may limit their suitability and selection of sample sizes for experimental studies, leading to possible type II statistical errors.

The aim of the present study was to assess the performance characteristics of rectal sensory testing in healthy humans, specifically the reproducibility of measurements of rectal sensation and compliance during standardized distensions with the barostat performed in two separate centres.

METHODS

Participants, study design and questionnaires

Thirty-four healthy volunteers [10 males, mean age 27.8 ± 1.4 years, mean body mass index (BMI) 25 ± 1] were enrolled in this study with 17 enrolled in each centre (Fig. 1). All participants signed informed consent, and this study was approved by the Mayo Clinic Institutional Review Board and South Manchester Research Ethics Committee. Current comorbidities, abdominal symptoms, previous abdominal surgery and concomitant somatization or psychological disorders were excluded by means of a validated bowel symptom questionnaire, a somatic checklist and a physical examination.[10] In order to test for a potential association between anxiety or fear of pain and the sensitivity thresholds or sensory ratings, subjects also completed the 30-item Fear of Pain Questionnaire (FPQ-III)[11] and a revised version of the 36-item Anxiety Sensitivity Index (ASI-R)[12] before each procedure. The rectal barostat procedures were performed using an identical sequence on three separate occasions at each centre (Fig. 1): twice on the first study day (4 h apart) and once 14–17 days later. The subjects were given a 150 mL glucose drink [Lemon Lucozade Sport, 118 kJ

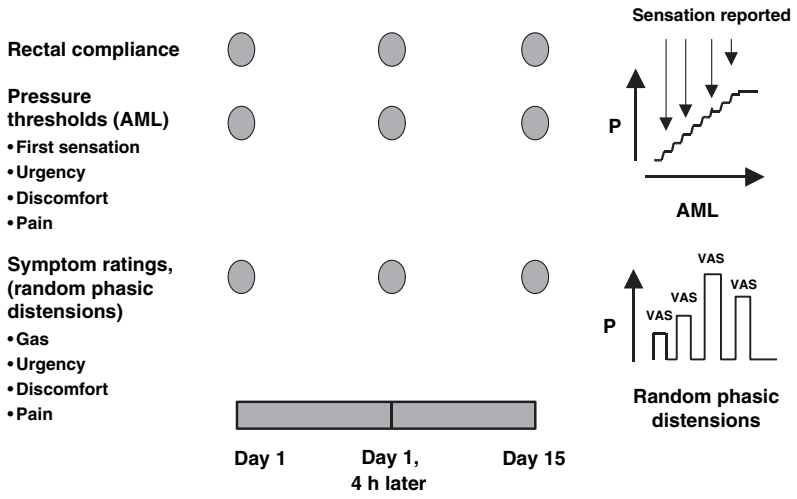

Figure 1 Experimental design.

(28 kcal) per 100 mL; SmithKline Beecham, Brentford, UK] 1 h prior to setting baseline operating pressure (BOP) on each of the three test occasions. This was performed to provide standard hydration and minimal calories in participants who would have fasted from midnight and were embarking on 9 h of studies on the first study day.

Rectal barostat

All subjects presented to the research centres of the Mayo Clinic and the University Hospital of South Manchester after bowel preparation (Fleet® phosphate enema, self-administered at least 1 h before reporting to the centres) and an overnight fast. A catheter (customized rectal barostat catheter, part no. C7-2CB-R-22F, MUI Scientific, Mississauga, ON, Canada), to which was attached a polyethylene bag (Pillow Type Rectal Barostat Balloon, part no. CT-BP600R; length, 22 cm; diameter, 15 cm; capacity 600 mL; MUI Scientific), was inserted into the rectum so that the middle of the balloon was located approximately 10 cm from the anal verge. To decrease the effects of abdominal viscera on the balloon volume, the subjects were placed in a semiprone position and the foot end of the bed elevated 15°. The bag was then unfolded by transiently inflating it with 75 mL of air and then deflating it completely. After a 20–30-min recovery period, the catheter was connected to a barostat (G&J Electronics Inc., Toronto, ON, Canada) and the pressure in the bag increased from 4 mmHg in steps of 1 mmHg for 1 min per step until respiratory excursions were observed. The BOP was defined as 2 mmHg above the minimal distension pressure at which respiratory excursions were clearly recorded from the barostat tracing. If respiratory variations were not seen by 18 mmHg, BOP was set at 12 mmHg. An initial 'conditioning' distension of the rectum was then performed in which the pressure was increased from 0 mmHg in steps of 4 mmHg for 15 s per step until 20 mmHg was reached. Previous studies have shown that an initial 'conditioning' distension to 20 mmHg renders subsequent assessments of compliance and perception more reproducible.[13] The bag was then deflated to 0 mmHg and the subjects were allowed to rest for 10 min before proceeding to the ascending method of limits.

Ascending method of limits: compliance and sensory thresholds

Rectal compliance and sensory thresholds were measured by ramp inflation, starting at 0 mmHg and increasing in steps of 4 mmHg for 1 min per step to a maximum of 60 mmHg. Thresholds for first sensation, urgency, discomfort and pain (Fig. 1) were indicated by the subjects by pressing a button at the distension pressure at which sensations were perceived. Ramp inflation was terminated as soon as the subjects reported the first sensation of pain. Following this procedure, the bag was deflated to BOP and the subjects allowed to rest for 10 min.

Random order phasic distensions: sensory ratings

After the ascending method of limits' protocol, phasic distensions of 12, 24, 36 and 48 mmHg above BOP were each applied once in random order. Each distension was maintained for 60 s with an interstimulus interval of 2 min during which the balloon was deflated to BOP. This approach has been shown to be reliable in multiple previous studies, as the distension pressure is associated with intensity ratings that are generally proportional to the magnitude of the distension pressures.[14–16] Study participants were blinded to the distension order, which was provided by the study statistician (A.R.Z.) and was also randomized between study days. Subjects were asked to mark four separate 100 mm VAS 30 s after the onset of the distension for the sensations of gas, urgency, discomfort and pain. These scales were anchored at each end by the descriptions 'unnoticeable' and 'unbearable'. Pressure was immediately released if the subject reported >80 mm of discomfort or pain on the VAS and higher distensions were not subsequently administered. During the assessment of sensation, the interaction between the subject and the study investigator was kept to a minimum.

Data analysis

The following measurements were derived: (i) the sensory thresholds for gas, urgency, discomfort and pain during ascending method of limits, (ii) the aggregate sensation score (gas, urgency, discomfort and pain) in response to the four random phasic distensions (12, 24, 36 and 48 mmHg above BOP) and (iii) rectal compliance.

Because rectal pressure–volume relationships are sigmoidal, compliance was summarized, as in previous studies,[13,15] by a power exponential model by plotting observed volume at each pressure divided by the maximum observed volume as a function of 1/pressure as follows: $\mathrm{Vol}/\mathrm{Vol}_{\max} = r + \exp[-(\kappa \times 1/\mathrm{Pr})^{\beta}]$, where r represents the minimum observed volume divided by the maximum observed volume ($\mathrm{Vol}_{\max}$), Pr is pres-

sure, and κ and β are constants that describe the compliance curve. A summary value for compliance was also computed from the estimated constants for each subject, specifically the pressure observed at one half of the maximum observed volume ($Pr_{1/2}$), where a smaller $Pr_{1/2}$ corresponds to higher compliance.

Statistical analysis

Sensory end points The assessment of differences in observed sensory ratings (rectal sensation VAS scores) at the two testing sites was based on ANCOVA for repeated measures (the four phasic distension levels by three distinct study sessions). Each type of VAS score was analysed individually (gas, urgency, discomfort and pain). The covariates in this analysis were gender and the corresponding pressure at each phasic distension level.

The differences in each of the rectal sensation thresholds observed at the two testing sites were assessed using a proportional hazards regression analysis separately for each type of sensation threshold (first sensation, urgency, discomfort and pain – using the PHREG procedure in SAS version 8). This analysis was required to account for the modest number of 'censored' values included in the analysis. This occurred when subjects failed to report a specific sensation threshold at the highest distension pressure used. In such a situation, we included a value in the analysis as a 'censored' observation, by the last observation carried forward. Covariates included in these analyses included gender and the maximum volume attained at the initial study session (day 1). To account for the multiple observations per subject, a multiple events method for computing the standard errors of the regression coefficients was used (the so-called 'sandwich' variance–covariance estimator) and a simultaneous test for differences in results between centres (the Lin-Wei-Weissfield method).

Rectal compliance A two-sample *t*-test was used to assess at the two study sites the differences in the computed compliance ($Pr_{1/2}$) values for each separate estimate in the three study sessions. The κ- and β-constants in the compliance curves were estimated using the NLIN procedure in the SAS software package.[17]

Coefficients of variation The intraindividual coefficients of variation (CV) were calculated as percentage of the ratio between the SD of the individual differences (day 15–day 1) to the individual mean measurements (day 1 + day 15/2). The interindividual CV were calculated as the percentage of the ratio of the (between subjects) SD and the mean of the measurements obtained at each testing session (day 1, day 1 plus 4 h, and day 15). These CV (%) were computed ignoring the censoring status for the rectal sensation threshold pressures.

Sample calculations for future studies Using the mean values and variances observed in this study, we calculated the sample size for parallel group and crossover design studies required to demonstrate differences in the pressure thresholds (ignoring the censoring status and pooling the data across centres) for first sensation, urgency, discomfort and pain during ascending method of limits and for the sensory ratings at 36 and 48 mmHg distensions.

RESULTS

All subjects completed the study. Participants' median BMI was 25.0 kg m^{-2} (range: 19.3–41.4). There were no differences in BMI across centres [median BMI 23.1 (IQR: 21–25.4)] for participants in Manchester, 24.4 (IQR: 21.6–29.2) for participants in Rochester ($P = 0.11$ by rank sum test). As the protocol required that distension sequences be interrupted if the subjects reported pain, data for sensation ratings at the highest pressures (during phasic distensions) were available for at least 27 of 34 participants at 48 mmHg and for at least 28 of 34 participants at 36 mmHg pressures. All participants completed the distensions during ascending method of limits.

Sensory thresholds using ascending method of limits

Comparison of results at two testing sites The mean (and 95% CI) overall pressure threshold for each of the four sensations: gas, urgency, discomfort and pain obtained during the ascending method of limits are reported in Fig. 2. For eight of a possible 12 comparisons between thresholds at the three testing dates, there was no difference in results obtained at the two testing centres. Overall, the average variances in discomfort ($P = 0.0004$) and pain ($P = 0.007$) were different between the two testing sites. Note in Fig. 2, the larger variances for discomfort and pain at the Rochester compared with the Manchester test site. These effects were most noticeable on the first study session, but were also evident on the second session on day 1 for the discomfort threshold level.

Significantly larger variances in the thresholds for discomfort in the first day testing sessions ($P = 0.0004$ and $P = 0.0043$ for the initial testing session and for the post-4 h session, respectively) and for pain

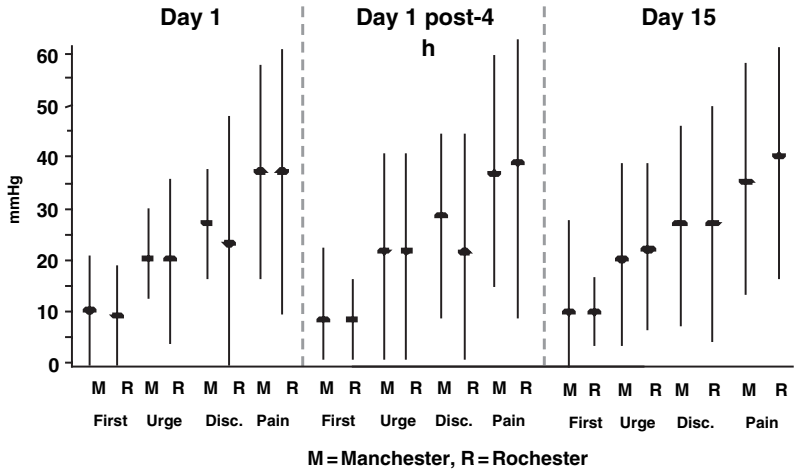

Figure 2 Mean sensory thresholds [95% confidence interval (CI)] in the two centres (M = Manchester, R = Rochester) on day 1, day 1 post-4 h and day 15. Despite the wide interindividual variability, there is good reproducibility of pressure thresholds between centres and between sessions. Disc., discomfort threshold.

($P = 0.012$ and $P = 0.055$ for the initial and post-4 h sessions, respectively) were observed in Rochester, compared with Manchester. However, note also that the mean sensory thresholds at these sites are not significantly different.

Reproducibility The overall analysis showed good reproducibility of the mean pressure thresholds for all the sensations across the three study sessions in the two centres. Although there are lower variances seen in Fig. 2 for first sensation and urgency, the intraindividual and interindividual variability in thresholds for first sensation, urgency and discomfort were numerically greater than the variability in pain sensation thresholds (Table 1), because the mean values of these thresholds are lower than the sensation threshold for pain (CV = SD/mean, in general).

There was no effect of age, gender, BMI, anxiety and fear of pain scales on the reproducibility of sensation thresholds.

Sensory ratings using random order phasic distensions

Comparison of results at two testing sites There was no overall difference in any of the four VAS sensation rating scores between the two testing sites. The only difference observed between study sessions is noted in Fig. 3, where in one centre (Rochester), the ratings of pain and urge at 36 mmHg distensions obtained on day 1, second test were somewhat (but not significantly) lower when compared with the first measurements on day 1. Ratings of pain on day 1, second test are also lower compared with the measurements on day 15. However, the mean ratings of pain at 36 mmHg distensions on day 15 were similar to those obtained during the first test on day 1 (Fig. 3). There were only minimal differences in urgency ratings between the three testing sessions ($P = 0.026$) and these appear to be driven by the higher ratings in day 1 session 1 in Rochester (Fig. 3A).

Reproducibility Each individual centre's data analysis showed reproducibility of the data. The sensation ratings using VAS for gas, urgency, discomfort and pain (Fig. 3) were highly consistent within individuals during the highest pressure distensions (48 mmHg). In the overall sample, the interindividual variability of VAS was lower at higher distending pressures. This is shown by the higher CV for 36 mmHg distensions compared with 48 mmHg distensions in Table 2.

Table 1 Intraindividual and interindividual coefficients of variation (CV) of pressure thresholds for first sensation, discomfort, urgency and pain

Sensation	CV interindividual (%, day 1)	CV interindividual (%, day 1 + 4 h)	CV interindividual (%, day 15)	CV intraindividual (%, day 1–day 15)
First sensation	56	77	64	43
Urgency	34	49	41	41
Discomfort	40	47	38	47
Pain	34	37	31	23

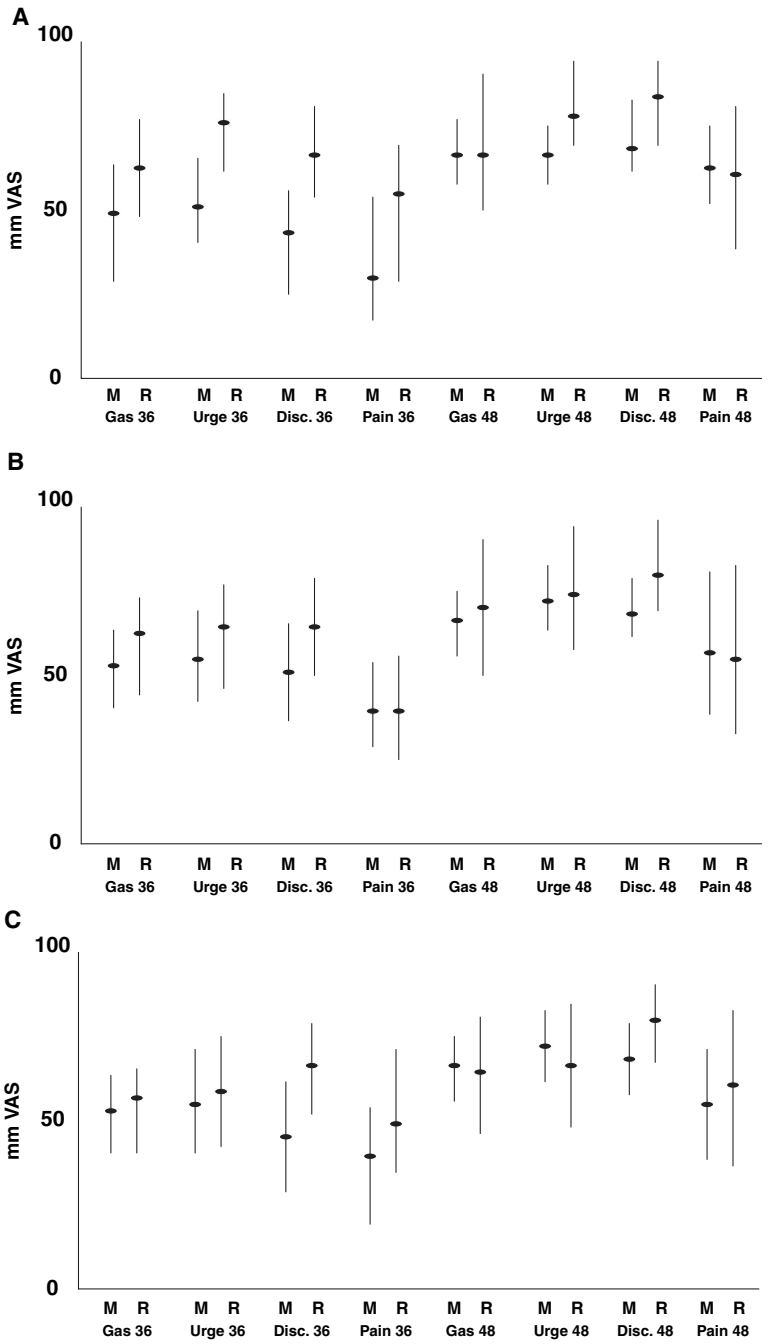

Figure 3 Mean sensory ratings [with 95% confidence interval (CI)] for the two centres (M = Manchester, R = Rochester) on day 1 (A), day 1 post-4 h (B) and day 15 (C).

Good intraindividual reproducibility was observed at 48 mmHg for the ratings of gas, urgency and discomfort (Table 2), while the ratings of pain had the highest interindividual and intraindividual variability of all sensory ratings at a specific distension level tested (Table 2).

The sensation ratings during 12 and 24 mmHg random order distensions showed limited reproducibility and are not presented in this report.

There was no effect of age, gender, BMI, anxiety and fear of pain scales on the reproducibility of sensory ratings.

Rectal compliance using ascending methods of limits

Fig. 4 shows the mean (95% CI) rectal compliance ($Pr_{1/2}$) on day 1, day 1 after 4 h, and on day 15 in the

Table 2 Intraindividual and interindividual coefficients of variation (CV) of sensory ratings during random order distensions for all participants across the two centres

Sensation and pressure applied (mmHg)	CV interindividual (%)			
	Day 1	Day 1 + 4 h	Day 15	Day 1–day 15
Gas 36	41	38	45	42
Gas 48	31	33	35	23
Urge 36	33	37	46	42
Urge 48	22	28	34	21
Discomfort 36	44	39	45	44
Discomfort 48	21	27	29	21
Pain 36	65	72	73	70
Pain 48	40	56	53	52

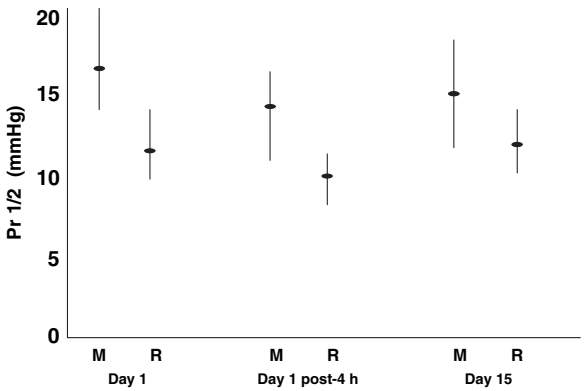

Figure 4 Mean [95% confidence interval (CI)] pressures observed at one half of the maximum observed volume ($Pr_{1/2}$) as summary values for the compliance curves in the different testing sessions across centres (M = Manchester, R = Rochester).

two testing sites. These measurements were reproducible in each centre, as shown by the minimal variation in results between the study sessions. However, a significant centre effect was observed consistently. The results of compliance were higher in Rochester, shown by the consistently lower $Pr_{1/2}$ ($P = 0.0004$) over the three study sessions. However, these mean intercentre differences in $Pr_{1/2}$ were within 5 mmHg and, therefore, of doubtful clinical relevance. Study session differences were also detected ($P = 0.033$), although the overall session differences in mean values were only 1.5–2.5 mmHg.

Sample sizes for demonstrating differences between two groups in future parallel and crossover designed studies

Table 3 shows the sample sizes (per treatment arm) required to demonstrate 25, 30 and 35% differences in these sensory end points between two hypothetical groups (parallel-group study) and in a crossover design study. The numbers calculated refer to 80% power and a two-sided $\alpha = 0.05$.

Table 3 Sample sizes for demonstrating 25, 30 and 35% difference between two hypothetical study groups in the pressure thresholds for sensation during ascending method of limits and in the sensory ratings

	Parallel-group design (%)			Crossover design (%)		
End points	35	30	25	35	30	25
Pressure thresholds						
First sensation	53	76	116	14	21	31
Urgency	20	32	48	11	15	23
Discomfort	22	29	44	9	13	19
Pain	12	18	27	3	4	6
Sensory ratings, random distensions (mmHg)						
Gas 36	14	21	31	8	11	17
Gas 48	12	18	27	3	4	6
Urgency 36	13	18	27	8	11	17
Urgency 48	9	13	20	2	3	4
Discomfort 36	14	19	29	9	13	19
Discomfort 48	6	9	13	2	3	4
Pain 36	49	71	108	24	34	52
Pain 48	32	46	70	15	21	32

Sample sizes are per study arm and are calculated using intraindividual and interindividual coefficients of variation from all participants and are presented for a hypothetical parallel-group and crossover study design, with 80% power and a two-sided α of 0.05.

The sample sizes needed for planning a crossover study were smaller than those required for a parallel-group study. Using a distension protocol with ascending method of limits, smaller sample sizes would be required for demonstrating differences in pressure thresholds for pain whether the design was for parallel groups or crossover. For random order distensions of 36 and 48 mmHg, smaller sample sizes would be required to show a difference in the ratings for non-painful sensations in both parallel group or crossover designs. Using a crossover design, the data show that for each dose tested in an experimental therapeutic study, a sample size of 21 would be sufficient to demonstrate a 30% change in all sensation thresholds, and in all sensory ratings except the rating for pain at 36 mmHg distension.

DISCUSSION

In the present study, we report substantial day-to-day and intercentre reproducibility for sensory end points using the rectal barostat in healthy subjects in two

testing centres, one in Europe and the other in North America. However, there were some centre effects in sensory thresholds. Reproducibility depends on the type of sensation, pressure level and method of distension. Thus, we observed high reproducibility (i.e. low CV) for the pressure sensitivity threshold of pain, but less reproducibility for non-painful sensations (Table 1, Fig. 2). For VAS ratings, reproducibility was better for the non-painful sensory ratings, particularly at higher distension pressures (Table 2, Fig. 3). The CV for all sensory ratings was lower at 48 mmHg compared with 36 mmHg.

Rectal sensation has been used as a biomarker in comparisons between health and IBS and in studies of experimental therapies in healthy volunteers and in patients with functional bowel disorders. Bouin *et al*. suggested good sensitivity and specificity and predictive values of pain threshold measurements to differentiate patients with functional dyspepsia or IBS compared with healthy controls.[18] The pain thresholds observed in the healthy control group of Bouin *et al*. and those in the healthy subjects reported in other studies[1,2,13,18,19] using ramp-like distension protocols or ascending method of limits are similar to the thresholds observed in the present study (Table 4). These data suggest that our findings are generalizable to other healthy populations. While the study did not specifically address the issue of sex-related differences, we have evaluated 24 females and 10 males, and this ratio reflects the gender distribution of IBS in epidemiological studies.

In the present study on rectal sensation, we found little intraindividual day-to-day variability in the measurement of the thresholds for the sensation of pain, but other end points such as the sensory ratings for urgency during random order distensions were more variable.

These observations apply to healthy subjects and may not be generalizable to IBS patients, in whom anxiety or fear of pain levels may be higher and in whom increased perception of non-noxious sensations has been reported.[20] Furthermore, there are data which suggest that a smaller percentage of IBS patients are insensitive to rectal distension, especially those with non-urge constipation-predominant IBS.[21,22] A study characterizing differences in visceral sensitivity between IBS subtypes using volume-based distensions of a latex balloon during anorectal manometry showed reproducible sensory thresholds. The level of reproducibility was not characterized in detail.[23]

It is also conceivable that variability may be significantly greater in IBS patients because of a variety of factors unrelated to the stimulus, such as anxiety, belief systems, response to verbal or non-verbal cues and instructions.

The rectal barostat study was recommended by multinational consensus for evaluating sensation;[5] however, reproducibility data are generally lacking. A study in three IBS patients has shown that measurements of rectal sensory end points is more reproducible during phasic than during ramp distensions.[24] In the present study with two widely used distension protocols, we found that both the ascending method of limits and random order distensions have good reproducibility although there are differences in the maximal reproducibility among sensations and methods of distensions. Previous data in health also suggest older subjects have increased rectal sensitivity thresholds for first sensation, urgency and pain.[25] We observed low reproducibility for sensory ratings at 12 and 24 mmHg phasic distensions, and this may be attributable either to the lower volumes of distension, or to the intrinsic difficulty to differentiate first sensation and urgency in these sensation studies. This should not be confused with the reproducibility shown with the stepwise distensions to identify thresholds (Table 1). The latter is commonly used to identify hypersensitivity (e.g. 28 mmHg in the Bouin *et al*. study[18]).

Table 4 Comparison of pressure thresholds for sensation in health across published studies using barostat-delivered ramp-like distensions

	End point					
Author	First sensation	Desire to defecate/stool	Urge	Pain	Discomfort	Compliance ($Pr_{1/2}$)
Bharucha *et al*.[19]	15 (2)	20 (2)	27 (2)	n.r.	n.r.	14 (1)
Bouin *et al*.[18]	n.r.	n.r.	n.r.	44 (5)	n.r.	n.r.
Hammer *et al*.[13]	6 (1)	14 (2)	21 (4)	n.r.	n.r.	n.r.
Lembo *et al*.[1]	n.r.	16 (1)	n.r.	n.r.	31 (3)	20
Mertz *et al*.[2]	n.r.	n.r.	n.r.	n.r.	30 (3)	n.r.
Naliboff *et al*.[28]	n.r.	34 (3)	37 (4)	39 (5)	38 (2)	n.r.
Present study	10 (1)	n.r.	20 (1)	37 (2)	25 (2)	12 (1)

Values are expressed as mean pressures (SEM), mmHg.
n.r., not recorded.

Although differences in rectal perception according to gender or BMI influence visceral sensation end points,[25,26] our goal in the present study was not to evaluate influence of gender or age, and further studies are needed to address these issues. While BMI values did not differ between centres, the study analysis included age and gender as covariates and, hence, any observed variability in the measurements is unlikely to be accounted for by such differences across centres.

We found high reproducibility of rectal compliance measurements within centres, consistent with previous single-centre smaller studies. Hammer *et al.* reported limited variability of rectal compliance and rectal sensitivity during two repeated measurements in five healthy volunteers.[13] A larger study of 19 healthy women reported low intraindividual variation of rectal compliance assessed on two separate occasions.[19] In the present study, the derived values of compliance ($Pr_{1/2}$) were significantly different between centres. The calculated $Pr_{1/2}$ is based on several previous reports,[15,16,19] and a lower $Pr_{1/2}$ implies higher compliance because it is the pressure at half of the maximal volume achieved at the highest pressure distension. However, the statistically significant differences across centres observed in $Pr_{1/2}$ are likely to have little relevance to rectal sensorimotor function, given the small absolute differences observed within- and between-centres.

A second major observation in our study refers to the calculated sample sizes to demonstrate differences between treatments in sensory end points, as shown in Table 3. As to be expected from the reproducibility data, smaller sample sizes are required to compare pain thresholds and non-painful sensory ratings, whether using a parallel or crossover design. The minimal sample sizes to detect a 30% difference between two treatment groups using a crossover study design for pain thresholds and sensory ratings for non-painful sensations (gas, urgency and discomfort) during phasic distensions at 48 mmHg is three to four subjects; however, with the same distension pressure, the sample size to show a change in pain is around 20. A sample size of 21 would be able to detect a 30% change in virtually all sensation thresholds and ratings.

For studies with parallel-group design, sample size estimates are considerably higher than for crossover studies. To place our findings in further perspective, we summarized, from the published literature, the percentage differences in the urgency and pain thresholds in several drug trials that used ramp-like distension (or ascending method of limits) delivered with the barostat in health and IBS (Table 5). This overview of previous data showed that parallel-group design studies needed much larger sample sizes than crossover trials to show similar magnitude of differences in thresholds. However, we noted that the drug effect size detected in pain thresholds is rarely ≥20% except with fentanyl[8] or octreotide.[27] The clinical significance of such a demonstrable effect is uncertain.

Responsiveness to medication of sensation ratings during phasic distensions has been evaluated in fewer studies and most utilized crossover design (Table 6). Larger percentage differences are observed in the average ratings in response to treatment (Table 6). This may suggest that the observable differences in sensation with phasic distensions are clinically signi-

Table 5 Differences in pressure thresholds for noxious sensations, expressed as percentage change in response to several therapeutic agents, in studies using pressure-based ramp-like distensions with rectal barostat

Author	Participants	*N* (per group)	Drug, route of administration	Difference in threshold(s) (%)	Study design
Hammer *et al.*[29]	Healthy	5	Ondansetron 0.15 mg kg^{-1} i.v.	Urgency 14	Postdrug *vs* predrug
Czimmer *et al.*[30]	IBS	15	Otilonium bromide 120 mg p.o.	Pain 10	Postdrug *vs* predrug
Sabate *et al.*[31]	Healthy	8	Cholecystokinin OP 40 ng kg^{-1} h^{-1} i.v.	Pain 20 Urgency 8	Crossover
Lembo *et al.*[8]	IBS	10	Fentanyl 56 μg i.v.	Pain 20	Crossover
Bradette *et al.*[27]	IBS	10	Octreotide 1.25 μg kg^{-1} s.c.	Discomfort 39 Pain 33	Crossover
Delvaux *et al.*[32]	IBS	20	Asimadoline 0.5 mg p.o.	Pain 13	Crossover
Kuiken *et al.*[33]	IBS	40	Fluoxetine 20 mg p.o.	Discomfort/pain 13	Parallel group
Delgado-Aros *et al.*[16]	Healthy	30	Asimadoline 1.5 mg p.o.	Pain 10	Parallel group
Delvaux *et al.*[34]	IBS	8	Alosetron 0.5–8 mg p.o.	Pain 5	Parallel group

The differences shown are between the higher drug dose used in each study and the placebo group or the predrug measurement.
IBS, irritable bowel syndrome; OP, operating pressure.

Table 6 Differences in sensory ratings (mm visual analogue scales) during phasic distensions, expressed as percentage change in response to several treatments, in studies using rectal barostat

Author, condition	Participants	*N* (per group)	Drug and route of administration	Differences in ratings	Study design
Van der Veek *et al.*[35]	Healthy	8	Neurotensin 5 pmol kg^{-1} min^{-1} i.v.	Urge at 30 mmHg – 17% Pain at 30 mmHg – 15%	Crossover
Law *et al.*[36]	Healthy	4	Neostigmine 1.5 mg i.v. Bethanechol 10 mg s.c.	Urgency at 32 mmHg – 27% Pain at 32 mmHg – 54% Urgency at 32 mmHg – 2% Pain at 32 mmHg – 37%	Crossover
Thumshirn *et al.*[37]	IBS	10	Alosetron 8 mg p.o. Placebo p.o.	Gas at 36 mmHg – 13% Urgency at 36 mmHg – 26% Pain at 36 mmHg – 57% Gas at 36 mmHg – 37% Urgency at 36 mmHg – 53% Pain at 36 mmHg – 74%	Crossover
Bharucha *et al.**[15]	Healthy	6–7 (10 in saline group)	Yohimbine 0.125 mg kg^{-1} i.v. Phenylephrine 0.4 μg kg^{-1} min^{-1} i.v. Ritodrine 100–300 μg min^{-1} i.v. Clonidine 0.3 mg p.o.	Gas at 32 mmHg – 10% Pain at 32 mmHg – 66% Gas at 32 mmHg – 33% Pain at 32 mmHg – 31% Gas at 32 mmHg – 39% Pain at 32 mmHg – 7% Gas at 32 mmHg – 4% Pain – 100%	Parallel group *vs* saline

*Colonic barostat.
IBS, irritable bowel syndrome.

ficant. However, to date, most of the differences have only seldom been statistically significant because of the high interindividual (in parallel-group studies) or intraindividual variability (in crossover studies). It is therefore conceivable that some of the previously published negative reports might have had insufficient statistical power, based on the CV observed in the present study. Taken together, the literature overview and our present data suggest that a parallel-group design may not be the most desirable option in the planning of future experimental therapeutic trials in healthy subjects. The crossover design appears preferable; however, the drawbacks of this design include the sequence and carryover effects and the potential for dropouts in the study, especially one that involves an invasive procedure.

In conclusion, rectal pain thresholds and sensory ratings with higher pressure distensions obtained with the barostat and polyethylene balloons are reproducible across two centres in healthy participants. These reproducibility data are, to our knowledge, the largest set and the only to have documented the two-centre comparison of test performance. Such assessment of variation/reproducibility is essential for the appropriate use of these end points in adequately designed and sized studies of experimental therapies performed either in single or multiple centres.

ACKNOWLEDGMENTS

The studies performed in the Mayo General Clinical Research Center were supported by grant no. RROO585 from National Institutes of Health and by a grant from GlaxoSmithKline. The studies performed in the Neurogastroenterology Unit, University Hospital of South Manchester were supported by a grant from GlaxoSmithKline. Dr Alpers is a consultant for GlaxoSmithKline. Dr Camilleri is supported by grants R01-DK54681, R01-DK67071 and K24-DK02638 from National Institutes of Health.

REFERENCES

1 Lembo T, Munakata J, Mertz H *et al.* Evidence for the hypersensitivity of lumbar splanchnic afferents in irritable bowel syndrome. *Gastroenterology* 1994; **107**: 1686–96.
2 Mertz H, Naliboff B, Munakata J, Niazi N, Mayer EA. Altered rectal perception is a biological marker of patients with irritable bowel syndrome. *Gastroenterology* 1995; **109**: 40–52.
3 Mearin F, Cucala M, Azpiroz F, Malagelada J-R. The origin of symptoms on the brain-gut axis in functional dyspepsia. *Gastroenterology* 1991; **101**: 999–1006.
4 Camilleri M. Management of the irritable bowel syndrome. *Gastroenterology* 2001; **120**: 652–68.
5 Whitehead WE, Delvaux M, The Working Team of Glaxo-Wellcome Research, UK. Standardization of barostat

procedures for testing smooth muscle tone and sensory thresholds in the gastrointestinal tract. *Dig Dis Sci* 1997; **42**: 223–41.
6 Microsoft, Inc. *Microsoft Bookshelf Dictionary*. Redmond, WA, USA: Microsoft, 1994.
7 Distrutti E, Azpiroz F, Soldevilla A, Malagelada J-R. Gastric wall tension determines perception of gastric distention. *Gastroenterology* 1999; **116**: 1035–42.
8 Lembo T, Naliboff BD, Matin K *et al.* Irritable bowel syndrome patients show altered sensitivity to exogenous opioids. *Pain* 2000; **87**: 137–47.
9 Sarnelli G, Vos R, Cuomo R, Janssens J, Tack J. Reproducibility of gastric barostat studies in healthy controls and in dyspeptic patients. *Am J Gastroenterol* 2001; **96**: 1047–53.
10 Talley NJ, Phillips SF, Melton J III, Wiltgen C, Zinsmeister AR. A patient questionnaire to identify bowel disease. *Ann Intern Med* 1989; **111**: 671–4.
11 Osman A, Breitenstein JL, Barrios FX, Gutierrez PM, Kopper BA. The fear of pain questionnaire III: further reliability and validity with nonclinical samples. *J Behav Med* 2002; **25**: 155–73.
12 Taylor S, Cox BJ. An expanded Anxiety Sensitivity Index: evidence for a hierarchic structure in a clinical sample. *J Anxiety Disord* 1998; **12**: 463–83.
13 Hammer HF, Phillips SF, Camilleri M, Hanson RB. Rectal tone, distensibility, and perception: reproducibility and response to different distensions. *Am J Physiol* 1998; **274**: G584–90.
14 Ford MJ, Camilleri M, Zinsmeister AR, Hanson RB. Psychosensory modulation of colonic sensation in the human transverse and sigmoid colon. *Gastroenterology* 1995; **109**: 1772–80.
15 Bharucha AE, Camilleri M, Zinsmeister AR, Hanson RB. Adrenergic modulation of human colonic motor and sensory function. *Am J Physiol* 1997; **273**: G997–G1006.
16 Delgado-Aros S, Chial HJ, Camilleri M *et al.* Effects of a kappa-opioid agonist, asimadoline, on satiation and GI motor and sensory functions in humans. *Am J Physiol* 2003; **284**: G558–66.
17 Anonymous. The NLIN procedure. In: *SAS/STAT User's Guide* (version 6). Cary, NC, USA: SAS Inst, Inc., 1989: 1135–94.
18 Bouin M, Plourde V, Boivin M *et al.* Rectal distention testing in patients with irritable bowel syndrome: sensitivity, specificity, and predictive values of pain sensory thresholds. *Gastroenterology* 2002; **122**: 1771–7.
19 Bharucha AE, Seide B, Fox JC, Zinsmeister AR. Day-to-day reproducibility of anorectal sensorimotor assessments in healthy subjects. *Neurogastroenterol Motil* 2004; **16**: 241–50.
20 Holtmann G, Goebell H, Talley NJ. Functional dyspepsia and irritable bowel syndrome: is there a common pathophysiological basis? *Am J Gastroenterol* 1997; **92**: 954–9.
21 Hammonds R, Houghton LA, Whorwell PJ. Whole gut sensitivity in constipation predominant irritable bowel syndrome: 80 percent of patients exhibit an abnormality. *Gastroenterology* 1998; **114**: 761.
22 Harraf F, Schmulson M, Saba L *et al.* Subtypes of constipation predominant irritable bowel syndrome based on rectal perception. *Gut* 1998; **43**: 388–94.
23 Prior A, Maxton DG, Whorwell PJ. Anorectal manometry in irritable bowel syndrome: differences between diarrhoea and constipation predominant subjects. *Gut* 1990; **31**: 458–62.
24 Plourde V, Lembo T, Shui Z *et al.* Effects of the somatostatin analogue octreotide on rectal afferent nerves in humans. *Am J Physiol* 1993; **265**: G742–51.
25 Lagier E, Delvaux M, Vellas B *et al.* Influence of age on rectal tone and sensitivity to distension in healthy subjects. *Neurogastroenterol Motil* 1999; **11**: 101–7.
26 Sun WM, Read NW. Anorectal function in normal human subjects: effect of gender. *Int J Colorectal Dis* 1989; **4**: 188–96.
27 Bradette M, Delvaux M, Staumont G, Fioramonti J, Bueno L, Frexinos J. Octreotide increases thresholds of colonic visceral perception in IBS patients without modifying muscle tone. *Dig Dis Sci* 1994; **39**: 1171–8.
28 Naliboff BD, Munakata J, Fullerton S *et al.* Evidence for two distinct perceptual alterations in irritable bowel syndrome. *Gut* 1997; **41**: 505–12.
29 Hammer J, Phillips SF, Talley NJ, Camilleri M. Effect of a 5HT3-antagonist (ondansetron) on rectal sensitivity and compliance in health and the irritable bowel syndrome. *Aliment Pharmacol Ther* 1993; **7**: 543–51.
30 Czimmer J, Suto G, Kiraly A, Mozsik G. Otilonium bromide enhances sensory thresholds of volume and pressure in patients with irritable bowel syndrome. *J Physiol Paris* 2001; **95**: 153–6.
31 Sabate JM, Gorbatchef C, Flourie B, Jian R, Coffin B. Cholecystokinin octapeptide increases rectal sensitivity to pain in healthy subjects. *Neurogastroenterol Motil* 2002; **14**: 689–95.
32 Delvaux M, Beck A, Jacob J, Bouzamondo H, Weber FT, Frexinos J. Effect of asimadoline, a kappa opioid agonist, on pain induced by colonic distension in patients with irritable bowel syndrome. *Aliment Pharmacol Ther* 2004; **20**: 237–46.
33 Kuiken SD, Tytgat GN, Boeckxstaens GE. The selective serotonin reuptake inhibitor fluoxetine does not change rectal sensitivity and symptoms in patients with irritable bowel syndrome: a double blind, randomized, placebo-controlled study. *Clin Gastroenterol Hepatol* 2003; **1**: 219–28.
34 Delvaux M, Louvel D, Mamet JP, Campos-Oriola R, Frexinos J. Effect of alosetron on responses to colonic distension in patients with irritable bowel syndrome. *Aliment Pharmacol Ther* 1998; **12**: 849–55.
35 van der Veek PP, Schots ED, Masclee AA. Effect of neurotensin on colorectal motor and sensory function in humans. *Dis Colon Rectum* 2004; **47**: 210–8.
36 Law NM, Bharucha AE, Undale AS, Zinsmeister AR. Cholinergic stimulation enhances colonic motor activity, transit, and sensation in humans. *Am J Physiol* 2001; **281**: G1228–37.
37 Thumshirn M, Coulie B, Camilleri M, Zinsmeister AR, Burton DD, Van Dyke C. Effects of alosetron on gastrointestinal transit time and rectal sensation in patients with irritable bowel syndrome. *Aliment Pharmacol Ther* 2000; **14**: 869–78.

