Quality of Life Research

Measuring what matters MOST: Validation of the Measure of Ovarian Symptoms and Treatment, a patient-reported outcome measure of symptom burden and impact of chemotherapy in recurrent ovarian cancer (ROC)

--Manuscript Draft--

Manuscript Number:	
Full Title:	Measuring what matters MOST: Validation of the Measure of Ovarian Symptoms and Treatment, a patient-reported outcome measure of symptom burden and impact of chemotherapy in recurrent ovarian cancer (ROC)
Article Type:	Special Section : Measuring What Matters (by invitation only)
Keywords:	ovarian cancer; recurrent ovarian cancer; platinum sensitive; platinum resistant; platinum refractory; symptom burden; symptom benefit; Magnitude of Clinical Benefit; Net Health Benefit; patient-reported outcome; PRO; patient-reported outcome measure; PROM.
Corresponding Author:	Madeleine T King, BSc(Hons), PhD University Of Sydney Sydney, NSW AUSTRALIA
Corresponding Author Secondary Information:	
Corresponding Author's Institution:	University Of Sydney
Corresponding Author's Secondary Institution:	
First Author:	Madeleine T King
First Author Secondary Information:	
Order of Authors:	Madeleine T King
	Martin Stockler
	Rachel O'Connell
	Luke Buizen
	Florence Joly
	Anne Lanceley
	Felix Hilpert
	Aikou Okamoto
	Eriko Aotani
	Jane Bryce
	Paul Donnellan
	Amit Oza
	Elisabeth Avall-Lundqvist
	Jonathan Berek
	Amanda Feeney
	Dominique Berton-Rigaud
	Jalid Sehouli
	Michael Friedlander
Order of Authors Secondary Information:	

Powered by Editorial Manager® and ProduXion Manager® from Aries Systems Corporation

Funding Information:	National Health and Medical Research Council (1063012)	Prof. Michael Friedlander
	National Health and Medical Research Council (570893)	Prof. Michael Friedlander
	National Institute for Health Research (C444/A15953)	Dr. Anne Lanceley

Abstract:

Purpose: Gynecologic Cancer Intergroup Symptom Benefit Study (GCIG-SBS) Stage 2 aimed to review, revise then validate the Measure of Ovarian Symptoms and Treatment concerns (MOST), developed in GCIG-SBS Stage 1 (MOSTv1, 35 items), and comprehensively document ROC symptom burden and benefit.

Methods: GCIG-SBS Stage 2 recruited patients with platinum resistant/refractory recurrent ovarian cancer (PRR-ROC) or potentially platinum sensitive ROC with ≥3 lines of prior chemotherapy (PPS-ROC≥3). Patients completed MOSTv1, QLQ-OV28 and FACT-O/FOSI at baseline and before cycle 3 of chemotherapy (pre-C3), and global assessments of change (MOST-Change) pre-C3. Clinicians rated patients' cancer-related symptoms, performance status and adverse events. Internal consistency (Cronbach's alpha), convergent and divergent validity (Spearman correlations), discriminative validity (effect sizes between groups classified by clinician-rated characteristics) and responsiveness (paired t-tests in patients expected to experience clinically meaningful change) were assessed.

Results: Of 948 recruits, 903 completed PROMs at baseline and 685 pre-C3. Baseline symptom burden was substantial, with few differences between PRR-ROC and PPS-ROC≥3. Eleven MOSTv1 items were excluded. MOSTv2 has 24 items and five multi-item scales: abdominal symptoms (MOST-Abdo), disease or treatment-related symptoms (MOST-DorT), chemotherapy-related symptoms (MOST-Chemo), psychological symptoms (MOST-Psych), and MOST-Wellbeing. Cronbach's alpha were >0.80. Correlations confirmed concurrent and divergent validity. Discriminative validity was confirmed by effect sizes that conformed with a priori hypotheses. MOST-Abdo was responsive to improvements in abdominal symptoms and MOST-Chemo detected the adverse effects of chemotherapy.

Conclusions: The MOSTv2 validly quantifies patient-reported symptom burden, adverse effects, and symptom benefit in ROC, making fit-for-purpose for clinical trials of palliative chemotherapy in ROC.

Suggested Reviewers:

Richard T Penson, MD, MRCP
Clinical Director, Medical Gynecologic Oncology, Massachusetts General Hospital
rpenson@partners.org
Dr Penson is a medical oncologist with interests in gynecologic oncology and psychosocial oncology, who treats patients with ovarian cancer. Clinician researcher involved in Gynecologic Oncology Group (GOG) studies, with a particular interest and involvement in HRQOL endpoints and analysis.

Ethan Basch, MD, MSC
Director, Cancer Outcomes Research Program, University of North Carolina Lineberger Comprehensive Cancer Center
ebasch@med.unc.edu
Dr. Basch is a medical oncologist and health services researcher. His clinical expertise is prostate cancer, and his research expertise includes patient-reported outcomes, drug regulatory policy, and comparative effectiveness research.

Powered by Editorial Manager® and ProduXion Manager® from Aries Systems Corporation

Measuring what matters MOST: validation of the Measure of Ovarian Symptoms and Treatment, a patient-reported outcome measure of symptom burden and impact of chemotherapy in recurrent ovarian cancer

Madeleine T King , Martin R Stockler, Rachel L. O'Connell, Luke Buizen, Florence Joly, Anne Lanceley, Felix Hilpert, Aikou Okamoto, Eriko Aotani, Jane Bryce, Paul Donnellan, Amit Oza, Elisabeth Avall-Lundqvist, Jonathan Berek, Jalid Sehouli, Mandy Feeney, Dominique Berton-Rigaud, and Michael Friedlander; for the GCIG Symptom Benefit group

Madeleine King, University of Sydney, Sydney, NSW, Australia; Australia New Zealand Gynaecological Oncology Group (ANZGOG)

Martin Stockler, National Health and Medical Research Council (NHMRC) Clinical Trials Centre, University of Sydney, Sydney, NSW, Australia; Australia New Zealand Gynaecological Oncology Group (ANZGOG)

Rachel O'Connell, National Health and Medical Research Council (NHMRC) Clinical Trials Centre, University of Sydney, Sydney, Australia

Luke Buizen, National Health and Medical Research Council (NHMRC) Clinical Trials Centre, University of Sydney, Sydney, Australia

Florence Joly, Centre Francois Baclesse, Caen, France; Group d'Investigateurs Nationaux pour l'Etude des Cancers Ovariens (GINECO)

Anne Lanceley, University College London, London, United Kingdom; National Cancer Research Institute (NCRI); Medical Research Council Clinical Trials Unit (MRC CTU)

Felix Hilpert, Onkologisches Therapiezentrum am Krankenhaus Jerusalem Hamburg, Hamburg, Germany; Arbeitsgemeinschaft Gynäkologische Onkologie Studiengruppe (AGO Study Group)

Aikou Okamoto, Jikei University School of Medicine, Tokyo, Japan; Japanese Gynecologic Oncology Group (JGOG)

Eriko Aotani, Global Health Research Coordinating Center, Kanagawa Academy of Science and Technology, Kanagawa, Japan; Gynecologic Oncology Trial and Investigation Consortium (GOTIC)

Jane Bryce, National Cancer Institute, Naples, Italy; Multicenter Italian Trials in Ovarian cancer and gynecologic malignancies group (MITO)

Paul Donnellan, Cancer Trials Ireland, Galway University Hospital, Galway, Ireland; All Ireland Cooperative Oncology Research Group (ICORG)

Amit Oza, Princess Margaret Hospital, Toronto, Canada; Princess Margaret Consortium (PMHC)

Elisabeth Avall-Lundqvist, Linkoping University, NSGO, Linkoping, Sweden; Nordic Society of Gynaecological Oncology (NSGO)

Jonathan Berek, Stanford Comprehensive Cancer Institute, Stanford, CA, United States; Cooperative Ovarian Cancer Group (COGi)

Amanda Feeney, University College London, London, United Kingdom; National Cancer Research Institute (NCRI); Medical Research Council Clinical Trials Unit (MRC CTU)

Dominique Berton-Rigaud, Institut de Cancerologie de l'Ouest (ICO), Centre René Gauducheau, Saint Herblain, France; Group d'Investigateurs Nationaux pour l'Etude des Cancers Ovariens (GINECO)

Jalid Sehouli, Department of Gynecology and Oncological Surgery, Charité, University of Berlin, Germany; Arbeitsgemeinschaft Gynäkologische Onkologie Studiengruppe (AGO Study Group)

Michael Friedlander, Prince of Wales Hospital, Sydney, Australia; Australia New Zealand Gynaecological Oncology Group (ANZGOG)

On behalf of the GCIG Symptom Benefit Study Group, Gynecologic Cancer Intergroup (GCIG), Kingston, Ontario, Canada.

Corresponding author:

Madeleine King

Quality of Life Office, Psycho-oncology Co-operative Research Group

Level 6 North, Chris O'Brien Lifehouse C39Z

University of Sydney NSW 2006 Australia

E-mail: madeleine.king@sydney.edu.au

P: +612 434 164 438

Abstract Click here to download Abstract ABSTRACT-FINAL.docx

Measuring what matters MOST: validation of the Measure of Ovarian Symptoms and Treatment, a patient-reported outcome measure of symptom burden and impact of chemotherapy in recurrent ovarian cancer

ABSTRACT (250 of 250 word limit)

Purpose: Gynecologic Cancer Intergroup Symptom Benefit Study (GCIG-SBS) Stage 2 aimed to review, revise then validate the Measure of Ovarian Symptoms and Treatment concerns (MOST), developed in GCIG-SBS Stage 1 (MOSTv1, 35 items), and comprehensively document ROC symptom burden and benefit.

Methods: GCIG-SBS Stage 2 recruited patients with platinum resistant/refractory recurrent ovarian cancer (PRR-ROC) or potentially platinum sensitive ROC with ≥3 lines of prior chemotherapy (PPS-ROC≥3). Patients completed MOSTv1, QLQ-OV28 and FACT-O/FOSI at baseline and before cycle 3 of chemotherapy (pre-C3), and global assessments of change (MOST-Change) pre-C3. Clinicians rated patients' cancer-related symptoms, performance status and adverse events. Internal consistency (Cronbach's alpha), convergent and divergent validity (Spearman correlations), discriminative validity (effect sizes between groups classified by clinician-rated characteristics) and responsiveness (paired t-tests in patients expected to experience clinically meaningful change) were assessed.

Results: Of 948 recruits, 903 completed PROMs at baseline and 685 pre-C3. Baseline symptom burden was substantial, with few differences between PRR-ROC and PPS-ROC≥3. Eleven MOSTv1 items were excluded. MOSTv2 has 24 items and five multi-item scales: abdominal symptoms (MOST-Abdo), disease or treatment-related symptoms (MOST-DorT), chemotherapy-related symptoms (MOST-Chemo), psychological symptoms (MOST-Psych), and MOST-Wellbeing. Cronbach's alpha were >0.80. Correlations confirmed concurrent and divergent validity. Discriminative validity was confirmed by effect sizes that conformed with a priori hypotheses. MOST-Abdo was responsive to improvements in abdominal symptoms and MOST-Chemo detected the adverse effects of chemotherapy.

Conclusions: The MOSTv2 validly quantifies patient-reported symptom burden, adverse effects, and symptom benefit in ROC, making it fit-for-purpose for clinical trials of palliative chemotherapy in ROC.

Manuscript

Click here to download Manuscript MOST-Validation-MWM-FINAL.docx

Click here to view linked References

Introduction

Recurrent ovarian cancer (ROC) – what's the problem?

The goals of treatment for recurrent ovarian cancer (ROC) are to reduce disease-related symptoms, delay disease progression, and prolong overall survival. Yet in clinical trials, the main metrics used to measure the benefit of palliative chemotherapy and for regulatory approval of drugs are overall survival and progression-free survival are. Trials rarely assess or document the overall symptom burden, the proportion of patients with specific cancer-related symptoms, or the extent to which symptoms improve with palliative chemotherapy. The burden of cancer-related symptoms, and their relief, should be key considerations for patients, clinicians, regulators, and policy makers.

Women with ROC are a heterogeneous group with highly variable time to progression. Disease relapsing later than 6 months after first-line platinum-based chemotherapy is classified platinum sensitive, tumour response rates are 25% to 67%, and median survival 2 to 3 years [1]. Disease progressing earlier is classified platinum resistant, with tumour response rates of 10% to 15%, and median survival 9 to 12 months [2]. Disease progressing during chemotherapy is classified as platinum refractory, and has response rates of 10% or less and median survival <6 months with chemotherapy [2]. The benefits of chemotherapy diminish with each successive line [3]: in a retrospective analysis of over 1600 patients, the median progression-free survival after first, second, and third treated recurrences was 10, 6, and 4 months, respectively [3].

The 3rd Ovarian Cancer Consensus Conference recognized the importance of symptom benefit as an endpoint for clinical trials, and recommended development and validation of an instrument to measure this in women with ROC, particularly those with platinum resistant-

refractory ovarian cancer (PRROC) [4]. This led to the symptom benefit working group of the Gynecologic Cancer Intergroup (GCIG) and, subsequently, the two-stage Symptom Benefit Study (GCIG-SBS). More recently, the European Society of Medical Oncology (ESMO) and American Society of Clinical Oncology (ASCO) have proposed standardized measures of the magnitude of clinical benefit [5] and the net health benefit [6] in clinical trials that incorporate effects on survival time, toxicity, and health-related quality of life to determine whether patients are living better and/or longer as a result of their anticancer treatment. The aims and objectives of GCIG-SBS were closely aligned with these recommendations.

GCIG-SBS Stage 1: developing MOST version 1 (MOSTv1)

The primary aim of GCIC-SBS Stage 1 was to document the most prevalent symptoms in women with PRROC (expected to result in the greatest symptom burden), and to seek an optimal instrument to measure subjective symptom benefit in clinical trials of palliative chemotherapy [7, 8]. We defined “optimal” as best able to provide efficient and focused measurement of symptom benefit as an endpoint in clinical trials, then operationalized this with four criteria derived from key elements of the Food and Drug Administration’s (FDA) recommended process for developing a patient-related outcome measure (PROM): content validity, recall period, item rating scale, and scoring [9]. We used mixed methods to identify the symptoms that were most noticed and most severe in the Stage 1 sample of 129 women [8], then assessed the extent to which these were covered by existing candidate PROMs. While the EORTC QLQ-C30 [10] and QLQ-OV28 [11] together cover all of these symptoms, their scoring algorithms split them into numerous scales, potentially dissipating effects, and requiring multiple testing when analyzing them for clinical trials. The FACT-O [12] covered all but two of the most prevalent cancer-related symptoms, but combined these with chemotherapy side-effects and other aspects of health-related quality of life, leading to potential dilution in multi-item summary scales. This was partly

addressed by the development of two FACT Ovarian Symptom Indexes (FOSI) [13-15], but even these are somewhat diluted by other issues, including preexisting adverse effects of chemotherapy and existential concerns that are unlikely to improve with chemotherapy. Further, the recall period for those PROMs is a week, while the period between chemotherapy cycles is typically 3-4 weeks. We therefore undertook to develop a new PROM to quantify the overall symptom burden, the benefit of chemotherapy, and the adverse effects of chemotherapy: the Measure of Ovarian Symptoms and Treatment concerns (MOST).

Steps in developing the MOST

In developing and validating the MOST, we followed the iterative process recommended by the FDA for developing a PROM [9], which comprises five steps. Stage 1 of the GCIG-SBS addressed the first two of these steps by developing a conceptual framework, and then drafting the MOSTv1 with patient input, piloting the MOSTv1, and determining its content validity (step 2) [8]. In this paper, we complete step 2 by assessing adjustments to our conceptual framework, and address step 3 using data collected in Stage 2 of the GCIG-SBS.

GCIG-SBS Stage 2: Aims

The main aims of GCIG-SBS Stage 2 were: 1. to confirm the MOST's final conceptual framework with scoring rules; 2. to assess the reliability of component scores; 3. to assess the ability of component scores to detect changes; and, 4. to comprehensively document the symptoms most prevalent in ROC at baseline.

Methods

GCIG-SBS Stage 2 study design, participants and data collection

This prospective, observational cohort study recruited women with PRROC and women with potentially platinum-sensitive ovarian cancer who had had three or more lines of chemotherapy

(PPS≥3) from 120 sites in 11 countries. Eligibility criteria were: ROC and; progression based on CA125 level, imaging, or clinical characteristics; Eastern Cooperative Oncology Group performance status 0 to 3; life expectancy >3 months; and ability to complete questionnaires independently. Chemotherapy and supportive care were at the discretion of treating physicians.

Treating physicians recorded their patients' baseline characteristics and the presence of ascites, cramping abdominal pain or intermittent or incomplete bowel obstruction. They assessed 'clinical benefit' every 6–8 weeks (2–3 cycles of chemotherapy) according to CA125, RECIST, and/or symptomatic improvement. PROMs included the MOSTv1 [8], EORTC QLQ-C30 [10], QLQ-OV28 [11], FACT-O [12], and FOSI [13-15]. Participants completed these PROMs at baseline (before their first cycle of chemotherapy), then every 3 to 4 weeks before chemotherapy, until disease progression. Participants completed the MOST-Change questionnaire 3–4 weeks after their second cycle of chemotherapy, immediately before cycle 3. This time point was chosen because any improvement in cancer-related symptoms due to chemotherapy (symptom benefit) was judged likely to be evident by then, and many patients stop chemotherapy after two cycles because of progression [16]. This corresponded with the first assessment of clinical benefit by the treating physician. The MOST-Change questionnaire asks participants for each aspect included in the MOSTv1, 'how things are now, compared with how things were before you started this course of chemotherapy 6–8 weeks ago, rated on a 5-point scale: much better, a little better, the same, a little worse, or much worse.

The MOSTv1 and MOST-Change questionnaires were provided in English, German, French, Italian, Japanese, Swedish after translation from English, based on the EORTC guidelines [17].

PROM development steps 2 and 3: review conceptual framework, content and scoring rules

Our initial conceptual framework (GCIG-SBS Stage 1, PROM development step 1) envisaged two indexes: one assessing disease-related symptoms, the other treatment-related concerns, so that the benefits and burdens of chemotherapy could be quantified separately. On reflection, this was an oversimplification, and we revised our conceptual framework into four groups of symptoms plus a global wellbeing group:

- Physical symptoms most likely due to ROC and not chemotherapy;
- Physical symptoms most likely due to chemotherapy and not ROC;
- Physical symptoms reasonably likely due to ROC and/or chemotherapy;
- Psychological symptoms;
- Global wellbeing (physical, emotional and overall).

Our approach remained predominantly clinimetric rather than psychometric [18], i.e. our goal was indexes to include key symptoms according to their likely causes, regardless of their co-occurrence and/or correlations.

Our guiding principle was to *measure what matters and no more*. We therefore excluded items that:

- had low prevalence at baseline and pre-cycle 3 and did not change markedly over time;
- were conceptually and/or functionally similar, and highly correlated (≥0.50) with an already included item;

- were symptoms of neither ROC nor chemotherapy (other existing PROMs in the EORTC and FACIT suites address these aspect of health-related quality of life).

As an evidence base for our PROM development considerations, and to address our objective of comprehensively documenting the symptoms most prevalent in ROC at baseline, we summarized the prevalence of each MOSTv1 item at baseline and pre-C3 with its mean, standard deviation, and proportions based on the MOST response format, a numeric rating scale with integers from zero to 10, with five verbal anchors: 'No trouble at all' (over 0), 'Mild' (over 1–3), 'Moderate' (over 4–6), 'Severe' (over 7–10), and 'Worst I can imagine' (over 10). We used chi-squared tests to assess differences in baseline symptoms between PRROC and PPS≥3. Associations between items were assessed at baseline and pre-C3 using Spearman's rank correlations (r_s).

A key requirement for using the MOST in clinical trials is to have scoring rules yielding PROM scales that are sensitive to both improvement (symptom benefit) and deterioration (burden of treatment and/or worsening of disease). To achieve this, we sought items that changed in the same direction (either all improved or all worsened) and assessed mean changes from baseline to pre-C3. On the basis of these considerations, multi-item symptom scales were scored by taking the average of the component items, with linear rescaling to an observable range of 0–100, with higher scores representing worse symptoms or better wellbeing.

Psychometric properties

Reliability: Given our clinimetric approach, we did expect or require high internal consistency between items within each multi-item scale [19], but nevertheless assessed it with Cronbach's alpha.

Construct validity: We assessed evidence that relationships among items, multi-item scales and concepts conformed to *a priori* hypotheses concerning logical relationships that should exist with other measures or patient characteristics, as follows.

Convergent and divergent validity: we hypothesized that the MOST multi-item scales should be strongly correlated (r=0.50 or higher, Cohen's threshold for strong correlation [20]) with related scales from the EORTC QLQ-C30/OV28, FACT-O and FOSI, and more strongly correlated with these scales than with other scales of the MOST and other PROMS that measured less related constructs [21].

Discriminative validity (difference between 'known groups'): We hypothesized the MOST would discriminate between clinically distinct groups, as specified in the following five hypotheses (H1–H5).

H1. Baseline MOST scores for physical symptoms of ROC would be worse for participants rated by clinicians at baseline as having cancer-related symptoms, ascites, abdominal pain/obstructive symptoms, or poor performance status (ECOG PS 2–3 versus 0–1), than for those not having these characteristics.

H2. Patterns for baseline wellbeing would be similar to those for H1, but with smaller gradients than for self-reported physical symptoms, because wellbeing is less directly related to these clinical characteristics.

H3. Psychological symptoms were not expected to differ according to clinicians' ratings of ascites, abdominal pain/obstructive symptoms, or performance status, as observed previously for the QLQ-C30 [22, 23].

H4. At pre-C3, patient self-report of specific symptoms, on average, would increase in a gradient across CTC-AE grades for analogous National Cancer Institute Common Terminology for Adverse Events (CTC-AE v4) symptoms.

H5. At pre-C3, MOST multi-item scales rating symptoms that may be caused by chemotherapy would be worse (higher) for participants with CTC-AE grades higher than the median.

For MOST multi-item scales, we assessed H1-H5 with mean differences, independent *t* tests and Cohen's D effect sizes, defined as the mean difference divided by its standard deviation [20], and 95% confidence intervals (CI). All scores were transformed using the square root function to achieve normality for the *t* test and calculation of Cohen's D. For MOSTv1 items that had a corresponding CTC-AE item, the Kruskall-Wallis test was used to test for differences in MOST ranks across CTC-AE grades, and gradients were assessed visually in boxplots.

Longitudinal validity (Responsiveness to expected change): We hypothesized that changes in MOST multi-item scale scores from baseline to pre-C3 would follow a gradient from most improved through to most worsened when grouped by patients' global ratings of change on conceptually-related MOST-Change items with correlations ≥0.30 [24]. We assessed this statistically with analysis of variance (ANOVA) and visually with boxplots.

We hypothesized that a MOST multi-item scale composed of physical symptoms most likely due to ROC and not chemotherapy would register an improvement with chemotherapy in women whose treating clinicians recorded baseline cramping abdominal pain or intermittent/incomplete bowel obstruction, and that this improvement would be most apparent in the subgroup who rated their overall wellbeing as much better on the MOST-Change

questionnaire at pre-C3. We hypothesized similar but larger effects when the patient-report of abdominal symptoms was used to define the group baseline burden and pre-C3 benefit.

We hypothesized that a MOST scale comprising physical symptoms most likely due to chemotherapy and not ROC would register the known toxic impact of chemotherapy, i.e. an increase from baseline to pre-C3, more so for PPS≥3 than for PRROC, as the former received combination chemotherapy while the latter received single-agent chemotherapy.

Statistical analyses were conducted in SAS and STATA.

Results

Stage 2 of GCIG-SBS recruited 948 participants including 903 (96%) who completed the MOSTv1 at baseline (Table 1) and 685 (72%) who completed MOSTv1 and MOST-Change pre-C3. Missing data rates for MOST items were low (0.8–3.5% at baseline, 0.3–5.1% at pre-C3).

Symptom burden at baseline

Baseline symptom burden was substantial (Figure 1), with little difference between patients with PRROC versus PPS≥3 ($p>0.05$ for all items except vomiting ($p=0.03$, 21% PRROC mild or moderate vs 13% PPS≥3) and hair loss ($p=0.003$, 22% PRROC moderate or severe vs 12% PPS≥3)). Most participants (86%) reported one or more symptoms as moderate or severe (MOST item score of 4–10), 50% reported 6 or more symptoms moderate or severe, and 27% reported 9 or more symptoms as moderate or severe. The most common and severe symptoms were abdominal, fatigue, anorexia and anxiety. The items for 'abdominal swelling, bloating and/or fullness' (MOSTv1 item 4) and for 'abdominal pain, discomfort, and/or cramps' (MOSTv1 item 5) often co-occurred: either or both were reported as severe by 28% of participants, moderate by 26%, mild by 31%, and no trouble at all by 15%.

Impact of chemotherapy

Among the 685 women who provided ratings at baseline and pre-C3, the symptoms that improved most (P<0.001) were: abdominal swelling, bloating, and/or fullness; abdominal pain, discomfort and/or cramps; and anxiety (Figure 2). Each improved on average by about half a point on the 0–10 point MOST item response scale. Of the seven symptoms that worsened most (P<0.001), four were due to chemotherapy (hair loss, altered sense of taste, sore mouth or throat, skin rash) and three to cancer or chemotherapy (fatigue, nausea, and difficulty swallowing).

Revision of conceptual framework, content and scoring rules: from MOSTv1 to MOSTv2

In our revisions of the MOSTv1 conceptual framework, content and scoring rules, 11 items were excluded (3 deemed redundant because they were conceptually and empirically related to retained items and 7 judged to be consequences rather than symptoms of ROC) (Table 2). 'Pain (all/anywhere)' and 'trouble concentrating' were judged too general, and could be due to other causes.

For MOSTv2, we proposed scoring rules that aggregated the 24 items retained into five multi-item scales: abdominal symptoms (MOST-Abdo), disease or treatment-related symptoms (MOST-DorT), chemotherapy-related symptoms (MOST-Chemo), psychological symptoms (MOST-Psych), and wellbeing (MOST-Wellbeing). Nine items that could be caused by either disease or treatment worsened on average from baseline to pre-C3, while two improved ('trouble sleeping', 'bladder problems'). We assessed 9 and 11 item versions (MOST-DorT-9, MOST-DorT-11), as it was unclear how 'trouble sleeping' and 'bladder problems' would affect the MOST-DorT scale's performance.

Psychometric properties

Internal consistency was excellent for all MOSTv2 multi-item scales (Cronbach's alpha ≥ 0.80) except MOST-DorT-11 (alpha<0.70) at baseline (Table 3) and pre-C3 (Table A, online resource).

Convergent and divergent validity were confirmed at baseline (Table 3) and pre-C3 (Table A, online resource). Each MOSTv2 multi-item scale substantially correlated with similar scales from other PROMs (r_s always >0.50, often ≥0.70), and typically this correlation was the highest of all correlations for that MOST scale. MOST-Abdo and MOST-DorT were always correlated >0.60 with FOSI-8 and FOSI-15, while MOST-Chemo correlated <0.50, and the correlation for MOST-Psych increased from about 0.5 at baseline to ≥0.60 pre-C3. Correlations among MOST scales were highest between MOST-Abdo and MOST-DorT (about 0.7), lowest with MOST-Chemo (0.25-0.52) and MOST-Psych (0.25-0.56), and generally increased from baseline to pre-C3.

Discriminative validity was confirmed (Table 4). Effect sizes for MOST-Abdo were larger for participants classified by clinician-rated symptoms than by clinician-rated performance status, and vice versa for MOST-Wellbeing. Effect sizes for MOST-DorT were smallest for clinician-rated abdominal pain/bowel obstruction (reflecting least similarity of content) and CTC-AE at baseline (reflecting low toxicity burden at baseline), with larger effect sizes pre-C3, reflecting the increase in toxicity burden for the MOST-DorT symptoms. Expected gradients were observed in the distributions of MOST items scores across CTC-AE grades for the 11 symptoms with comparable CTC-AE items (Figures A and B, online resource).

Longitudinal validity: Figure 3 shows a clear gradient for MOST-Abdo (grouped by the average of the two abdominal symptom items on the MOST-Change form, *r*=0.41), MOST-DorT-9 and -11 (grouped by the physical wellbeing change item, item *r*=0.39, 0.40), and MOST-Wellbeing (grouped by the overall wellbeing change item, item *r*=-0.36). Correlations for MOST-Psych and

the emotional wellbeing change item (r= 0.25) and MOST-Chemo and the physical wellbeing change item (r= 0.22) were less than the required 0.30 threshold. The MOST-Change scores at pre-C3 explained a significant proportion of the change from baseline to pre-C3 in all MOST multi-item scores (all P<0.001, ANOVA).

Responsiveness: The MOST-Abdo scale was responsive to expected improvements with chemotherapy among patients recorded to have symptoms of cramping abdominal pain or intermittent/incomplete bowel obstruction at baseline (Table 5), and very responsive in subgroups expected to show improvements (Table 6), more so for the patient-based than the clinician-based definition. The MOST-Chemo scale was responsive to expected deteriorations due to chemotherapy toxicity, registering the expected changes from baseline to pre-C3, more so for women with PPS≥3 (predominantly treated with combination chemotherapy) than with PRROC (predominantly treated with single-agent chemotherapy) (Table 5).

Results for the other MOST multi-item scales presented in Table 5 demonstrate their clinical utility in describing the symptom benefit of chemotherapy and toxicity burden, for participants who were symptomatic versus asymptomatic, and with PRROC versus PPS≥3.

Discussion

These results from a large, international initiative focusing on symptom burden, treatment benefit, and treatment toxicity in ROC complete steps 2 and 3 of the iterative process recommended by the FDA for developing PROMs, by adjusting the conceptual framework of the MOST, confirming it with scoring rules, and assessing the reliability, validity and responsiveness of resultant multi-item scales. The MOSTv2 contains 24 items and yields five multi-item scales: MOST-Abdo, MOST-DorT, MOST-Chemo, MOST-Psych, MOST-Wellbeing. All scales have excellent internal consistency, except the MOST-DorT 11-item version. Convergent and

divergent validity were confirmed for all scales. Discriminant validity was confirmed for all scales except the MOST-Psych, for which we lacked a suitable clinical anchor. The MOST's ability to detect clinically important changes was demonstrated. The responsiveness of the MOST-Abdo to improvements in the abdominal symptoms of ROC makes it a good candidate as a key endpoint in trials of palliative chemotherapy for ROC, as these are the defining and predominant symptoms of ROC. The MOST-Chemo detected the adverse effects of chemotherapy and specific toxicities.

Some symptoms of ROC may be caused by both the cancer and its treatment - this is the most challenging aspect of the impact of chemotherapy to measure and interpret. The MOST-DorT increased (worsened) with chemotherapy in asymptomatic participants, and we surmise that this was predominantly because of the adverse effects of chemotherapy. Similarly, we surmise that the larger increase in MOST-DorT for PPS≥3 versus PRROC was also because of the greater adverse effects of multi-agent chemotherapy compared with a single agent. We have reported two versions of the MOST-DorT—with and without 'trouble sleeping' and 'bladder problems'—as these were the only two symptoms due to disease or treatment that improved with chemotherapy. While this may be a sample-specific finding, reduction of abdominal bloating should reduce bladder symptoms, and this along with reduced anxiety and pain could improve sleeping. We did not show substantial differences in the performance of these two versions, except that the 11-item version lacked internal consistency. Further comparisons in other datasets are required to determine whether these two items should be reported as single items or combined with the other nine symptoms of disease or treatment.

We have also addressed the secondary aim of Stage 2 of the GCIG-SBS, by comprehensively documenting the symptom burden of ROC before chemotherapy, and further, have used MOST to estimate the impact of chemotherapy in terms of benefits and harms in this cohort.

GCIG-SBS Stage 2 had some limitations for comprehensive validation of MOSTv2. Test-retest reliability was not built into the data collection design of the study; this needs to be addressed in future studies. We lacked suitable anchors for assessing the MOST-Psych's discriminative validity and responsiveness; future studies that target or identify clinical cases of anxiety and depression are required. The use of clinician-rated CTC-AEs was somewhat limited for MOST-Chemo validation because of differences between the patient-rated and clinician-rated items in the specification of symptoms [25].

Although thousands of women with ROC have participated in clinical trials of palliative chemotherapy, whether chemotherapy palliates symptoms is so far unsubstantiated. The development of MOST was driven by the imperative for a measure of symptom benefit, in direct response to the third consensus conference [4]. The GCIG-SBS illustrates the theory and methods of developing and validating a PROM designed to measure what matters most from the perspective of patients with ROC. We followed the FDA's iterative process for developing and validating a fit-for-purpose PROM in over 1000 women with ROC who participated in GCIG-SBS, Stages 1 and 2. MOST is now ready for inclusion in clinical trials of palliative chemotherapy in ROC, and will meet the need for symptom benefit to be included as a trial endpoint, as identified by the fourth consensus conference [26].

About 20% of study participants were reported asymptomatic at baseline, with clinicians stating the aim of chemotherapy was to delay the development of symptoms. It is unclear how often this is achieved, because the delay in development of cancer-related symptoms has not yet

been measured in randomized trials. For asymptomatic patients, symptom benefit is not relevant and yet treatment toxicity is likely; chemotherapy cannot make these patients feel better but it will probably make them feel worse. This clinically important postulation should be addressed in future clinical trials, and could be assessed with MOST.

This study clearly demonstrates the substantial baseline symptom burden reported by patients with ROC: 86% reported at least one of the MOST symptoms as moderate or severe, 50% rated 6 or more symptoms moderate or severe, and 27% rated 9 or more symptoms moderate or severe. These findings underscore the importance of documenting the severity of symptoms at baseline and the effects of treatment on these symptoms. The predominant and disease-defining symptoms of ROC were 'abdominal pain, discomfort and/or cramps'; and 'abdominal swelling, bloating and/or fullness', the symptoms most likely to be improved by chemotherapy. These observations suggest that the MOST-Abdo is a suitable endpoint for clinical trials of palliative chemotherapy in ROC.

The MOST is a flexible instrument that can be adapted and modified, depending on the aims and objectives of the clinical trial. Symptoms within the MOST-Chemo and MOST-DorT—attributable to chemotherapy, or to disease and/or treatment, respectively—may be analyzed separately item by item, or together in overall symptom indexes. The preferred options will depend on the target population, treatments, and patient-reported outcome hypotheses specified a priori.

We developed the MOST as a PROM fit for the purpose of assessing the benefit and harms of palliative chemotherapy in ROC. We have established key aspects of the MOSTv2's reliability, validity, and responsiveness, and identified evidence gaps for future research. The 24-item MOSTv2 is designed to take less than 5 minutes to complete, and to provide a comprehensive assessment of the burden of symptoms and effects of palliative chemotherapy,

both beneficial and adverse, in ROC. MOST focuses exclusively on symptoms and aspects of wellbeing; researchers wishing to assess health-related quality of life should use MOST together with the EORTC and/or FACIT measures. In conclusion, we are now positioned to meet the ASCO and ESMO directive to “measure what matters” in trials of palliative chemotherapy.

Compliance with ethical standards

Conflicts of interest

[to come, if required by QOLR]

Ethical standard

GCIG-SBS was led and coordinated by the Australian New Zealand Gynecological Oncology Group (ANZGOG) and National Health and Medical Research Council (NHMRC) Clinical Trials Centre, University of Sydney, in collaboration with the GCIG Symptom Benefit Committee. The trial was registered on the Australian New Zealand Clinical Trials Registry (ANZCTR 12607000603415). The study was performed in accordance with the NHMRC Statement on Ethical Conduct in Research Involving Humans and the Declaration of Helsinki, with ethics approval at all participating sites, and signed, written, informed consent from all participants.

Funding

In Australia the study was supported by NHMRC grants 1063012 and 570893. In the United Kingdom (UK), this was a National Institute for Health Research (NIHR) study jointly funded by Target Ovarian Cancer and the Cancer Research UK and UCL Cancer Trials Centre (Programme Grant C444/A15953). Professor King is supported by the Australian Government through Cancer Australia. Dr Anne Lanceley was assisted by the National Institute for Health Research (NIHR) UCLH/UCL Biomedical Research Centre which is supported by the Department of Health.

Acknowledgments

In Australia, the study was coordinated by the NHMRC Clinical Trials Centre, University of Sydney.The Cancer Research UK and UCL Cancer Trials Centre coordinated UK participation in the study.

References

1. Friedlander, M., et al., *Clinical trials in recurrent ovarian cancer.* International Journal of Gynecological Cancer, 2011. **21**(4): p. 771-5.
2. Davis, A., A.V. Tinker, and M. Friedlander, *"Platinum resistant" ovarian cancer: What is it, who to treat and how to measure benefit?* Gynecologic Oncology, 2014. **133**(3): p. 624-631.
3. Hanker, L.C., et al., *The impact of second to sixth line therapy on survival of relapsed ovarian cancer after primary taxane/platinum-based therapy.* Annals of Oncology, 2012. **23**(10): p. 2605-12.
4. du Bois, A., et al., *2004 consensus statements on the management of ovarian cancer: final document of the 3rd International Gynecologic Cancer Intergroup Ovarian Cancer Consensus Conference (GCIG OCCC 2004).* Annals of Oncology, 2005. **16**: p. 7-12.
5. Cherny, N.I., et al., *A standardised, generic, validated approach to stratify the magnitude of clinical benefit that can be anticipated from anti-cancer therapies: the European Society for Medical Oncology Magnitude of Clinical Benefit Scale (ESMO-MCBS).* Annals of Oncology. **26**(8): p. 1547-73.
6. Schnipper, L.E., et al., *American Society of Clinical Oncology Statement: A Conceptual Framework to Assess the Value of Cancer Treatment Options.* Journal of Clinical Oncology. **33**(23): p. 2563-77.
7. Friedlander, M.L., et al., *Symptom burden and outcomes of patients with platinum resistant/refractory recurrent ovarian cancer: a reality check: results of stage 1 of the gynecologic cancer intergroup symptom benefit study.* International Journal of Gynecological Cancer, 2014. **24**(5): p. 857-64.
8. King, M.T., et al., *Development of the measure of ovarian symptoms and treatment concerns: aiming for optimal measurement of patient-reported symptom benefit with chemotherapy for symptomatic ovarian cancer.* International Journal of Gynecological Cancer, 2014. **24**(5): p. 865-73.
9. FDA, *Food and Drug Administration. Guidance for Industry on Patient-Reported Outcome Measures: Use in Medical Product Development to Support Labeling Claims.* Federal Register, 2009. **74**(235): p. 65132-65133.
10. Aaronson, N.K., et al., *The European Organization for Research and Treatment of Cancer QLQ-C30: a quality-of-life instrument for use in international clinical trials in oncology.* J Natl Cancer Inst, 1993. **85**(5): p. 365-76.
11. Greimel, E., et al., *An international field study of the reliability and validity of a disease-specific questionnaire module (the QLQ-OV28) in assessing the quality of life of patients with ovarian cancer.* Eur J Cancer, 2003. **39**(10): p. 1402-8.
12. Basen-Engquist, K., et al., *Reliability and validity of the functional assessment of cancer therapy-ovarian.* J Clin Oncol, 2001. **19**(6): p. 1809-17.
13. Beaumont, J., et al., *Validation of the Functional Assessment of Cancer Therapy-Ovarian (FACT-O) Symptom Index (FOSI) in a phase II clinical trial of pertuzumab in patients with*

advanced ovarian cancer. . 2007 ASCO Annual Meeting Proceedings Part I. Vol 25, No. 18S (June 20 Supplement), 2007. **25**: p. 18S.

14. Cella, D., et al., *What are the most important symptom targets when treating advanced cancer? A survey of providers in the National Comprehensive Cancer Network (NCCN).* Cancer Investigation, 2003. **21**(4): p. 526-35.
15. Jensen, S.E., et al., *A new index of priority symptoms in advanced ovarian cancer.* Gynecologic Oncology, 2011. **120**(2): p. 214-9.
16. Sehouli, J., et al., *Topotecan Weekly Versus Conventional 5-Day Schedule in Patients With Platinum-Resistant Ovarian Cancer: a randomized multicenter phase II trial of the North-Eastern German Society of Gynecological Oncology Ovarian Cancer Study Group.* Journal of Clinical Oncology, 2011. **29**(2): p. 242-8.
17. Kulis, D., et al., *EORTC Quality of Life Group Translation Procedure*. 4th ed. 2016, Brussels: EORTC Quality of Life Group.
18. Wright, J.G. and A.R. Feinstein, *A comparative contrast of clinimetric and psychometric methods for constructing indexes and rating scales.* Journal of Clinical Epidemiology, 1992. **45**(11): p. 1201-18.
19. Streiner, D.L., *Being inconsistent about consistency: when coefficient alpha does and doesn't matter.* Journal of Personality Assessment, 2003. **80**(3): p. 217-22.
20. Cohen, J., *Statistical Power Analysis for the Behavioral Sciences*. 1988: Routledge.
21. Campbell, D.T. and D.W. Fiske, *Convergent and discriminant validation by the multitrait-multimethod matrix.* Psychological Bulletin, 1959. **56**(2): p. 81-105.
22. Cocks, K., et al., *Evidence-based guidelines for determination of sample size and interpretation of the European Organisation for the Research and Treatment of Cancer Quality of Life Questionnaire Core 30.* Journal of Clinical Oncology, 2011. **29**(1): p. 89-96.
23. King, M.T., *The interpretation of scores from the EORTC quality of life questionnaire QLQ-C30.* Quality of Life Research, 1996. **5**(6): p. 555-67.
24. Revicki, D., et al., *Recommended methods for determining responsiveness and minimally important differences for patient-reported outcomes.* Journal of Clinical Epidemiology, 2008. **61**(2): p. 102-9.
25. Basch, E., et al., *Adverse symptom event reporting by patients vs clinicians: relationships with clinical outcomes.* J Natl Cancer Inst, 2009. **101**(23): p. 1624-32.
26. Stuart, G.C., et al., *2010 Gynecologic Cancer InterGroup (GCIG) consensus statement on clinical trials in ovarian cancer: report from the Fourth Ovarian Cancer Consensus Conference.* International Journal of Gynecological Cancer, 2011. **21**(4): p. 750-5.

Table
Click here to download Table MOSTtables_21Nov2016_FINAL.docx

Table 1. Baseline characteristics of patients who completed at least one item of the MOST questionnaire at baseline (n=903)

Characteristic	*n* (%)
Age (years, mean 62.5)	
23–49	103 (11.5)
50–59	245 (27.1)
60–69	298 (33.0)
70–89	257 (28.5)
Type of resistance, PRROC	
Primary platinum-refractory	60 (6.6)
Primary platinum-resistant	223 (24.7)
Secondary platinum-refractory	88 (9.7)
Secondary platinum-resistant	169 (18.7)
Potentially platinum sensitive ≥3[a]	363 (40.2)
ECOG performance status	
0	306 (33.9)
1	498 (55.1)
2	92 (10.2)
3	7 (0.8)
Lines of previous treatment for ovarian cancer	
1	178 (19.7)
2	336 (37.2)
3	203 (22.5)
≥4 (maximum 10)	186 (11.5)
Response to most recent line	
Complete response	115 (12.8)
Progressive disease	362 (40.2)
Partial response	246 (27.3)
Stable disease	150 (16.7)
Unknown (not assessed)	27 (3.0)
Clinician rating of symptoms at baseline	
Cancer-related symptoms	650 (72.1)
Symptomatic ascites	202 (22.4)
Symptoms of cramping abdominal pain, or intermittent or incomplete bowel obstruction	369 (40.9)
Reasons for planned chemotherapy	
Symptom control or palliation	635 (70.6)
If asymptomatic, to delay the development of symptoms	278 (35.9)
Rising CA125	534 (63.0)
Radiological evidence of progression	602 (76.3)

a. Potentially platinum sensitive disease but had 3 or more prior lines of chemotherapy.

Abbreviations: MOST, Measure of Ovarian Symptoms and Treatment; PRROC, platinum resistant or refractory recurrent ovarian cancer.

Table 2. Revision of the MOST version 1 (35 items) to MOST version 2 (24 items)

MOSTv1 symptoms and problems*	Attribution[a-e]	Change[f]	MOSTv2[g]	Reason, if excluded
Abdominal pain, discomfort and/or cramps	ROC[a]	I	MOST-Abdo	
Abdominal swelling, bloating and/or fullness	ROC[a]	I	MOST-Abdo	
Fatigue (tiredness)	ROC/Chemo [b]	W	MOST-DorT	
Trouble eating	ROC/Chemo [b]	W	MOST-DorT	
Indigestion	ROC/Chemo[b]	W	MOST-DorT	
Nausea	ROC/Chemo[b]	W	MOST-DorT	
Vomiting	ROC/Chemo[b]	W	MOST-DorT	
Diarrhoea	ROC/Chemo[b]	W	MOST-DorT	
Constipation	ROC/Chemo[b]	W	MOST-DorT	
Shortness of breath	ROC/Chemo[b]	W	MOST-DorT	
Difficulty swallowing	ROC/Chemo[b]	W	MOST-DorT	
Trouble sleeping	ROC/Chemo[b]	I	DorT or Single?	
Bladder problems	ROC/Chemo[b]	I	DorT or Single?	
Pain (all/anywhere)	ROC/Chemo[b]	I	Excluded	Too general, redundant (r=0.66 abdominal pain)
Poor appetite (or feeling full quickly)	ROC/Chemo[b]	0	Excluded	Redundant (r=0.78 trouble eating, r=0.70 loss of appetite)
Loss of appetite	ROC/Chemo[b]	W	Excluded	Redundant (r=0.73 trouble eating, r=0.79 poor appetite)
Leg swelling	ROC/Chemo[b]	0	Excluded	Uncommon, didn't change
Trouble concentrating	ROC/Chemo[b]	I	Excluded	Consequence not symptom, too general, R (r=0.64 anxiety)
*Problems doing what I wanted	ROC/Chemo[b]	W	Excluded	Consequence not symptom
*Problems for my family or friends	ROC/Chemo[b]	I	Excluded	Consequence not symptom
Altered sense of taste	Chemo[c]	W	MOST-Chemo	
Sore mouth or throat	Chemo[c]	W	MOST-Chemo	
Hair loss	Chemo[c]	W	MOST-Chemo	
Skin rash	Chemo[c]	W	MOST-Chemo	
Numbness or pins and needles	Chemo[c]	W	MOST-Chemo	
Sore hands and feet	Chemo[c]	W	MOST-Chemo	
*Problems taking tablets		W	Excluded	Consequence not symptom
*Problems with needles or injections		I	Excluded	Consequence not symptom
*Inconvenience of treatment		W	Excluded	Consequence not symptom
*Thought of actually having treatment		I	Excluded	Consequence not symptom
Anxiety (feeling worried)	Psych[d]	I	MOST-Psych	
Depression (feeling sad)	Psych[d]	I	MOST-Psych	
Physical well-being	Wellbeing[e]	I	MOST-Wellbeing	
Emotional wellbeing	Wellbeing[e]	I	MOST-Wellbeing	
Overall wellbeing	Wellbeing[e]	I	MOST-Wellbeing	

a Physical symptoms most likely due to ROC and not chemotherapy.

b Physical symptoms most likely due to chemotherapy and not ROC.

c Physical symptoms reasonably likely due to ROC and/or chemotherapy.

d Psychological symptoms.

e Global wellbeing (physical, emotional and overall).

f Shaded cells indicate correlations ≥ 0.50.

f Direction of change from baseline to cycle 3: I = improved, W = worsened, 0 = zero change

g MOSTv2 scoring rules yield five multi-item scales: MOST-Abdo, MOST-DorT, MOST-Chemo, MOST-Psych, MOST-Wellbeing, with two items (trouble sleeping, bladder problems) to either be included in MOST-DorT or retained as single items, pending psychometric performance of nine and eleven item versions (MOST-DorT-9, MOST-DorT-11).

Abbreviations: MOST, Measure of Ovarian Symptoms and Treatment; ROC, recurrent ovarian cancer

Table 3. Internal consistency (Cronbach alpha) and convergent and divergent validity (Spearman correlation[a] of MOST domains with each other and EORTC, FACT-O and FOSI domains) at baseline (*N*=903)

Domains	MOST-Abdo[a]	MOST-Chemo[b]	MOST-DorT-9[c]	MOST-DorT-11[c]	MOST-Psych[d]	MOST-Wellbeing[e]
Internal consistency	0.80	0.84	0.84	0.65	0.86	0.91
Convergent validity						
EORTC						
OV28-Abdo	**0.77**[f,g]	0.28	0.70	0.71	0.26	-0.42
OV28-Chemo	0.25	**0.60**	0.38	0.40	0.23	-0.30
OV28-Peripheral neuropathy	0.18	**0.67**	0.27	0.28	0.18	-0.24
C30-Fatigue	0.52	0.37	0.69	0.70	0.37	-0.59
C30-Nausea/vomiting	0.43	0.23	0.63	0.61	0.22	-0.33
C30-Diarrhea	0.23	0.15	0.32	0.31	0.14	-0.15
C30-Constipation	0.35	0.22	0.45	0.44	0.17	-0.26
C30-Shortness of breath	0.33	0.25	0.49	0.47	0.18	-0.30
C30-Appetite	0.40	0.22	0.59	0.58	0.22	-0.42
C30-Emotional functioning	-0.34	-0.22	-0.41	-0.45	**-0.71**	0.55
C30-Global Health/QOL	-0.50	-0.33	-0.62	-0.62	-0.39	**0.66**
FACT-O						
Ovarian Cancer Additional concerns	-0.53	-0.35	-0.64	-0.63	-0.34	0.54
Trial Outcome Index	-0.56	-0.42	**-0.71**	**-0.72**	-0.48	0.70
Emotional wellbeing	-0.22	-0.18	-0.27	-0.30	**-0.72**	0.48
FACT-O total	-0.48	-0.39	-0.63	-0.65	-0.59	**0.70**
FOSI						
FOSI-8	-0.72	-0.34	-0.76	-0.77	-0.47	0.61
FOSI-15[h]	-0.63	-0.43	-0.73	-0.75	-0.53	0.68
Divergent						
MOST-Chemo	0.26					
MOST-DorT-9	0.68	0.40				
MOST- DorT-11	0.69	0.40	0.98			
MOST-Psych	0.28	0.25	0.35	0.39		
MOST-Wellbeing	-0.43	-0.30	-0.58	-0.59	-0.57	

a Physical symptoms most likely due to ROC and not chemotherapy.

b Physical symptoms most likely due to chemotherapy and not ROC.

c Physical symptoms reasonably likely due to ROC and/or chemotherapy.

d Psychological symptoms.

e Global wellbeing (physical, emotional and overall).

f Shaded cells indicate correlations ≥ 0.50.

f Shaded cells indicate correlations ≥ 0.50.

g Bolding indicates correlations between MOST and scales from other patient-related outcome measures with similar content or intent.

h FOSI-15 contains 15 of the 18 items in FOSI-18, which was published after GICG-SBS commenced. We used the 15 items available in FACT-O to create FOSI-15 as the closest possible approximation to FOSI-15.

Abbreviations: MOST, Measure of Ovarian Symptoms and Treatment; EORTC, European Organisation for Research and Treatment of Cancer; FACT-O, Functional Assessment of Cancer Therapy-Ovarian; FOSI, FACT Ovarian Symptom Index; QOL, quality of life, ROC, recurrent ovarian cancer.

Table 4. Discriminative validity for MOST multi-item scales *

		Expect poorer outcomes		Expect better outcomes				
Scaled subscale score	**Anchor**	***n***	**Mean (SD)**	***n***	**Mean (SD)**	**Difference in means (95% confidence limits)**	***P***	**Effect size (95% confidence limits)**
Baseline								
MOST-Abdo	Cancer symptoms (yes/no)	643	40.8 (27.46)	249	16.7 (19.53)	24.04 (20.30, 27.77)	<0.001	1.00 (0.84, 1.16)
	Ascites (yes/no)	200	55.3 (26.67)	693	27.9 (24.78)	27.36 (23.39, 31.34)	<0.001	0.96 (0.79, 1.13)
	Abdominal pain or bowel obstruction (yes/no)[a]	364	45.3 (27.14)	529	26.4 (25.31)	18.90 (15.42, 22.38)	<0.001	0.73 (0.59, 0.87)
	ECOG: 2-3 vs 0-1[b]	98	45.4 (28.57)	795	32.7 (27.24)	12.74 (6.99, 18.50)	<0.001	0.42 (0.21, 0.64)
MOST-DorT-9	Cancer symptoms (yes/no)	638	23.2 (17.28)	250	11.9 (13.24)	11.27 (8.90, 13.65)	<0.001	0.82 (0.67, 0.98)
	Ascites (yes/no)	200	30.2 (19.75)	689	17.1 (14.88)	13.13 (10.59, 15.67)	<0.001	0.77 (0.60, 0.94)
	Abdominal pain or bowel obstruction (yes/no)[a]	361	25.6 (18.39)	528	16.2 (14.84)	9.43 (7.24, 11.63)	<0.001	0.59 (0.46, 0.73)
	ECOG: 2-3 vs 0-1[b]	97	32.9 (18.52)	792	18.4 (16.13)	14.46 (11.00, 17.93)	<0.001	0.85 (0.62, 1.09)
	CTC-AE score >median[c]	429	23.8 (17.42)	458	16.4 (15.57)	7.39 (5.21, 9.56)	<0.001	0.53 (0.39, 0.66)
MOST-DorT-11	Cancer symptoms (yes/no)	638	23.5 (16.28)	250	12.5 (12.78)	11.02 (8.77, 13.28)	<0.001	0.85 (0.69, 1.00)
	Ascites (yes/no)	200	29.3 (18.58)	689	17.8 (14.38)	11.51 (9.07, 13.94)	<0.001	0.71 (0.54, 0.88)
	Abdominal pain or bowel obstruction (yes/no)[a]	361	25.7 (17.19)	528	16.7 (14.31)	8.96 (6.88, 11.04)	<0.001	0.60 (0.46, 0.73)
	ECOG: 2-3 vs 0-1[b]	97	32.0 (17.23)	792	19.0 (15.42)	13.09 (9.79, 16.39)	<0.001	0.81 (0.57, 1.04)
	CTC-AE score >median[c]	429	23.7 (16.27)	458	17.2 (15.17)	6.47 (4.40, 8.54)	<0.001	0.48 (0.35, 0.62)
MOST-Psych	Cancer symptoms (yes/no)	645	34.1 (27.07)	250	32.2 (27.48)	1.92 (-2.05, 5.90)	0.18	0.10 (-0.05, 0.25)
	Ascites (yes/no)	201	34.8 (27.39)	695	33.2 (27.11)	1.56 (-2.71, 5.83)	0.39	0.07 (-0.09,0.23)
	Abdominal pain or bowel obstruction (yes/no)[a]	366	35.5 (28.13)	530	32.2 (26.43)	3.33 (-0.29, 6.94)	0.096	0.12 (-0.02,0.25)
	ECOG: 2-3 vs 0-1[b]	97	36.5 (31.13)	799	33.2 (26.65)	3.28 (-2.45, 9.02)	0.60	0.06 (-0.15,0.27)

		Expect poorer outcomes		Expect better outcomes				
Scaled subscale score	**Anchor**	***n***	**Mean (SD)**	***n***	**Mean (SD)**	**Difference in means (95% confidence limits)**	***P***	**Effect size (95% confidence limits)**
MOST Well-being index	Cancer symptoms (yes/no)	629	55.0 (19.83)	247	64.5 (18.16)	-9.50 (-12.36, -6.65)	<0.001	-0.49 (-0.64, -0.34)
	Ascites (yes/no)	195	52.0 (21.57)	682	59.3 (19.00)	-7.29 (-10.41, -4.17)	<0.001	-0.37 (-0.53, -0.21)
	Abdominal pain or bowel obstruction (yes/no)[a]	357	53.6 (20.41)	520	60.5 (18.93)	-6.85 (-9.48, -4.21)	<0.001	-0.35 (-0.49, -0.21)
	ECOG: 2-3 vs 0-1[b]	94	45.8 (21.74)	783	59.1 (19.10)	-13.34 (-17.49, -9.18)	<0.001	-0.69 (-0.92, -0.45)
Pre-cycle 3								
MOST-DorT-9	CTC-AE score >median[c]	238	28.9 (18.65)	435	16.9 (14.43)	11.96 (9.42, 14.50)	<0.001	0.75 (0.58, 0.91)
MOST-DorT-11	CTC-AE score >median[c]	238	27.6 (17.52)	435	17.2 (14.00)	10.41 (7.98, 12.83)	<0.001	0.69 (0.52, 0.85)
MOST-Chemo	CTC-AE score >median[d]	305	24.6 (17.10)	370	16.9 (14.68)	7.73 (5.32, 10.13)	<0.001	0.52 (0.36, 0.67)

a Clinician-rated symptoms of cramping abdominal pain or intermittent/incomplete bowel obstruction

b Performance status: poor [ECOG 2-3] vs good [ECOG 0-1]

c Score includes: nausea, vomiting, diarrhea, constipation, anorexia, fatigue (representing 6 of the 9–11 MOST-DorT items)

d Score includes: neuropathy-sensory, hand-foot syndrome, alopecia, erythema, mucositis/stomatitis (representing 5 of the 6 MOST-Chemo items)

Abbreviations: MOST, Measure of Ovarian Symptoms and Treatment; ECOG, Eastern Cooperative Oncology Group performance status; CTC-AE, Common Terminology Criteria for Adverse Events.

Table 5. Mean change from baseline to pre-cycle 3 of chemotherapy for the MOSTv2 multi-item scales, grouped by clinician-rated baseline abdominal symptoms and type of recurrent ovarian cancer

MOST multi-item scale	*n*	Mean (95% CL)[a]	*P*[b]	*n*	Mean (95% CL)[a]	*P*[b]	*P*[c]
Baseline abdominal symptoms[d]		**Yes**			**No**		
MOST-Abdo	249	-11.1 (-14.9, -7.29)	<0.001	407	-0.59 (-2.90, 1.73)	0.62	<0.001
MOST-Chemo	252	6.46 (4.34, 8.58)	<0.001	412	7.01 (5.42, 8.60)	<0.001	0.68
MOST-DorT-9	247	0.56 (-1.81, 2.92)	0.64	409	5.02 (3.55, 6.49)	<0.001	<0.001
MOST- DorT-11	247	-0.08 (-2.33, 2.16)	0.94	409	4.05 (2.71, 5.38)	<0.001	<0.001
MOST-Psych	253	-4.11 (-7.24, -0.98)	0.01	414	-2.78 (-4.89, -0.67)	0.01	0.47
MOST-Wellbeing	233	-0.64 (-3.48, 2.21)	0.60	389	1.79 (-0.15, 3.72)	0.07	0.15
ROC type[e]		**PRROC**			**PPS≥3**		
MOST-Abdo	375	-4.84 (-7.65, -2.03)	<0.001	281	-4.23 (-7.31, -1.15)	0.007	0.78
MOST-Chemo	380	4.49 (2.86, 6.12)	<0.001	284	9.90 (7.93, 11.87)	<0.001	<0.001
MOST-DorT-9	376	2.35 (0.65, 4.05)	0.01	280	4.67 (2.71, 6.64)	<0.001	0.08
MOST- DorT-11	376	1.61 (0.02, 3.20)	0.05	280	3.67 (1.86, 5.49)	<0.001	0.09
MOST-Psych	380	-4.12 (-6.53, -1.71)	<0.001	287	-2.18 (-4.76, 0.41)	0.10	0.29
MOST-Wellbeing	351	0.38 (-1.68, 2.45,)	0.71	271	1.52 (-1.05, 4.08)	0.25	0.49

a Negative change indicates improvement, positive change indicates worsening (the direction of the MOST-Wellbeing scale has been reversed so it can be interpreted consistently with the MOST symptom scales).

b Paired *t* test for change over time within a group.

c 2-sample *t* test for difference between groups in change over time.

d Clinician-rated symptoms of cramping abdominal pain or intermittent or incomplete bowel obstruction.

e PRROC or potentially platinum sensitive ROC with ≥3 lines of prior chemotherapy (PPS≥3).

Abbreviations: MOST, Measure of Ovarian Symptoms and Treatment; ROC, recurrent ovarian cancer; PRROC, platinum-resistant or refractory ROC; PPS, potentially platinum sensitive ROC.

Table 6. Responsiveness to change in ovarian cancer symptoms

Group expected to change	Experienced symptoms at baseline	Experienced improvement with chemotherapy	*n*	Mean change	SD (change)	*P*	Effect size (95% confidence limits)
Definition 1	Scores GE 40 on MOST-Abdo scale (mean of 4 or more on items 4 and 5)	Abdo change items (4 and 5): average change of 1 or 1.5	57	-43.1	21.4	<0.001	-2.01 (-2.52, -1.51)
Definition 2	Rated by clinicians as having abdominal symptoms	Scored 1 (much better) on the global wellbeing item (18)	48	-26.3	25.9	<0.001	-1.02 (-1.39, -0.64)

Click here to download Figure Figure legend.docx

Figure legends and footnotes:

Fig. 1 Baseline prevalence of physical symptoms and other treatment related problems* assessed by the MOST version 1[a]
* Treatment-related problems

Fig. 2 Mean change from baseline for each MOSTv1 item (*n*=685 who completed MOST prior to Cycle 3 chemotherapy)

Fig. 3 Boxplots of change scores for each MOST multi-item scale grouped by patient-rated change category from the most relevant MOST-Change item

Fig. A Boxplots for MOST Disease or Treatment (MOST-DorT) items corresponding to Common Terminology for Adverse Events (CTC-AE) Nausea, Vomiting, Diarrhea, Constipation, Anorexia, Fatigue, plotted by CTC-AE grade

Fig. B Boxplots for MOST items caused solely by chemotherapy (MOST-Chemo) corresponding to Common Terminology for Adverse Events (CTC-AE) Neuropathy – sensory, Hand-foot syndrome, Alopecia, Erythema, Mucositis/Stomatitis, plotted by CTC-AE grade

Figure 1

Click here to download Figure Fig 1.tif

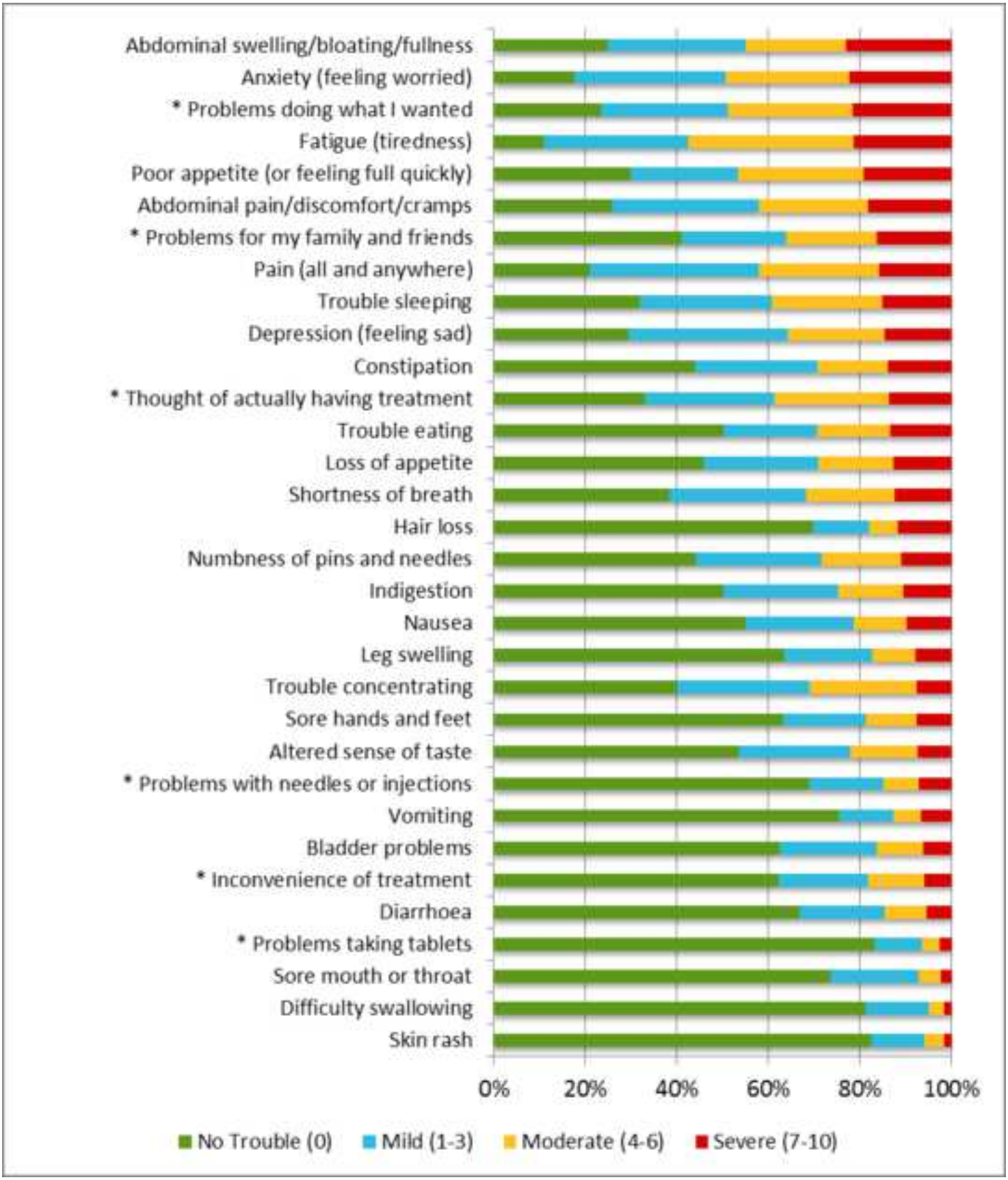


Figure 2

Click here to download Figure Fig 2.tif

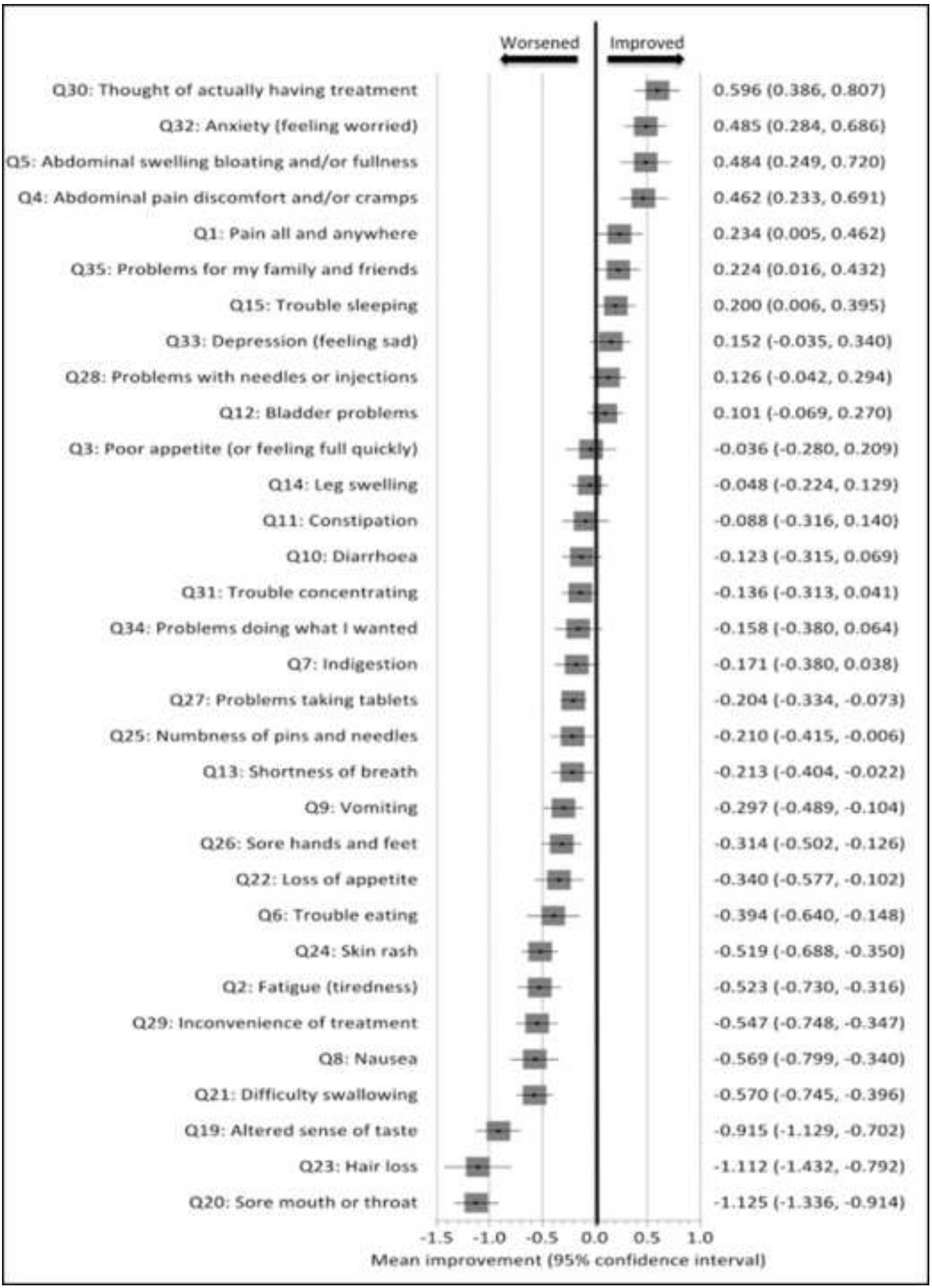
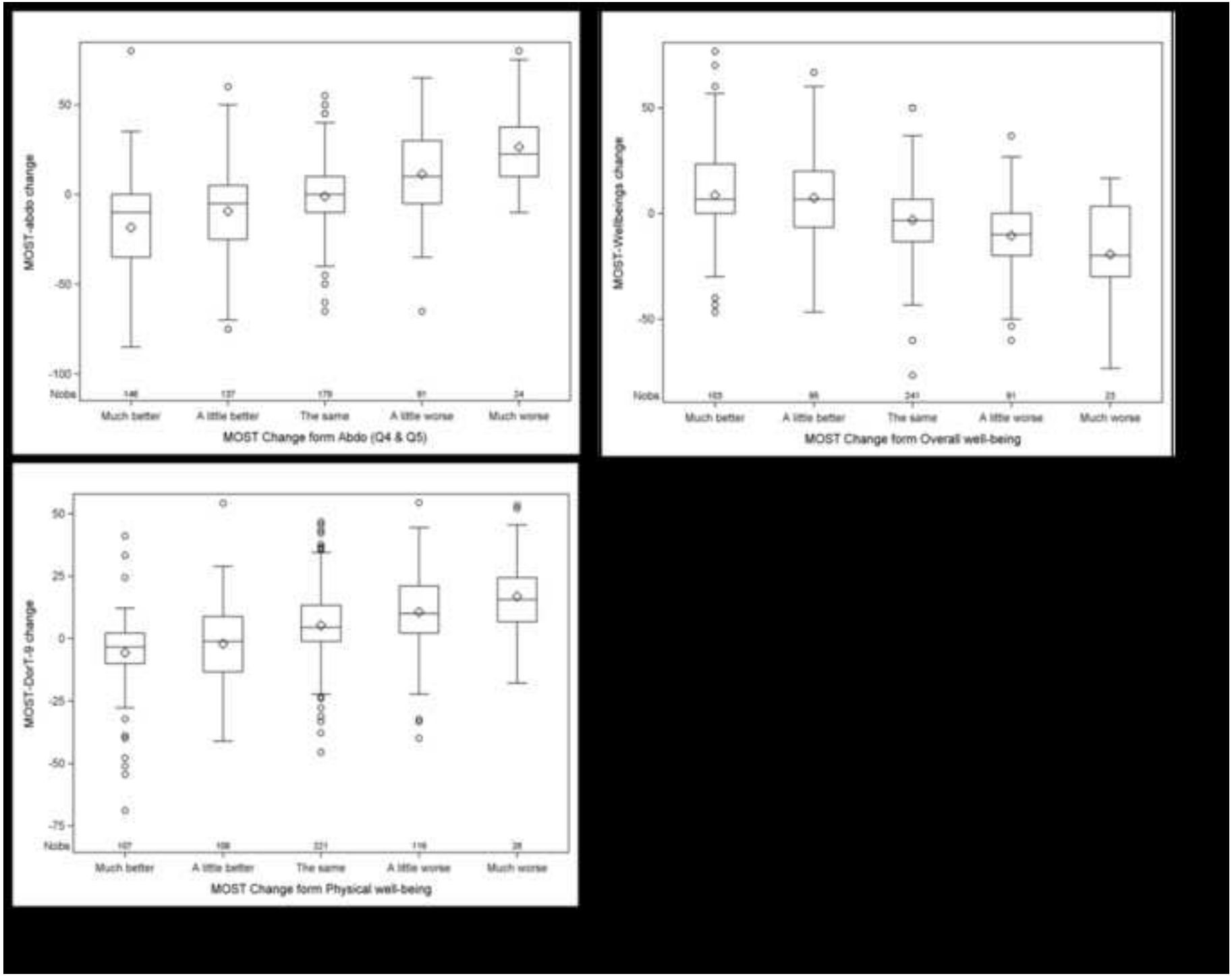


Figure 3

Click here to download Figure Fig 3.tif



Electronic supplementary material

Click here to download Electronic supplementary material ESM-1.pdf

Online supplement

Measuring what matters MOST: Validation of the Measure of Ovarian Symptoms and Treatment, a patient-reported outcome measure of symptom burden and impact of chemotherapy in recurrent ovarian cancer (ROC)

Quality of Life research

Madeleine T King*, Martin R Stockler, Rachel O'Connell, Luke Buizen, Florence Joly, Anne Lanceley, Felix Hilpert, Aikou Okamoto, Eriko Aotani, Jane Bryce, Paul Donnellan, Amit Oza, Elisabeth Avall-Lundqvist, Jonathan Berek, Jalid Sehouli, Mandy Feeney, Dominique Berton-Rigaud, and Michael Friedlander for the GCIG Symptom Benefit group

*University of Sydney, madeleine.king@sydney.edu.au

Table A. Internal consistency (Cronbach alpha) and convergent validity (Spearman correlation of MOST domains with EORTC, FACT-O and FOSI domains) at pre-cycle 3 (N=685)

	MOST-Abdo[a]	MOST-Chemo[b]	MOST-DorT-9[c]	MOST-DorT-11[c]	MOST-Psych[d]	MOST-Wellbeing[e]
Internal consistency	0.82	0.84	0.85	0.66	0.90	0.95
Convergent validity						
EORTC						
OV28-Abdo	**0.75**[f,g]	0.32	0.62	0.64	0.41	-0.40
OV28-Chemo	0.32	**0.60**	0.39	0.43	0.32	-0.31
OV28-Peripheral neuropathy	0.27	**0.54**	0.32	0.33	0.29	-0.25
C30-Fatigue	0.45	0.35	0.64	0.64	0.46	-0.58
C30-Nausea/vomiting	0.32	0.21	0.52	0.49	0.23	-0.28
C30-Diarrhea	0.22	0.12	0.27	0.26	0.13	-0.15
C30-Constipation	0.33	0.19	0.35	0.36	0.22	-0.21
C30-Shortness of breath	0.33	0.29	0.47	0.47	0.30	-0.35
C30-Appetite	0.40	0.30	0.63	0.61	0.36	-0.46
C30-Emotional functioning	-0.38	-0.31	-0.45	-0.49	**-0.75**	0.57
C30-Global health/QOL	-0.42	-0.32	-0.56	-0.57	-0.48	**0.68**
FACT-O						
Ovarian Cancer Additional concerns	-0.56	-0.37	-0.62	-0.62	-0.50	0.59
Trial Outcome Index	-0.57	-0.41	**-0.70***	**-0.71***	-0.62	0.74
Emotional wellbeing	-0.36	-0.21	-0.38	-0.41	**-0.74***	0.59
FACT-O total	-0.53	-0.37	-0.63	-0.65	-0.69	**0.76***
FOSI						
FOSI-8	-0.70	-0.39	-0.72	-0.73	-0.60	0.64
FOSI-15[2]	-0.62	-0.45	-0.70	-0.72	-0.67	0.72
MOST						
MOST-Chemo	0.35					
MOST-DorT-9	0.67	0.50				
MOST- DorT-11	0.69	0.52	0.98			
MOST-Psych	0.44	0.39	0.52	0.56		
MOST-Wellbeing	-0.43	-0.33	-0.59	-0.61	-0.67	

a Physical symptoms most likely due to ROC and not chemotherapy.

b Physical symptoms most likely due to chemotherapy and not ROC.

c Physical symptoms reasonably likely due to ROC and/or chemotherapy.

d Psychological symptoms.

e Global wellbeing (physical, emotional and overall).

f Shaded cells indicate correlations ≥ 0.50.

g Bolding indicates correlations between MOST and scales from other patient-related outcome measures with similar content or intent.

h FOSI-15 contains 15 of the 18 items in FOSI-18, which was published after GICG-SBS commenced. We used the 15 items available in FACT-O to create FOSI-15 as the closest possible approximation to FOSI-15.

Abbreviations: QOL, quality of life; ROC, recurrent ovarian cancer.

Fig. A Boxplots for MOST Disease or Treatment (MOST-DorT) items corresponding to CTC-AE Nausea, Vomiting, Diarrhea, Constipation, Anorexia, Fatigue, plotted by CTC-AE grade

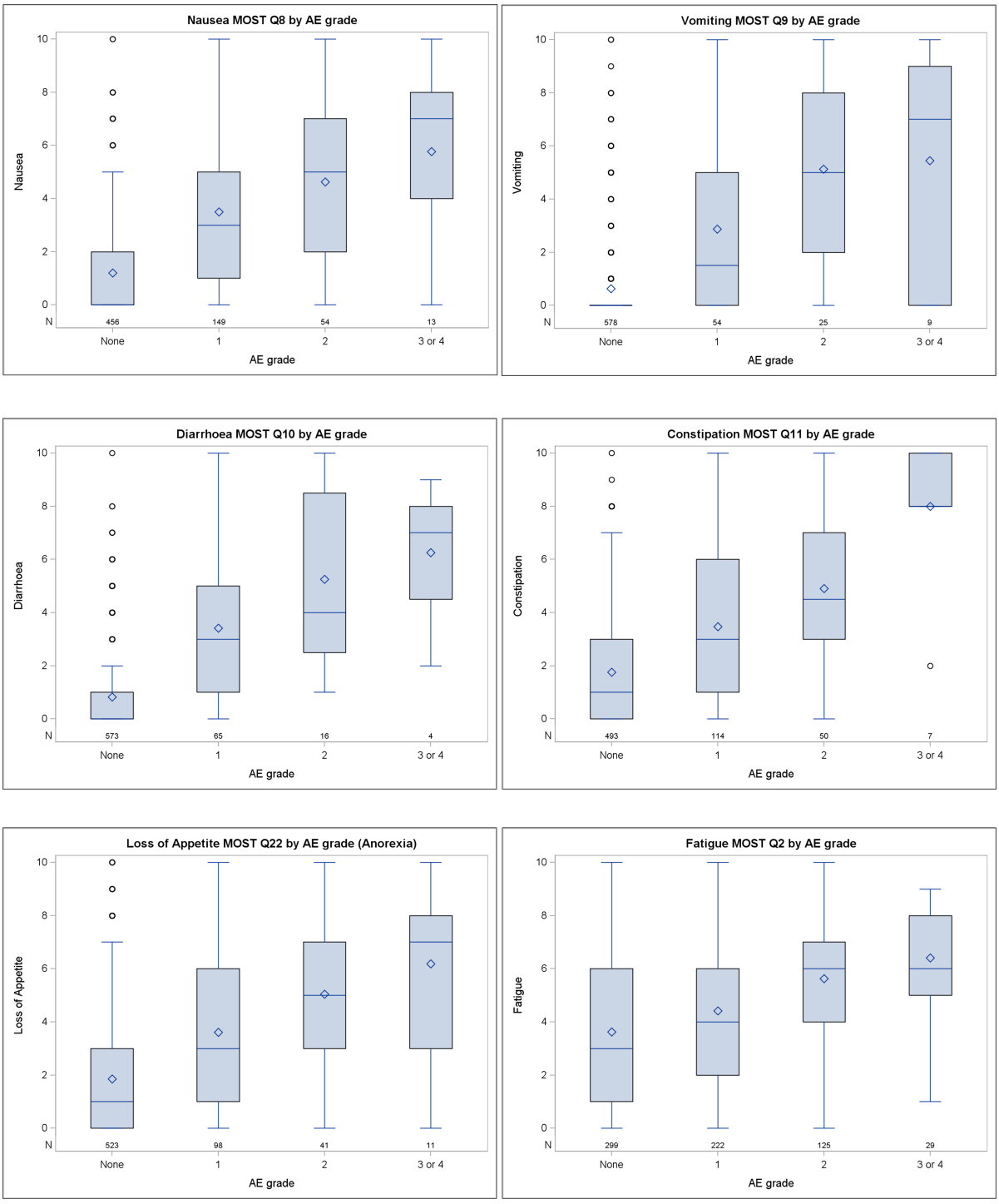
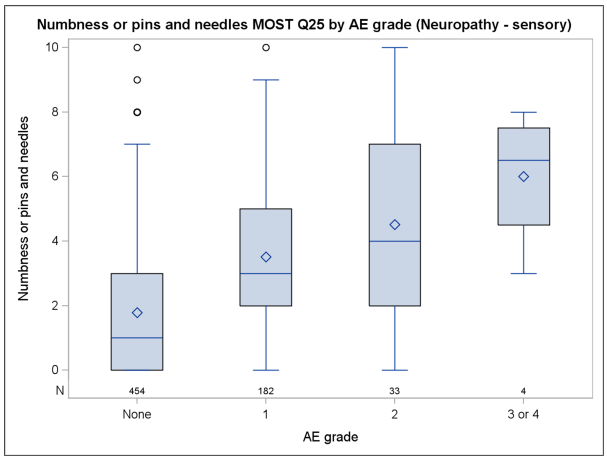
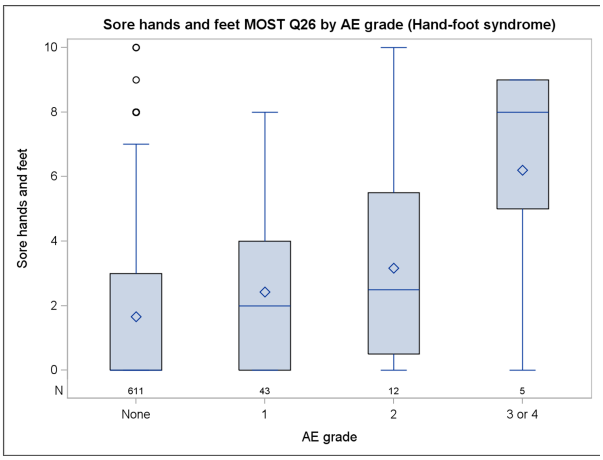
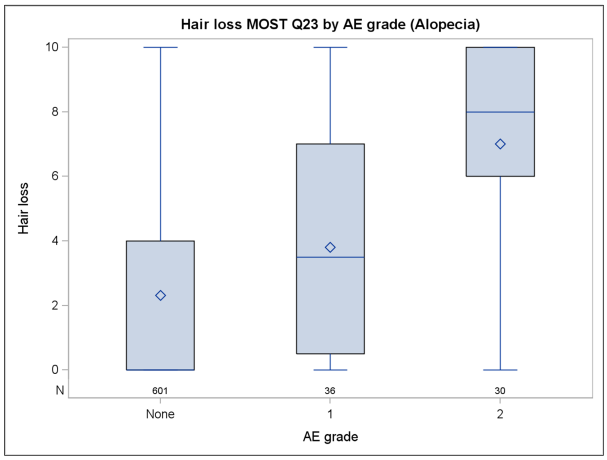
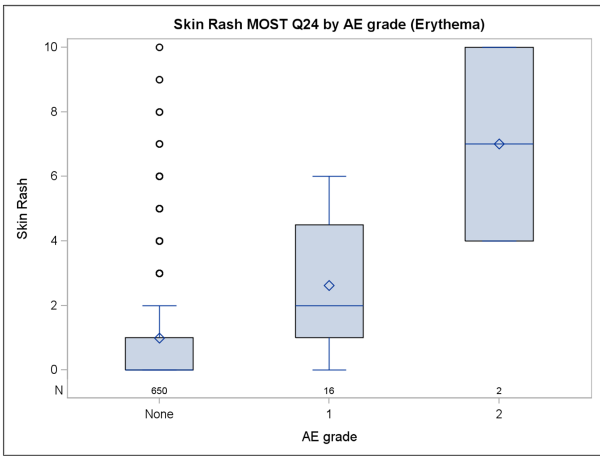
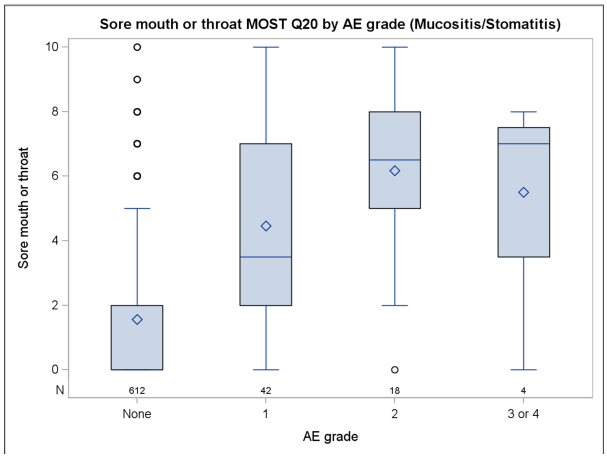



Fig. B Boxplots for MOST items caused solely by chemotherapy (MOST-Chemo) corresponding to CTC-AE Neuropathy – sensory, Hand-foot syndrome, Alopecia, Erythema, Mucositis/Stomatitis, plotted by CTC-AE grade