

Development and validation of the self-completed ascites impact measure to understand patient motivation for requesting a paracentesis

Bruce Crawford¹
Elizabeth Pault²
Walter Gotlieb³
Florence Joulain⁴

¹Mapi Values, Tokyo, Japan; ²Mapi Values, Boston, MA, USA; ³McGill University, Montreal, Quebec, Canada; ⁴Sanofi, Paris, France

Background: The Ascites Impact Measure (AIM) was developed to record patients' daily experiences of symptoms that trigger a request for a paracentesis.

Methods: Development of the AIM followed a rigorous step-wise approach, including a review of the literature, expert opinions, and qualitative research involving patients who experience symptomatic malignant ascites. The AIM's measurement properties were assessed using data from two international trials, including 59 ovarian cancer patients with symptomatic malignant ascites.

Results: Following the literature review and expert discussions to develop the conceptual model, ten patients with symptomatic malignant ascites were interviewed in the item elicitation phase, resulting in a draft questionnaire with four questions. Validation analyses consisted of 59 patients pooled from two international trials. Inter-items correlations for the AIM were good ($r > 0.60$), except for the Pain item. Internal consistency reliability ($\alpha = 0.89$) improved after removing the Pain item from the Total Symptom score (TSS). Test-retest reliability was sufficient. Scores significantly improved after paracentesis except for the Pain item. Preliminary estimates indicate that a two-point improvement on the three-item TSS (without the Abdominal Pain item) could be interpreted as clinically meaningful.

Conclusion: The Abdominal Pain item appears to behave differently than the other three items, and could be more related to cancer. While the validity of the AIM TSS (four-item) is acceptable, removing the Pain item from the TSS scoring algorithm demonstrated better construct validity. In addition, test-retest reliability and responsiveness were found to be similar to the results for the four-item AIM TSS. The Pain item should be used as a supplemental item to the three-item AIM TSS, as it provides additional information about the underlying cancer state.

Keywords: ascites impact measure, symptoms, ovarian cancer, paracentesis

Introduction

Ascites is characterized by an abnormal accumulation of fluid in the peritoneal cavity. Malignant ascites accounts for approximately 10% of all ascites cases.¹ Common primary tumor sites for cancers that result in malignant ascites are intra-abdominal, including the pancreas, stomach, lymphocytes, uterus, and ovaries, accounting for 30%–54% of all cases.^{2–8}

Malignant ascites is usually associated with a poor prognosis. Overall, patients with ovarian or hematologic cancers have better mean survival (32 and 58 weeks, respectively) compared to gastrointestinal cancer (12–20 weeks).^{9–11}

Symptom severity is associated with the increasing volume of ascites. The fluid pressure on abdominal organs results in abdominal and/or lower back pain, dyspnea,

Correspondence: Bruce Crawford
Mapi Values, Arco Tower 12F,
1-8-1 Shimomeguro, Meguro,
Tokyo 153-0064, Japan
Tel +81 3.5436.2070
Fax +81 3.5436.2071
Email bruce.crawford@mapivalues.com

dysphagia, and organ malfunction, including reduced bowel movements and increased urinary frequency.¹² Physical functioning can also be affected; for example, patients may not be able to tolerate even limited activity, such as bending over or standing up. The accumulation of fluid results in increased weight and a distended abdomen, which significantly affect the patients' body image and emotional well-being.^{13,14}

Malignant ascites does not respond to salt restriction or diuretic therapy. In addition, surgical insertion of peritoneovenous shunts is associated with severe side effects. Therefore, palliative paracentesis is the only therapeutic option for chemotherapy-resistant patients.¹⁵ In a survey of physicians who treat symptomatic malignant ascites, paracentesis was employed by 98% and was felt to be effective by 89%.¹⁶ The limitation of paracentesis is that its mean duration of symptomatic relief is approximately ten days, so that repeated paracentesis is required due to the reaccumulation of ascites. Furthermore, repeated drainage of protein-rich ascites over time depletes protein reserves, leading to more leakage of fluids in the peritoneal cavity and quicker reaccumulation of ascites. The risk of bowel or vascular injury exists.^{17,18} As a result of the symptoms, most patients repeatedly seek medical advice and notify their healthcare provider when fluid starts to reaccumulate. Symptoms that drive patients' decisions to seek paracentesis for relief have not been documented using a standardized approach.

An examination of published self-completed instruments identified the Edmonton Symptom Assessment Scale (ESAS)-Ascites Modification (AM) questionnaire as a relevant instrument for patients with ascites.¹⁹ The ESAS-AM is a modified version of the ESAS that includes two ascites-specific symptoms: the ability to move normally and the amount of abdominal bloating or discomfort. The two additional symptoms were the most common symptoms of a list of eight symptoms derived from a survey of ascites patients. In a longitudinal pre- and post-intervention cohort study, the ESAS-AM demonstrated acceptable sensitivity, reliability, and responsiveness to change. However, as the items from the ESAS-AM might not have been spontaneously elicited from patients, the content validity of this questionnaire might be questioned. For instance, it might be difficult to document how patients interpret "No abdominal discomfort/bloating", as this incorporates two concepts.

The objective in this article is to report on the development and validation process of a self-completed instrument, the Ascites Impact Measure (AIM), in order to record the severity of symptoms and what triggers a patient to request a paracentesis.

Methods

Instrument development

Development of the AIM started in 2004 and followed a rigorous step-wise approach, including a review of the literature, expert opinions, and qualitative research involving patients experiencing symptomatic malignant ascites (SMA).

Classic biomedical databases (Medline, PubMed, Ovid) and an in-house Quality of Life Instruments database were searched for relevant literature. Publications referenced in key articles were gathered.

Symptoms retrieved from the literature were submitted for review by a clinician and were used to generate the preliminary items. To ensure linguistic and conceptual equivalence of the AIM, items were generated simultaneously in English and French by native speakers of each language.

Patients' spontaneous and unbiased opinions were then collected via face-to-face or telephone interviews in order to record the symptoms that best characterize their perceptions of the condition.

Recruitment of sample population

Adult patients with cytologically confirmed malignant ascites (International Classification of Diseases [ICD]-10 CM code C78.6) who required at least one to four paracenteses per month to manage abdominal discomfort attributable to SMA and resistance (relapse or progression of disease during or after treatment or drug intolerance) to at least two lines of prior chemotherapy treatments were considered for inclusion. Exclusion criteria included pseudomyxoma peritonei, spinal cord compression, or carcinomatous meningitis. A recruitment agency used cancer-related websites and mass emails and three oncologists spoke with patients at their clinics to ultimately recruit ten patients over a ten-month period. Each patient consented to participate in the study, in accordance with Health Insurance Portability and Accountability Act requirements. An oncologist confirmed their eligibility.

Qualitative interviews and data collection

A semistructured interview guide, including open-ended questions for concept elicitation, was developed based on a review of the literature. Patients were asked to provide spontaneous responses to define their symptoms, gauge the severity of the symptoms that prompted them to request a paracentesis, and to answer confirmatory questions regarding the symptoms described in the literature. The face and content validity of the draft AIM were tested during the interview as a second phase to verify the relevance, comprehensiveness, applicability, acceptability, and patient comprehension of

the items, as well as to confirm that the items appeared to measure what they intended to measure.

After all interviews were completed and transcribed, qualitative analyses following grounded theory methodology were conducted.^{20,21} Spontaneously reported concepts (for example, symptoms attributed to ascites that drive patients' decisions to undergo a paracentesis) were compared with concepts reported in the literature and served to finalize the AIM conceptual framework. Responses from the face and content validity phase of each interview were collated and grouped by item and response option. Patterns and inconsistencies in patients' understanding of the items were identified and reported. Specific patient quotes were documented to substantiate patients' decisions to keep, modify, or delete an item. Any questions that patients found difficult to understand or interpret were subsequently revised. The AIM questionnaire was then finalized after a review by clinical experts.

Psychometric validation of the AIM

The psychometric performance of the AIM was documented based on data pooled from a multicenter, open-label, single-arm study and on blinded data from a multicenter, randomized, double-blind, placebo-controlled study of intravenous aflibercept (VEGF Trap) for the treatment of recurrent SMA in patients with advanced ovarian cancer.

Sample population

Patients with recurrent symptomatic malignant ascites were registered the day after a paracentesis (pre-registration). The AIM was completed in the evening at approximately the same time for two months, including the day prior to the pre-registration for the paracentesis. The Functional Assessment of Cancer Therapy-Ovarian Cancer questionnaire (FACT-O)^{22,23} was completed at the beginning of each 21-day cycle. The FACT-O is a self-report measure that assesses four dimensions of well-being: physical (seven items), functional (seven items), social/family (seven items), and emotional (six items) and also includes 11 additional concerns that relate specifically to ovarian cancer symptoms.

Statistical analyses

Gold standard psychometric tests^{24,25} were used to define the scoring algorithm for a total symptom score (referred to as AIM TSS) and then to document the cross-sectional and longitudinal properties of the AIM. Distributional characteristics of the AIM individual symptom scores and TSS were documented.

The construct validity of the TSS was assessed using multi-trait, multi-item analyses, including inter-item correlations. A correlation of lower than 0.80 indicated there was no redundancy between items. Internal consistency reliability was considered to be sufficient with a Cronbach's alpha coefficient (α) greater than or equal to 0.70.^{26,27} Data collected at the pre-registration paracentesis were used in these analyses.

Concurrent validity of the items and the TSS was supported if a logical pattern of convergent and discriminant correlations was found with the relevant concepts from the FACT-O. Data from Day 1/Cycle 1, when patients had completed the two instruments, were used. Items with similar content should correlate more highly (Pearson correlation > 0.40) than items with less similar content.

Clinical validity was assessed based on correlations calculated between the clinical measures (weight, abdominal girth) and the items and the TSS of the AIM using changes in scores. The AIM scores and clinical measures were generated the day before and the day after the pre-registration paracentesis.

Test-retest correlations were assessed seven days apart at time points when the patient was anticipated to be in the equivalent clinical state, on Day 8 of Cycle 1 (test) and Day 1 of Cycle 2 (retest). Each treatment cycle lasted 14 days. In order to ensure the test-retest analysis was restricted to a clinically stable sample, the analysis was performed on data from patients who had maintained a stable weight ($\pm 5\%$) or the same abdominal girth ($\pm 5\%$) at the two assessments and who also completed the AIM. It is of note that the threshold of 5% was deemed clinically relevant to account for natural worsening of the condition.

As a paracentesis is expected to relieve the symptoms of ascites, the ability of the AIM individual scores and the TSS to discriminate between patients' status was evaluated using data before and after the registration paracentesis. Paired *t*-tests were used to compare the mean change between the day before and the day after registration paracentesis. The level of change in the domain was quantified using Cohen's standardized effect size (SES). Within-group effect sizes were obtained by taking the mean change in scores (of the domain) between the day before and the day after the registration paracentesis and dividing it by the standard deviation (SD) obtained from the scores on the day before the registration paracentesis. The SES was expected to be moderate (in other words, ≥ 0.50).^{28,29}

AIM score interpretation

The smallest change in scores that is clinically meaningful, referred to as the minimal important difference (MID), was

assessed for exploratory purposes. Using a distributional approach, the MID was estimated with the 1/2 SD method³⁰ using data from the day before pre-registration paracentesis, Day 1 of Cycle 1, and Day 1 of Cycle 2. Since a global assessment of change was not included in the two trials, the MID could not be estimated using an anchor-based approach.

Results

Instrument development

Sample population

A total sample of ten patients with cancer-related SMA participated in the study, of which 80% spoke US English and 20% spoke Canadian French. The primary cancer site was the ovary in most patients (60%) (Table 1).

Development of the conceptual framework and item generation

Following expert discussion, it was hypothesized that the main symptoms driving patients to request a paracentesis are abdominal pain, abdominal bloating, and abdominal discomfort. Other symptoms identified in the literature, including lack of appetite and early satiety, low back pain, the urge to urinate, urinary frequency, constipation, difficulty breathing/dyspnea, nausea, vomiting, and ankle swelling, were disregarded for inclusion in the draft questionnaire as they were determined to be less clinically relevant, infrequently reported, or otherwise considered distal effects of ascites. In addition, it was assumed that the patients' abilities to perform daily activities and move normally would be primary impairments in the patients' physical functioning that would lead to a request for paracentesis and therefore were added. Recognizing that all symptoms deemed relevant to the patients may not have been recorded in the literature or in the clinician's feedback, one item was left "blank" in the draft six-item AIM questionnaire. The questions were designed to record the severity of each item identified using a six-point Likert severity scale that ranged from 0 = No symptoms to 5 = Very severe symptoms.

Table 1 Sample characteristics for the development phase of AIM (n = 10)

	US English (n = 8)	Canadian French (n = 2)	Total (n = 10)
Female (%)	7 (85.7)	2 (100.0)	9 (90.0)
Primary cancer site (%)			
Ovary	5 (62.5)	1 (50.0)	6 (60.0)
Colorectum	1 (12.5)		1 (10.0)
Peritoneal cavity	2 (25.0)		1 (10.0)
Fallopian tube		1 (50.0)	1 (10.0)

Considering that ascites-related symptoms fluctuate, subjective reporting of these symptoms was determined to be "time-dependent." As multiple direct reports of symptoms on a diary-type questionnaire have the advantage of integrating "symptom evolution" into the assessment of patients' decision-making processes, the AIM was adapted for daily completion using a 24-hour recall period.

The ascites symptoms spontaneously reported by patients with cancer-related SMA during the in-depth qualitative interviews included abdominal discomfort (n = 10), abdominal bloating (n = 10), abdominal pain (sharp pain, pain on the side) (n = 10), urinary frequency (n = 3), difficulty breathing (n = 3), digestion problems (n = 6), lower back pain (n = 6), difficulty moving (n = 9), and difficulty performing daily activities (n = 4). Most patients (n = 7) differentiated "pain" from "discomfort" (for example, "Discomfort you can live with. Pain really hurts, and you don't want to live with that"; and "Discomfort is something that's just like an annoyance and it will go away or come and go. Pain is, to me, it's something that it just doesn't go away") (Table 2). Patients' symptoms were described as evolving at various rates (for example, abdominal bloating was noticeable within approximately ten days after a paracentesis and abdominal pain became more severe a few days prior to the paracentesis and often subsided for a few days after).

Table 2 Examples of patient quotes for the development phase of AIM (n = 10)

Patient quotes	AIM questionnaire item
"Most of the time it's generally discomfort." "You're just really uncomfortable." "It's very hard to find a comfortable position either sitting or standing or laying in bed." "My stomach goes from 38 to 44 [inches]." "A feeling of bloating and it just continues and continues until it just distends the stomach so much." "I felt like I was going to burst." "The bloating in the stomach is just horrible" "It's like being 12 months pregnant, it seems that I am going to explode."	Average level of abdominal discomfort due to ascites (n = 10) Average level of abdominal bloating due to ascites (n = 10)
"During the night I might get sharp pains." "Just the pain was excruciating." "When I just can't stand the pain and pressure anymore, when it's just constant 24 hours a day." "My walking is so bad, one of my daughters pushes me in a wheelchair into the hospital." "Getting out of bed, moving." "I couldn't walk across the room." "I couldn't roll over in bed." "I just can't hardly manipulate around."	Average level of abdominal pain due to ascites (n = 10) Impact of ascites on ability to move normally (n = 9)

Patients mentioned that their symptoms were worse in the afternoon and evening ($n = 6$) and after eating ($n = 4$). Patients also concurred that daily evening completion of a short questionnaire would not be too burdensome, yet was pertinent with regard to the rapid evolution of the ascites-related symptoms between two successive paracenteses ($n = 9$) (for example, “I think I would find it interesting [to complete this questionnaire daily] because I don’t monitor [my symptoms] myself.”).

Content validity

Most patients ($n = 8$) confirmed their involvement in the decision to undergo paracentesis. When presented with the symptoms and impact of ascites that were spontaneously reported in the first part of the interview, patients indicated that the severity of their abdominal discomfort ($n = 10$), abdominal bloating ($n = 9$), abdominal pain ($n = 7$), the impact of ascites on the ability to move normally ($n = 10$), and the impact of ascites on usual daily activities ($n = 7$) were relevant factors triggering their request for a paracentesis. In addition, patients found the response options to be clear, sufficiently detailed, and appropriate to adequately record the change in the severity of their symptoms.

Psychometric validation of the AIM

Sample population

A total sample of 59 patients who participated in either an open-label study ($n = 14$) or in a randomized, placebo-controlled, double-blind study ($n = 45$) completed the AIM at least once between the day before the pre-registration paracentesis and the day after the first post-registration paracentesis.

The sample consisted predominantly of Caucasian females (81.36%). The average age of patients was 58.73 years ($SD = 11.31$), ranging between 33 and 88 years. Approximately two thirds of the patients were diagnosed with stage III ovarian cancer (69.5%). Patients most commonly had serious ovarian cancer and the majority of tumors were poorly differentiated (57.6%) (Table 3). Patients were heavily pretreated with a median of four prior lines of chemotherapy. After the pre-registration paracentesis, the patients’ weight decreased by -2.49 kg ($SD = 1.89$, range -7.00 – 1.40) on average, and their abdominal girth decreased by -6.09 cm ($SD = 10.08$, range -62.4 – 1.0).

Item psychometric performance

Distributional characteristics

The day before the registration paracentesis, most of the patients indicated moderate to severe abdominal discomfort

Table 3 Sample characteristics for the validation phase of AIM ($n = 59$)

	Pivotal study ($n = 45$)	Open-label study ($n = 14$)	Pooled data ($n = 59$)
Age			
Mean (SD)	58.82 (12.01)	58.43 (9.09)	58.73 (11.31)
Median	56.00	57.50	57.00
Min, max	33.00, 88.00	40.00, 71.00	33.00, 88.00
Race			
Caucasian	34 (75.56%)	14 (100.0%)	48 (81.36%)
Black	1 (2.22%)	0 (0.00%)	1 (1.69%)
Asian, oriental	10 (22.22%)	0 (0.00%)	10 (16.95%)
Primary site			
Ovaries	38 (84.44%)	14 (100.0%)	52 (88.14%)
Fallopian tube	1 (2.22%)	0 (0.00%)	1 (1.69%)
Peritoneum	6 (13.33%)	0 (0.00%)	6 (10.17%)
History			
Adenocarcinoma	3 (6.67%)	0 (0.00%)	3 (5.26%)
Serous carcinoma	33 (73.33%)	7 (58.33%)	40 (70.18%)
Mucinous carcinoma	1 (2.22%)	0 (0.00%)	1 (1.75%)
Endometrioid carcinoma	3 (6.67%)	3 (25.00%)	6 (10.53%)
Clear cell (mesonephroid) carcinoma	1 (2.22%)	1 (8.33%)	2 (3.51%)
Unclassified carcinoma	3 (6.67%)	0 (0.00%)	3 (5.26%)
Other	1 (2.22%)	1 (8.33%)	2 (3.51%)
Missing/no response	0 (0.00%)	2 (14.29%)	2 (3.39%)
Grade			
Unknown	10 (22.22%)	5 (35.71%)	15 (25.42%)
Moderately differentiated	6 (13.33%)	1 (7.14%)	7 (11.86%)
Poorly differentiated	26 (57.78%)	8 (57.14%)	34 (57.63%)
Well differentiated	3 (6.67%)	0 (0.00%)	3 (5.08%)
Staging at diagnosis			
Stage IA	11 (24.44%)	1 (7.14%)	12 (20.34%)
Stage III	1 (2.22%)	1 (7.14%)	2 (3.39%)
Stage IIIA	1 (2.22%)	0 (0.00%)	1 (1.69%)
Stage IIIB	18 (40.00%)	1 (7.14%)	19 (32.32%)
Stage IIIC	13 (28.89%)	8 (57.14%)	21 (35.71%)
Stage IV	1 (2.22%)	3 (21.43%)	4 (6.78%)
Missing/no response/unknown	0 (0.00%)	0 (0.00%)	0 (0.00%)

Notes: Percentages are based on the number of non-missing responses; percentages of missing/no response are based on the overall number of patients.

(68.42%), abdominal bloating (57.90%), and the ability to move normally (63.16%). Patients’ ratings for the Abdominal Pain item were spread across severity levels (Table 4).

Inter-item correlations

The Abdominal Discomfort, Abdominal Bloating, and Ability to Move Normally items appeared to be related but not redundant concepts (Pearson inter-item correlation range: 0.69–0.75). Abdominal Pain yielded smaller correlations

Table 4 Item response distribution using the day before registration paracentesis (n = 59)

Item	No	Very mild	Mild	Moderate	Severe	Very severe
Abdominal discomfort	0 (0.0%)	4 (10.53%)	4 (10.53%)	12 (31.58%)	14 (36.84%)	4 (10.53%)
Abdominal bloating	3 (7.89%)	2 (5.26%)	5 (13.16%)	11 (28.95%)	11 (28.95%)	6 (15.79%)
Abdominal pain	6 (15.79%)	7 (18.42%)	7 (18.42%)	10 (26.32%)	5 (13.16%)	3 (7.89%)
Ability to move normally	3 (7.89%)	3 (7.89%)	6 (15.79%)	12 (31.58%)	12 (31.58%)	2 (5.26%)

with the previously mentioned items (Pearson inter-items correlation range: 0.28–0.49) (Table 5).

Concurrent validity

The correlation between the FACT-O Functional Well-Being (r range: –0.31–0.14) and Physical Well-Being scale (r range: –0.53–0.31) with each AIM item score was low to moderate, indicating they measure different concepts (Table 6). The FACT-O Pain item and Cramp item yielded moderate correlations (r = –0.54 and r = –0.67, respectively) with the AIM Abdominal Pain item.

Clinical validity

Low to moderate correlations were found between the changes in scores for Abdominal Bloating, Abdominal Discomfort, and Ability to Move Normally with changes in weight and changes in abdominal girth before and after a paracentesis. Changes in abdominal pain showed no correlation with changes in weight (r = –0.03) and a low correlation with abdominal girth (r = 0.35).

Test-retest reliability

Using weight to define stable patients (n = 20), the test-retest correlations for the Abdominal Discomfort, Abdominal Bloating, and Ability to Move Normally items were above 0.70 (range: 0.71–0.83). The lowest test-retest correlation was shown for the Abdominal Pain item (r = 0.62). Using abdominal girth to define stable patients (n = 12), the test-retest correlations for all individual items were below 0.70 (range: 0.24–0.64) (Table 7).

Responsiveness

All items showed a statistically significant improvement of scores after the pre-registration paracentesis

(Abdominal Discomfort item: $P < 0.0001$, SES = –1.25; Abdominal Bloating item: $P < 0.0001$, SES = –1.14; and Ability to Move Normally item: $P < 0.0001$, SES = –0.97). The change in scores for the Abdominal Pain item was significant ($P = 0.0084$), yet yielded a low SES (SES = –0.33) (Figure 1).

Total symptom score scoring algorithm

The AIM TSS was originally defined as the sum of the scores of the four items: Abdominal Discomfort, Abdominal Bloating, Abdominal Pain, and Ability to Move Normally. The AIM TSS (four items) mean score was 11.53 (SD = 4.39; median = 11.5; possible score range 0–20) and did not yield any floor or ceiling effects. Following examination of the psychometric performance of each item, a discussion with three clinician oncologists resulted in reconsidering including the Abdominal Pain item in the scoring algorithm. After removing the Abdominal Pain item, the AIM TSS (three items) mean score was 9.26 (SD = 3.50; median = 10; possible score range 0–15) and did not yield any floor or ceiling effects. Removing the Abdominal Pain item from the total score increased the Cronbach's alpha to 0.89 ($\Delta = 0.05$), indicating improvement in the internal consistency reliability. Removal of any of the other three items resulted in lower α values (Table 7). The psychometric performance of the four-item TSS and its shorter version (three-item TSS) were compared to further assess whether the Pain item should be removed from the TSS.

Total symptom score psychometric performance

The AIM TSS showed a small correlation with the FACT-O Additional Concerns scale, which is specific to ovarian

Table 5 Item score and inter-item correlations the day before registration paracentesis (n = 59)

Item	Mean score (standard deviation)	Abdominal discomfort	Abdominal bloating	Abdominal pain	Ability to move normally
Abdominal discomfort	3.50 (1.10)	1.00			
Abdominal bloating	3.50 (1.14)	0.75	1.00		
Abdominal pain	2.45 (1.57)	0.49	0.28	1.00	
Ability to move normally	2.95 (1.25)	0.75	0.69	0.45	1.00

Table 6 Concurrent validity of AIM with FACT-O at day 1 of cycle 1 (n = 59)

AIM/ FACT-O	Abdominal discomfort	Abdominal bloating	Abdominal pain	Ability to move normally	AIM TSS (three-item)	AIM TSS (four-item)
Physical well-being (n = 42)	-0.40	-0.31	-0.39	-0.53	-0.45	-0.49
Social/family well-being (n = 41)	0.01	-0.10	0.15	-0.09	-0.07	-0.01
Emotional well-being (n = 41)	-0.09	-0.14	-0.03	-0.28	-0.17	-0.14
Functional well-being (n = 41)	-0.30	-0.14	-0.23	-0.31	-0.28	-0.28
Additional concerns (n = 41)	-0.32	-0.20	-0.27	-0.31	-0.31	-0.33
Have pain (n = 41)	-0.37	-0.26	-0.54	-0.48	-0.37	-0.48
Have cramps in my stomach area (n = 41)	-0.27	-0.22	-0.67	-0.27	-0.25	-0.42
Swelling in my stomach area (n = 40)	-0.48	-0.42	-0.31	-0.54	-0.49	-0.50

cancer ($r = -0.31$ and $r = -0.33$ for the three- and four-item AIM TSS, respectively), the FACT-O Functional Well-Being scale ($r = -0.28$), and the Physical Well-Being scale ($r = -0.45$ and 0.49) (Table 6). Low to moderate correlations were found between the change in AIM TSS with a change in weight ($r = 0.31$ and $r = 0.24$ for the three- and four-item AIM TSS, respectively) and abdominal girth ($r = 0.25$ and 0.29) (Figure 2).

Test-retest correlations for the three- and four-item AIM TSS were 0.82 and 0.77, respectively, when using weight to define stable patients, and were 0.64 and 0.53, respectively, when using abdominal girth to identify stable patients (Table 8). The three- and four-item AIM-TSS showed statistically significant ($P < 0.0001$) improvement of scores on the day after registration paracentesis compared to the day before registration paracentesis (mean change [SD] = 4.39 [4.29] and 5.00 [5.06] for the three- and four-item AIM TSS, respectively), yielding a large effect size (SES = -1.22 and -1.09, respectively).

Interpretation of total symptom score

The 1/2 SD on the day before the registration paracentesis, Day 1 of Cycle 1, or Day 1 of Cycle 2 for the three- and four-item AIM TSS varied from 1.75 to 2.22 and from 2.20 to 2.82, respectively. These results indicated that the MID, based on a distribution approach, could be set at 2 for the three-item AIM TSS and at 2.5 for the four-item AIM TSS (Table 9).

Table 7 Internal consistency reliability the day before registration paracentesis (n = 59)

Item	Alpha	
	Using four items	Using three items
AIM TSS	0.84	0.89
Abdominal discomfort	0.73	0.82
Abdominal bloating	0.79	0.86
Abdominal pain	0.89	/
Ability to move normally	0.76	0.86

Discussion

The primary goal of this study was to document patient motivation for considering a therapeutic paracentesis to treat ascites and provide symptom relief.

The low prevalence of the condition, a high mortality rate, the advanced condition of subjects, and strict study inclusion criteria were challenging factors for the qualitative research. Ten patients were recruited in the US and Canada over a ten-month period, despite extensive efforts to recruit additional patients (80 investigators were contacted and 31,650 advanced cancer patients in the US were approached via e-mail). In order to account for these difficulties, a modified approach in developing the AIM questionnaire was employed. Sensitizing concepts identified in published evidence and by expert opinion were used to hypothesize the conceptual framework and draft preliminary items. Confirmatory concept elicitation interviews, and face and content validity testing performed during a 1-hour interview allowed the conceptual framework and AIM to be finalized.

The low prevalence of the condition was equally as challenging in documenting the psychometric performance of the new instrument. The validation analyses were thus performed on data collected in two trials that had the same schedule of observations for instruments and clinical variables. In an effort to increase the sample size, the studies were run internationally over three years across 24 countries (26 linguistic versions of the AIM were developed for the purpose of the studies). Examination of the score distribution, percentage of missing data, item convergent validity, inter-item and item-score correlations, and internal consistency reliability of the original English version of the AIM compared with the other linguistic versions combined did not yield any discrepancies (data not shown). These results, combined with the thorough translation of the AIM involving forward and backward translations^{31,32} and consistent interpretation of the items from patients who participated in 30-minute qualitative interviews for linguistic validation purposes

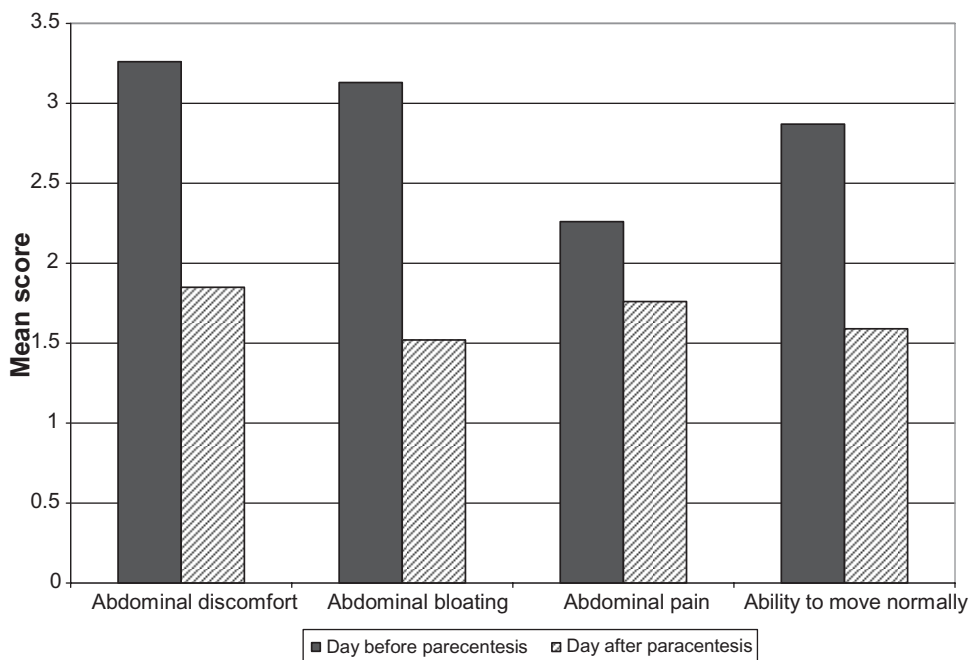


Figure 1 Responsiveness of AIM item scores (n = 59).

(one Swedish, one Israeli Hebrew, one Italian, and one UK English), supported the conceptual equivalence and content validity of the linguistic versions of the AIM. In addition, results for the psychometric analyses were either consistent between the two studies or their interpretation was otherwise limited by the small sample size, thus warranting caution regarding any conclusions.

The Abdominal Pain item behaved differently from the other items. The fact that the Abdominal Pain item showed poor responsiveness after a paracentesis and that at least six patients did not report any pain while eight patients reported worsening pain after paracentesis led a working group of experts to reconsider the clinical relevance of this item as a symptom relevant to ascites and whether to include it in

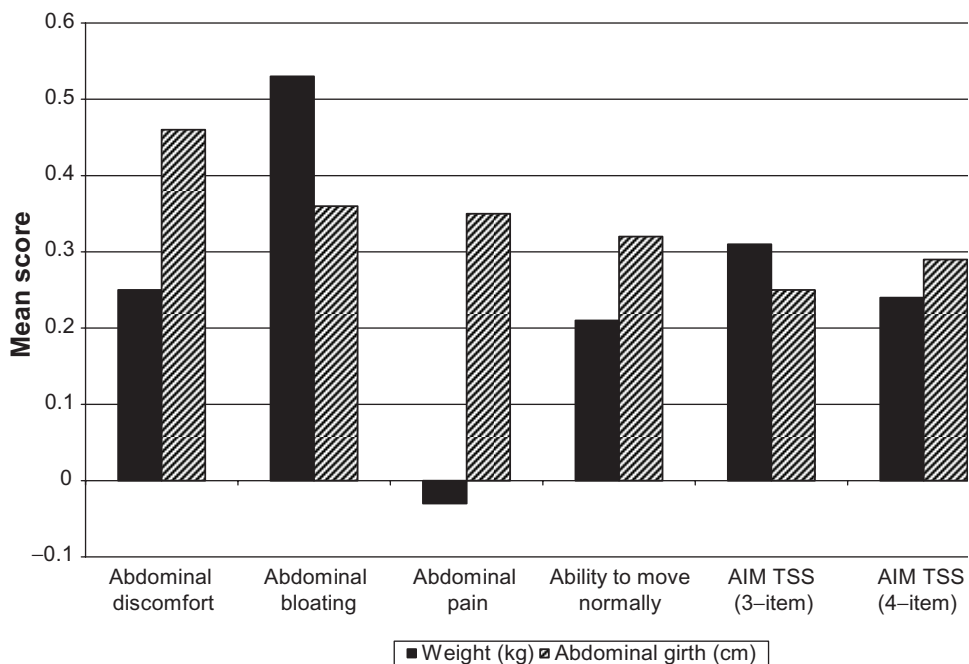


Figure 2 Clinical validity of AIM: changes in AIM item scores correlated with changes in clinical markers for the three- and four-item TSS.

Table 8 Test-retest correlation of AIM item scores at cycle 1/ day 8 and cycle 2/day 1

Item	Correlation for patients with stable weight n = 20 ^a	Correlation for patients with stable abdominal girth n = 12 ^b
Abdominal discomfort	0.71	0.53
Abdominal bloating	0.81	0.51
Abdominal pain	0.62	0.24
Ability to move normally	0.83	0.56
AIM TSS (three-item)	0.82	0.64
AIM TSS (four-item)	0.77	0.53

Notes: ^aResults only for stable patients with non-missing value for variable at both time points. Stable patients defined as patients who have the same weight ($\pm 5\%$) at both time points; ^bresults only for stable patients with non-missing value for variable at both time points. Stable patients defined as patients who have the same abdominal girth ($\pm 5\%$) at both time points.

the AIM TSS. Even if the AIM (four-item) is a valid and reliable questionnaire, removing the Abdominal Pain item from the AIM TSS allowed reaffirmation of the instrument's content validity. This decision was corroborated by the findings of another study, specifically that pain, as measured by the Edmonton Symptom Assessment System-Ascites Modification (ESAS-AM), the Memorial Symptom Assessment Scale-Short Form (MSAS-SF), and the Quality of Life Questionnaire-Core 30 (QLQ-C30), was only minimally relieved after a paracentesis for patients with advanced intra-abdominal malignancy.³³ Furthermore, the TSS showed more robust construct validity results when the Abdominal Pain item was removed. Nevertheless, changes in the individual score for Abdominal Pain should be considered in order to better understand patients' experience with ovarian cancer and the AIM TSS.

The small sample size (n = 59) and use of proxy assessments to define stable patients (weight and abdominal girth) justify the need to confirm the results in another study; it is

Table 9 Minimum important difference (distributional approach) for the three- and four-item TSS

TSS	Time point	1 standard deviation	0.5 standard deviation
Three-item AIM-TSS	Day before pre-registration paracentesis	3.50	1.75
	Cycle 1/day 1	4.04	2.02
	Cycle 2/day 1	4.44	2.22
Four-item AIM-TSS	Day before pre-registration paracentesis	4.39	2.20
	Cycle 1/day 1	5.05	2.52
	Cycle 2/day 1	5.64	2.82

also generally helpful to use a variety of methods to confirm the choice of an MID^{34,35} and additional studies should be conducted with appropriate patient-based anchors (for example, a Patient Global Impression of Change scale) to generate more data before finalizing the MID. The development and validation of a rating scale is an iterative process; every clinical research study will strengthen the evidence regarding the usefulness of the instrument in this population and presumably for other ascites-related oncology conditions.

Disclosure

This study was supported by Sanofi. Walter Gotlieb received honoraria and support for travel and manuscript activities from Sanofi and has acted as a consultant and advisor to Sanofi. Bruce Crawford and Elizabeth Pault are employed by Mapi Values, a consultancy firm retained by Sanofi for the development and validation of the instrument. Florence Joulain is an employee of Sanofi.

References

- Runyon BA. Care of patients with ascites. *N Engl J Med.* 1994;330:337–342.
- Parsons SL, Lang MW, Steele RJ. Malignant ascites: a 2-year review from a teaching hospital. *Eur J Surg Oncol.* 1996;22:237–239.
- Wilailak S, Linasmita V, Srivannaboon S. Malignant ascites in female patients: a seven-year review. *J Med Assoc Thai.* 1999;82:15–19.
- Ringenberg QS, Doll DC, Loy TS, Yarbrow JW. Malignant ascites of unknown origin. *Cancer.* 1989;64:753–755.
- Mackey JR, Venner PM. Malignant ascites: demographics, therapeutic efficacy and predictors of survival. *Can J Oncol.* 1996;6:474–480.
- Garrison RN, Kaelin LD, Galloway RH, Heuser LS. Malignant ascites. Clinical and experimental observations. *Ann Surg.* 1986;203:644–651.
- Appelqvist P, Silvo J, Salmela L, Kostianen S. On the treatment and prognosis of malignant ascites: is the survival time determined when the abdominal paracentesis is needed? *J Surg Oncol.* 1982;20:238–242.
- Hird V, Thomas H, Stewart JS, Epenetos AA. Malignant ascites: review of the literature, and an update on monoclonal antibody-targeted therapy. *Eur J Obstet Gynecol Reprod Biol.* 1989;32:37–45.
- Ariel IM, Oropeza R, Pack GT. Intracavitary administration of radioactive isotopes in the control of effusions due to cancer. Results in 267 patients. *Cancer.* 1966;19:1096–1102.
- Garrison RN, Kaelin LD, Galloway RH, Heuser LS. Malignant ascites. Clinical and experimental observations. *Ann Surg.* 1986;203:644–651.
- Jackson GL, Blosser NM. Intracavitary chromic phosphate (32P) colloidal suspension therapy. *Cancer.* 1981;48:2596–2598.
- Gotlieb WH, Feldman B, Feldman-Moran O, Zmira N, Kreizer D, Segal Y, et al. Intraperitoneal pressures and clinical parameters of total paracentesis for palliation of symptomatic ascites in ovarian cancer. *Gynecol Oncol.* 1998;71:381–385.
- Husain A, Bezjak A, Eason A. Malignant ascites symptom cluster in patients referred for paracentesis. *Ann Surg Oncol.* 2010;17:461–469.
- Woope H, Schouli J. Current and future options in the treatment of malignant ascites in ovarian cancer. *Anticancer Res.* 2009;29:3353–3359.
- Adam RA, Adam YG. Malignant ascites: past, present, and future. *J Am Coll Surg.* 2004;198:999–1011.

16. Ross GJ, Kessler HB, Clair MR, Gatenby RA, Hartz WH, Ross LV. Sonographically guided paracentesis for palliation of symptomatic malignant ascites. *AJR Am J Roentgenol.* 1989;153:1309–1311.
17. Gines P, Tito L, Arroyo V, Planas R, Panes J, Viver J, et al. Randomized comparative study of therapeutic paracentesis with and without intravenous albumin in cirrhosis. *Gastroenterology.* 1988;94:1493–1502.
18. Gough IR, Balderson GA. Malignant ascites. A comparison of peritoneovenous shunting and nonoperative management. *Cancer.* 1993;71:2377–2382.
19. Easson AM, Bezjak A, Ross S, Wright JG. The ability of existing questionnaires to measure symptom change after paracentesis for symptomatic ascites. *Ann Surg Oncol.* 2007;14:2348–2357.
20. Mays B, Pope C. Rigor and qualitative research. *Br Med J.* 1997; 109–112.
21. Strauss A, Corbin J. Basics of qualitative research: Techniques and procedures for developing grounded theory. London: Sage Publications, Inc; 1998.
22. Basen-Engquist K, Bodurka-Bevers D, Fitzgerald MA, Webster K, Cella D, Hu S, et al. Reliability and validity of the functional assessment of cancer therapy-ovarian. *J Clin Oncol.* 2001;19:1809–1817.
23. Cella DF, Tulsky DS, Gray G, et al. The Functional Assessment of Cancer Therapy (FACT) scale: Development and validation of the general measure. *J Clin Oncol.* 1993;11:570–579.
24. Atkinson MJ, Sinha A, Hass SL, Colman SS, Kumar RN, Brod M, et al. Validation of a general measure of treatment satisfaction, the Treatment Satisfaction Questionnaire for Medication (TSQM), using a national panel study of chronic disease. *Health Qual Life Outcomes.* 2004;2:12.
25. Nunnally J, Bernstein I. *Psychometric theory.* New York: McGraw Hill; 1994.
26. Cronbach LJ. Coefficient alpha and internal structure of tests. *Psychometrika.* 1951;16:297–334.
27. Nunnally J. The assessment of reliability. In Bernstein I, editor. *Psychometric Theory.* New York: McGraw Hill; 1994:248–292.
28. Cohen J. *Statistical Power Analysis for the Behavioral Sciences.* Hillsdale, NJ: Lawrence Erlbaum Associates; 1988.
29. Kazis LE, Anderson JJ, Meenan RF. Effect sizes for interpreting changes in health status. *Med Care.* 1989;27:S178–S189.
30. Norman GR, Sloan JA, Wyrwich KW. Interpretation of changes in health-related quality of life: the remarkable universality of half a standard deviation. *Med Care.* 2003;41:582–592.
31. Chassany O, Sagnier P, Marquis P, et al. Patient-Reported Outcomes: The Example of Health-Related Quality of Life – A European guidance document for the improved integration of Health-Related Quality of Life assessment in the drug regulatory process. *DIA Journal.* 2002;36:209–238.
32. Wild D, Grove A, Martin M, Eremenco S, McElroy S, Verjee-Lorenz A, et al. Principles of Good Practice for the Translation and Cultural Adaptation Process for Patient-Reported Outcomes (PRO) Measures: report of the ISPOR Task Force for Translation and Cultural Adaptation. *Value Health.* 2005;8:94–104.
33. Easson AM, Bezjak A, Ross S, Wright JG. The ability of existing questionnaires to measure symptom change after paracentesis for symptomatic ascites. *Ann Surg Oncol.* 2007;14:2348–2357.
34. Crosby RD, Kolotkin RL, Williams GR. Defining clinically meaningful change in health-related quality of life. *J Clin Epidemiol.* 2003;56: 395–407.
35. Revicki D, Hays RD, Cella D, Sloan J. Recommended methods for determining responsiveness and minimally important differences for patient-reported outcomes. *J Clin Epidemiol.* 2008;61:102–109.

Patient Related Outcome Measures

Publish your work in this journal

Patient Related Outcome Measures is an international, peer-reviewed, open access journal focusing on treatment outcomes specifically relevant to patients. All aspects of patient care are addressed within the journal and practitioners from all disciplines are invited to submit their work as well as healthcare researchers and patient support groups. Areas covered will

Submit your manuscript here: <http://www.dovepress.com/patient-related-outcome-measures-journal>

include: Quality of life scores; Patient satisfaction audits; Treatment outcomes that focus on the patient; Research into improving patient outcomes; Hypotheses of interventions to improve outcomes; Short communications that illustrate improved outcomes; Case reports or series that show an improved patient experience; Patient journey descriptions or research.

Dovepress