

Hospital, Gifu University, Yamaguchi University, Kyoundo Hospital, Niigata City General Hospital, Cancer Institute Hospital, Takamatsu Red Cross Hospital, Gunma University, Himeji Red Cross Hospital, Juntendo University Urayasu Hospital, Toyama University, Hamamatsu Medical Center,

Kushiro Rosai Hospital, Kawasaki Medical School Hospital, Asahikawa City Hospital, Kansai Medical University, Yokohama Municipal Citizen's Hospital, Osaka Medical College Hospital, Okazaki City Hospital, Tottori Prefectural Central Hospital, Saitama Medical Center Jichi Medical University.

Annals of Oncology 25: 257–264, 2014

doi:10.1093/annonc/mdt409

Published online 19 November 2013

Physician-assessed and patient-reported outcome measures in chemotherapy-induced sensory peripheral neurotoxicity: two sides of the same coin

P. Alberti^{1,†}, E. Rossi^{2,†}, D. R. Cornblath³, I. S. J. Merkies^{4,5}, T. J. Postma⁶, B. Frigeni¹, J. Bruna⁷, R. Velasco⁷, A. A. Argyriou⁸, H. P. Kalofonos⁸, D. Psimaras⁹, D. Ricard¹⁰, A. Pace¹¹, E. Galìè¹¹, C. Briani¹², C. Dalla Torre¹², C. G. Faber⁵, R. I. Lalisang¹³, W. Boogerd¹⁴, D. Brandsma¹⁴, S. Koeppen¹⁵, J. Hense¹⁵, D. Storey¹⁶, S. Kerrigan¹⁶, A. Schenone¹⁷, S. Fabbri¹⁷, M. G. Valsecchi² & G. Cavaletti^{1*} the CI-PeriNomS Group[‡]

¹Department of Surgery and Translational Medicine; ²Center of Biostatistics for Clinical Epidemiology, Department of Health Sciences, University of Milano-Bicocca, Monza, Italy; ³Department of Neurology, Johns Hopkins University School of Medicine, Baltimore, USA; ⁴Department of Neurology, Spaarne Hospital, Hoofddorp; ⁵Department of Neurology, Maastricht University Medical Center, Maastricht; ⁶Department of Neurology, VU University Medical Center, Amsterdam, The Netherlands; ⁷Unit of Neuro-Oncology, Department of Neurology, University Hospital of Bellvitge-ICO Duran I Reynals, L'Hospitalet, Barcelona, Spain; ⁸Division of Clinical Oncology, Department of Medicine, University Hospital of Patras, Patras, Greece; ⁹Service de Neurologie Mazarin, Hôpital de la Pitié-Salpêtrière, AP-HP, Paris; ¹⁰Service de Neurologie, Hôpital du Val-de-Grâce, Service de Santé des Armées, Paris, France; ¹¹Neurology Unit, National Cancer Institute Regina Elena, Rome; ¹²Department of Neurological, Psychiatric, Sensorial, Reconstructive and Rehabilitative Sciences, University of Padova, Padova, Italy; ¹³Division of Medical Oncology, Department of Internal Medicine, GROW-School of Oncology and Developmental Biology, Maastricht University Medical Center, Maastricht; ¹⁴Department of Neuro-Oncology, Netherlands Cancer Institute, Amsterdam, The Netherlands; ¹⁵Department of Neurology and West German Cancer Center, University of Essen, Essen, Germany; ¹⁶Edinburgh Centre for Neuro-Oncology and Edinburgh Cancer Research Center, Western General Hospital, Edinburgh, UK; ¹⁷Department of Neuroscience, Ophthalmology and Genetic, Center of Excellence for Biomedical Research, University of Genova, Genova, Italy

Received 27 May 2013; revised 23 August 2013; accepted 26 August 2013

Background: The different perception and assessment of chemotherapy-induced peripheral neurotoxicity (CIPN) between healthcare providers and patients has not yet been fully addressed, although these two approaches might eventually lead to inconsistent, possibly conflicting interpretation, especially regarding sensory impairment.

Patients and methods: A cohort of 281 subjects with stable CIPN was evaluated with the National Cancer Institute—Common Toxicity Criteria (NCI-CTC v. 2.0) sensory scale, the clinical Total Neuropathy Score (TNSc©), the modified Inflammatory Neuropathy Cause and Treatment (INCAT) sensory sumscore (mISS) and the European Organization for Research and Treatment of Cancer CIPN specific self-report questionnaire (EORTC QOL-CIPN20).

Results: Patients' probability estimates showed that the EORTC QLQ-CIPN20 sensory score was overall more highly related to the NCI-CTC sensory score. However, the vibration perception item of the TNSc had a higher probability to be scored 0 for EORTC QLQ-CIPN20 scores lower than 35, as vibration score 2 for EORTC QLQ-CIPN20 scores between 35 and 50 and as grade 3 or 4 for EORTC QLQ-CIPN20 scores higher than 50. The linear models showed a significant trend between each mISS item and increasing EORTC QLQ-CIPN20 sensory scores.

*Correspondence to: Prof. Guido Cavaletti, Department of Surgery and Translational Medicine, University of Milano-Bicocca, via Cadore 48, 20900 Monza, Milano-Bicocca, Italy. Tel: +39-02-6448-8039; Fax: +39-02-6448-8250; E-mail: guido.cavaletti@unimib.it

[†]Equally contributed to the paper.

[‡]See appendix for the complete list of participating centers and investigators.

Conclusion: None of the clinical items had a perfect relationship with patients' perception, and most of the discrepancies stood in the intermediate levels of CIPN severity. Our data indicate that to achieve a comprehensive knowledge of CIPN including a reliable assessment of both the severity and the quality of CIPN-related sensory impairment, clinical and PRO measures should be always combined.

Key words: chemotherapy, neuropathy, assessment, patient-reported outcome measure, neurotoxicity

introduction

The best way to measure chemotherapy-induced peripheral neurotoxicity (CIPN) [1] remains unclear. The most frequently used physician-assessed measure is the National Cancer Institute—Common Toxicity Criteria (NCI-CTC, here v. 2.0 was applied) sensory and motor scale (http://ctep.cancer.gov/protocolDevelopment/electronic_applications/ctc.htm#ctc_40). As it has well-known limitations in the evaluation of CIPN [2, 3], other physician-derived tools have been developed. The Total Neuropathy Score[®] and its clinical version (TNSc[®]) address many of issues regarding NCI-CTC [4–6]. The modified Inflammatory Neuropathy Cause and Treatment sensory subscore (mISS) is a widely used, standardized scale for scoring sensory neuropathies [7], recently investigated for its validity and reliability in assessment of stable CIPN in comparison with NCI-CTC and TNSc[®] [8].

Patient-reported outcomes (PROs) have also been used and have been proposed for use in medical product development to support labeling claims (<http://www.fda.gov/downloads/Drugs/Guidances/UCM193282.pdf>). Several different quality-of-life (QoL) questionnaires have been used in cancer patients [3] and the European Organization for Research and Treatment of Cancer (EORTC)-QLQ30 [9] is one of the most widely used. It is intended to be supplemented by additional condition-specific subscales, such as EORTC QLQ-CIPN20 to evaluate the impact of CIPN on QoL [8, 9].

The possible different perceptions and assessments of CIPN between physicians and patients have not yet been fully addressed. Prior studies have suggested that these two approaches might lead to inconsistent results, particularly in the interpretation of sensory impairment [10–12].

This secondary analysis of the CI-PeriNomS study database [8] compares results of sensory examination obtained by physicians through NCI-CTC, TNSc[®] and mISS with EORTC QLQ-C30 and sensory submodule EORTC QLQ-CIPN20 questionnaires results.

patients and methods

Patients analyzed in the present study were the 281 subjects with stable CIPN enrolled in the primary CI-PeriNomS study (see supplementary Appendix 2, available at *Annals of Oncology* online) [3, 8].

assessment methods

The mISS scoring method was used to assess pinprick, vibration, light touch and joint position sensations in arms and legs. Vibration sense was assessed using the graduated Rydel-Seiffer tuning fork and its reported normative data [13]. For two-point discrimination test, new normative values were used [14]. Light touch and pinprick were assessed by the use of standardized 10 g monofilaments and disposable Neurotip as part of the calibrated Neuropen (Owens Mumford, Woodstock, UK). Pinprick, vibration and light

touch proximal-to-distal impairments were recorded and subdivided in four scores in both the upper and lower limbs as follows: normal = 0, disturbed perception up to the distal phalanx of the index finger or hallux = 1, up to the ulnar styloid process or medial malleolus = 2, up to the medial humerus epicondyle or patella = 3, up to the acromio-clavicular joint or anterior superior iliac spine = 4. Two-point discrimination was assessed and scored as normal/abnormal at the index finger. Moreover, the NCI-CTC grade and the TNSc[®] score regarding pin and vibration thresholds (TNSc[®] PIN and TNSc[®] VIBRATION, respectively) were calculated as previously described [4, 5].

At the 'QoL level', EORTC QLQ-C30 and disease-specific CIPN20 questionnaires were selected and used in the patient's language [8].

data handling

EORTC QLQ-C30 questionnaire results and EORTC QLQ-CIPN20 sensory subscale results were converted into a 0–100 scale (0 = no sensory impairment, 100 = worst sensory impairment) [13]. The intra- and interexaminer agreement of each mISS item was confirmed by means of weighted K-Cohen coefficients and 95% confidence intervals. Agreement between evaluations was described according to Landis and Koch [15]. This had already been done for TNSc[®], NCI-CTC, overall mISS scales and QLQ-CIPN20 sensory submodules [8]. Therefore, in the secondary analysis presented in this article only the first evaluation for each patient was considered. When analyzing TNSc[®], NCI-CTC, mISS items, score 3 and 4 in each ordinal scale were combined due to the low number of patients in these classes. See supplementary Appendix 2, available at *Annals of Oncology* online for more details.

statistical analysis

To test the relationship between EORTC QLQ-CIPN20 sensory sub-module and the evaluation of CIPN done by the physician (by means of NCI-CTC, TNSc[®] sensory items and mISS scale), different statistical methods were used. First, the presence of a linear trend was assessed by means of a linear regression model of EORTC QLQ-CIPN20 sensory score over each of the scores assigned by the physician, assuming NCI-CTC, TNSc[®] sensory items and mISS scale as continuous variables. Afterward, to evaluate difference in EORTC QLQ-CIPN20 sensory score between contiguous scores of the three physician assessed scores, these latter were considered as ordinal categorical regressors and pairwise comparisons between their contiguous scores were evaluated by means of the ANOVA regression, with Bonferroni adjustment for multiple comparisons (*P* values were significant if <0.05).

To describe the patient probability to be assigned to higher (or lower) sensory levels in the NCI-CTC, TNSc[®] sensory scores and mISS items, according to his/her EORTC QLQ-CIPN20 sensory scores, ordinal logistic models were applied, where EORTC QLQ-CIPN20 sensory score was the independent variable and NCI-CTC, the TNSc[®] sensory scores and mISS items were the ordinal dependent variables, one by one. In Figures 1 and 2 (lower panels), each curve represents the estimated patient's probabilities to be assigned to different scores of the physician scales according to levels of EORTC QLQ-CIPN20 sensory score. When the curve of a specific score is higher than the others, it represents the most probable grade/score in which a patient could be assigned when he/she presents the corresponding QLQ-

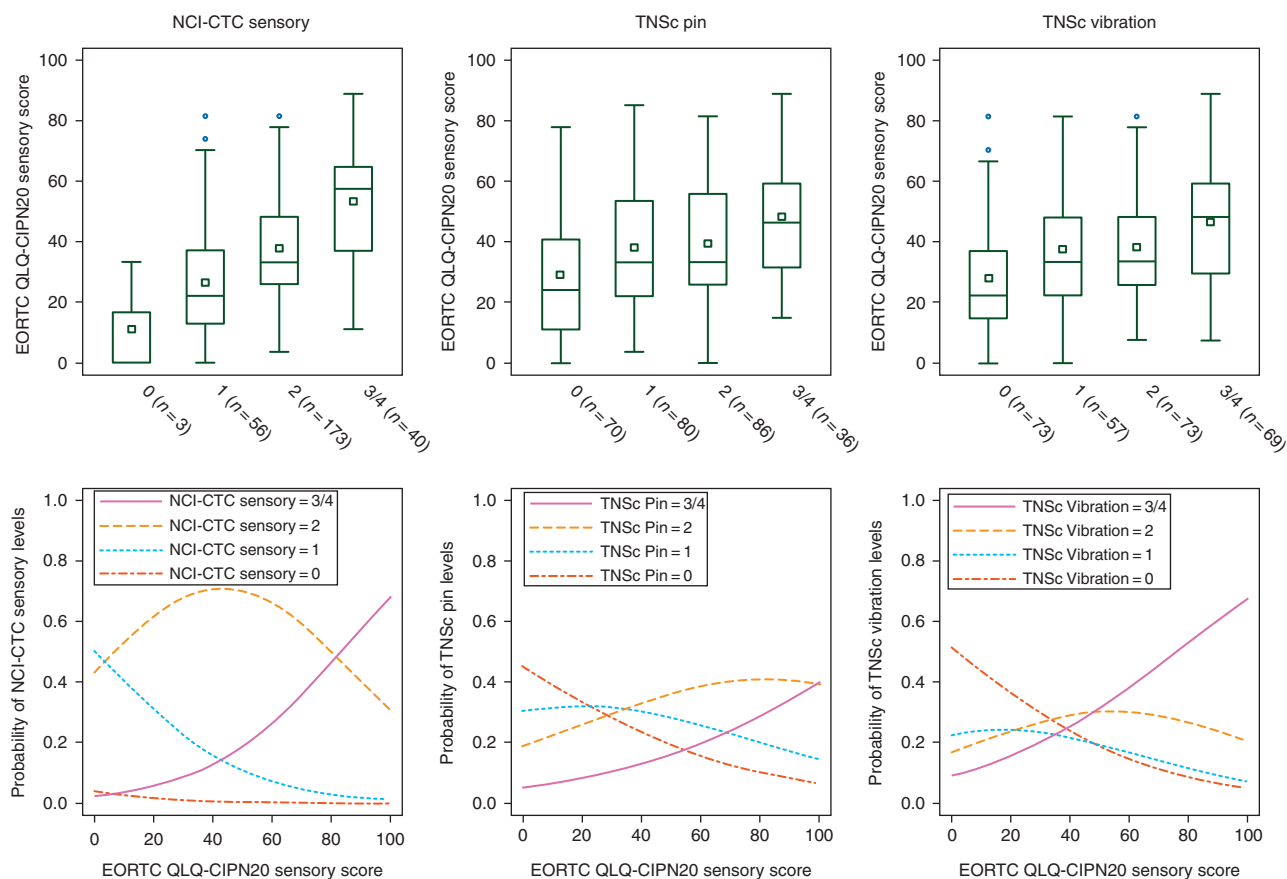


Figure 1. EORTC QLQ-CIPN20 sensory score distribution by NCI-CTC sensory, TNSc pin and vibration item scores (upper panel, in brackets is reported the number of patients available for comparison). Upper panel: each box represents the first and third quartile, the central line the median, the white square the mean and the whiskers are located at the maximum and minimum observation (outside observations indicated with dots are those out of the $1.5 \times$ interquartile range). Lower panel: the probability for a patient, with a given EORTC Lower panel: QLQ-CIPN20 sensory score, to be classified as score 0 to 4 of NCI-CTC sensory, TNSc PIN and VIBRATION item is reported.

CIPN20 sensory score. Analyses were carried out by means of the statistical software SAS v.9.2 (SAS Institute, Inc., Cary, NC).

results

At the overall QoL level, the analysis of the EORTC QLQ-C30 results confirmed that a significant proportion of patients in our cohort reported an impairment in both physical (60% having a score lower than 80) and emotional functioning (50% having a score lower than 80, supplementary Figure S1, available at *Annals of Oncology* online).

reliability analysis on mISS

Out of the entire cohort, 248 subjects underwent the mISS assessment and were available for the reliability analysis, which ruled out the null hypothesis of no agreement in both inter- and intrarater results ($P < 0.001$). Strength of correlation was different for different items: a 'substantial' (i.e., >0.60) inter- and intra-agreement was evidenced for most items, with 'almost perfect' (i.e., >0.80) and 'moderate' (i.e., >0.40) results for a few of them (Supplementary Table S1, available at *Annals of Oncology* online).

comparison among the NCI-CTC, TNSc© and EORTC QLQ-CIPN20 results

According to linear regression analysis, EORTC QLQ-CIPN20 scores tended to significantly increase at increasing grades/scores of each scale (NCI-CTC sensory and TNSc© PIN/VIBRATION items, Table 1, linear trend test). Pairwise comparisons show that this increase in EORTC QLQ-CIPN20 is more evident when contrasting pathological grades in NCI-CTC sensory item and when contrasting the normal versus score 1 on TNSc© (Table 1, Pairwise comparisons and Figure 1, upper panel).

Based on ordinal logistic model, the EORTC QLQ-CIPN20 sensory score was highly related to the NCI-CTC sensory grade, because patients with EORTC QLQ-CIPN20 sensory scores close to 0 had a higher probability to be considered as NCI-CTC sensory grade 1, while increasing scores, approximately up to 80, mainly had a high probability of grade 2 and higher scores had a very high probability to be considered as grade 3/4. Grade 0 had an extremely low probability to be associated with any extent of increase in EORTC QLQ-CIPN20 sensory scores (but this was estimated based on only 3 patients with grade 0) (Figure 1 lower panel).

A less clear relationship was found in the TNSc© models. The TNSc© PIN score 0 achieved the higher probability for EORTC

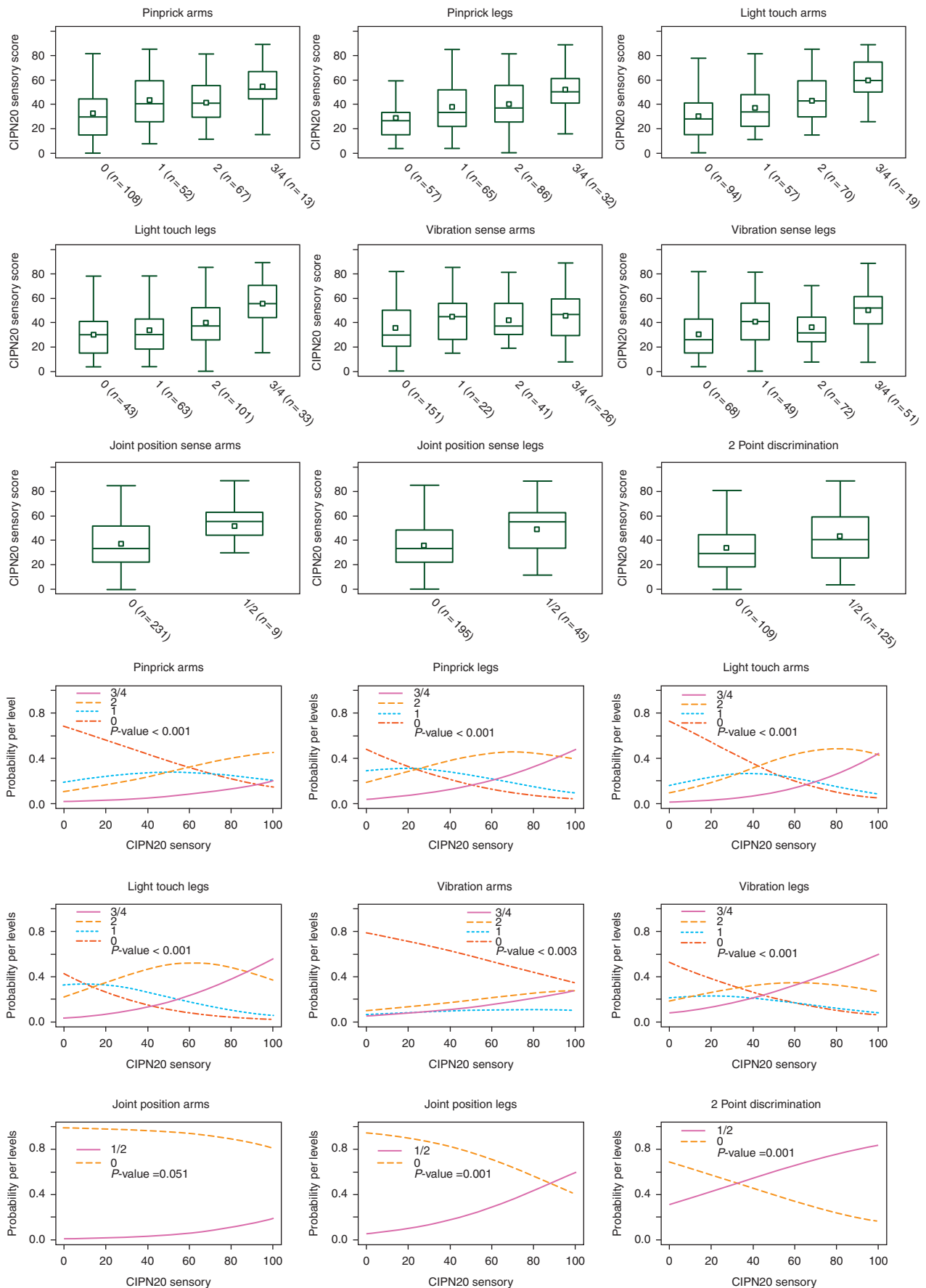


Table 1. Regression models of EORTC QLQ-CIPN20 score over each scale assessed by the physician (NCI-CTC sensory, TNSc pin and TNSc vibration and single mISS items): pairwise comparisons, trend test (β) and P -values

Sensory scales	Pairwise comparisons between item grades (adjusted P -value)			Linear trend Test β (P -value)
	G0 versus G1	G1 versus G2	G2 versus G3/G4	
NCI-CTC sensory	0.532	<0.001	<0.001	13.3 (<0.001)
TNSc PIN	0.020	1.000	0.090	5.5 (<0.001)
TNSc VIBRATION	0.016	1.000	0.032	5.7 (<0.001)
mISS items	Pairwise comparison between item grades (adjusted P -value)			Linear trend test β (P -value)
	G0 versus G1	G1 versus G2	G2 versus G3/G4	
Pinprick arms	0.002	1.000	0.081	5.6 (<0.001)
Pinprick legs	0.043	1.000	0.012	6.7 (<0.001)
Light touch arms	0.091	0.214	0.002	8.0 (<0.001)
Light touch legs	1.000	0.098	<0.001	7.4 (<0.001)
Vibration arms	0.108	1.000	1.000	3.5 (0.017)
Vibration legs	0.012	0.548	<0.001	5.4 (<0.001)
Joint position arms	0.043	–	–	14.0 (0.043)
Joint position legs	<0.001	–	–	13.4 (<0.001)
2-point discrimination	<0.001	–	–	9.5 (<0.001)

QLQ-CIPN20 score up to ~ 25 , while score 2 was the most probable choice for patients with EORTC QLQ-CIPN20 score higher than 25. TNSc[®] PIN score 1 had very similar probability to be assigned as score 0, along a wide range of EORTC QLQ-CIPN20 scores and TNSc[®] PIN score 3/4 had always a lower probability to be assigned than TNSc[®] PIN score 2, even for very high values of EORTC QLQ-CIPN20 score. Virtually none of the TNSc[®] PIN grades reached a probability higher than 40% of being selected along the entire range of EORTC QLQ-CIPN20 score. The TNSc[®] VIBRATION had a better performance than TNSc[®] PIN, since patients had a higher probability to be scored 0 for EORTC QLQ-CIPN20 scores lower than 35, as 2 for EORTC QLQ-CIPN20 scores approximately between 35 and 50 and as 3 or 4 for EORTC QLQ-CIPN20 scores higher than 50, with a probability higher than 50% for scores higher than 75. The probability for TNSc[®] VIBRATION to be scored as 1 was low and distributed along the entire EORTC QLQ-CIPN20 score range.

comparison between the mISS and the EORTC QLQ-CIPN20 results

The EORTC QLQ-CIPN20 sensory submodule distribution by scores for each individual mISS item is reported in Figure 2.

When regressing the EORTC QLQ-CIPN20 sensory score over each mISS item (Table 1), a positive slope significantly different from zero was estimated, meaning that worst patient perception of their status corresponded to the worst physician evaluations. Pairwise comparisons between scores showed that for the mISS pinprick arms item only the difference between 0 and 1 was statistically significant, while 1, 2 and 3/4 were attributed to patients with more similar EORTC QLQ-CIPN20 sensory

score. The light touch arms and light touch legs items appeared to be very similar on EORTC QLQ-CIPN20 sensory score, except for very severe conditions (i.e., 3/4), that statistically differed from 2. The mISS pinprick legs and vibration legs items presented statistically different mean values of EORTC QLQ-CIPN20 score only when comparing normal versus mISS pinprick score 1 and mISS pinprick score 2 versus very severe conditions. The vibration arms item scores appeared not statistically different. Patients with abnormal (i.e., score 1) joint position arms and legs and two-point discrimination results presented significantly higher EORTC QLQ-CIPN20 sensory score compared with normal (score 0).

The estimated probabilities of each patient to be assigned to different mISS item score in relationship with his self-evaluated EORTC QLQ-CIPN20 is shown in Figure 2. For CIPN20 sensory scores up to 60, patients had a higher probability to be classified as 0 in the pin arms item, while for scores >60 the favorite was 2, not 3/4, also given the small number in this class. Similar results (but with different cutoffs) were observed for the pinprick legs (cutoff ≈ 25) and light touch arms (cutoff ≈ 40) items. When analyzing the light touch legs item, the most probable score was 2, along all the range of EORTC QLQ-CIPN20 sensory scores. Only for extreme scores (i.e., <15 and >85), the most probable mISS score were 0 and 3/4. The vibration arms assignment was only slightly influenced by the EORTC QLQ-CIPN20 sensory score, because the most probable vibration score for each EORTC QLQ-CIPN20 sensory score was 0. In contrast, vibration legs item was clearly more associated with EORTC QLQ-CIPN20 sensory score, because for scores up to 30 the most probable score was 0, for scores between 30 and 65 the most probable vibration leg score was 2 and for scores >65 vibration leg score 3/4 were the favorite assignment.

Figure 2. EORTC QLQ-CIPN20 sensory submodule results distribution by each mISS item (see first three upper lines of panels; in brackets is reported the number of patients available for comparison). The box represents the first and third quartile, the central line the median, the white square the mean and the whiskers are located at the maximum and minimum observation (outside observations indicated with dots are those out of the $1.5 \times$ interquartile range). Probability for a patient, with a given EORTC QLQ-CIPN20 sensory score, to be classified as score 0 to 4 of each mISS item, is reported in the last three lines of panels.

When assessing the items with only two possible scores (Table 1, Figure 2), the joint position arms item was unrelated to EORTC QLQ-CIPN20 sensory score (score 0 was always the score with the highest probability to occur). The joint position legs item presented a better probability curve, since for very high EORTC QLQ-CIPN20 sensory scores, joint position leg score 1 had higher probability to occur. Two-point discrimination at the index finger appeared to have the best capacity to reflect the simple presence or absence of CIPN as measured with the EORTC QLQ-CIPN20 sensory scores since for values lower than 30, patients had a higher probability of being classified as 0, while the others were likely classified as 1 for two-point discrimination.

discussion

Currently, it is unknown if physician assessments reflect patients' perception of CIPN impact on their QoL [10–12]. The aim of our study was not to demonstrate the superiority of a given method, but rather to provide evidence for a critical selection of assessment tools based on the knowledge of their qualities and capacities.

We focused on sensory impairment because it has the major impact on QoL; thus, we selected as clinical tools the NCI-CTC sensory scale, clinical sensory items of the TNSc© and mISS. NCI-CTC has significant limitation in assessing CIPN, particularly in discriminating grades 2 and 3 [4]. The TNSc© allows a more accurate assessment, while the mISS is time-consuming and complicated for this use. However, the mISS is one of the most accurate and valid neurological scales to grade sensory polyneuropathies and is, therefore, a valuable 'benchmark' for our study. Among different PRO measures, EORTC QLQ-C30 questionnaire has been selected for its superior efficacy in reporting the QoL of cancer patients in comparison to other similar PRO instruments [16]. Accordingly, EORTC QLQ-CIPN20-specific submodule has already been used in clinical trials and is a valid tool in sensory CIPN detection [8, 17–19].

None of NCI-CTC, mISS and TNSc© sensory items had a perfect relationship with patients' perception and most of discrepancies stood in intermediate grades of CIPN severity.

Comparing NCI-CTC or TNSc© results with EORTC QLQ-CIPN20 scores, the former highly agreed with the PRO measure, except for patients with normal results. This is probably due to the role of the subjective reporting of sensory impairment by patients in NCI-CTC grading. However, formal assessment of vibration threshold according to TNSc© was able to discriminate at least three different levels of CIPN severity, with the best results in differentiating score 0 versus score 1 and score 2 versus score 3/4.

When PRO results were compared with items of mISS, both vibration and pinprick assessments in legs were able to detect significant differences in patient-reported CIPN severity when comparing scores 0 versus 1 and 2 versus 3/4. However, this discriminating capacity was not equivalent, since severe pinprick perception had a high probability to correlate only with very high EORTC CIPN scores (i.e., >90), while increasing impairment in vibration perception identified two distinct EORTC QLQ-CIPN20 thresholds (~30 and 65).

Interestingly, a valuable discriminating power was demonstrated by two-point discrimination at index finger (normal up to an EORTC QLQ-CIPN20 score <30), while impairment in joint

position perception in legs could only identify patients with an EORTC QLQ-CIPN20 score higher than 80. Negative result obtained with joint position arms item is likely due to the remarkably low number of patients with abnormal results (3.6% of the entire population), while the incidence in legs is higher—as expected—in a length-dependent polyneuropathy such as CIPN.

Our data strongly suggest that to achieve a comprehensive measure of CIPN, clinical and PRO measures should be used and indicate that the combined use of NCI-CTC and of EORTC QLQ-CIPN20 questionnaire is very likely to provide overlapping data, while only the formal neurological assessment can discriminate among different types and extent of nerve damage. These suggestions would be even more relevant in the era of personalized medicine and pharmacogenomics [20]. In fact, even the most refined and advanced tool would be powerless if not supported by a valid assessment to define the study population.

acknowledgements

The participation and availability of the subjects who agreed to participate in the study are gratefully acknowledged.

funding

This work was supported in part by a CI-PeriNomS research grant from the University of Milan Bicocca Rectoral Office and an unrestricted research grant from the 'Fondazione Banca del Monte di Lombardia' (GC), by a research grant from the 'Fondazione per la Ricerca Farmacologica Gianni Benzi' (BF), by a research grant (PI070493) from Istituto de Salud Carlos III (JB) and by a National Institute of Health/National Institute of Nursing Research grant [P30 NR011396] (SGD).

disclosure

The authors have declared no conflicts of interest.

references

1. Grisold W, Cavaletti G, Windebank AJ. Peripheral neuropathies from chemotherapeutics and targeted agents: diagnosis, treatment, and prevention. *Neuro Oncol* 2012; 14(Suppl 4): iv45–iv54.
2. Frigeni B, Piatti M, Lanzani F et al. Chemotherapy-induced peripheral neurotoxicity can be misdiagnosed by the National Cancer Institute Common Toxicity scale. *J Peripher Nerv Syst* 2011; 16: 228–236.
3. Cavaletti G, Frigeni B, Lanzani F et al. Chemotherapy-Induced Peripheral Neurotoxicity assessment: a critical revision of the currently available tools. *Eur J Cancer* 2010; 46: 479–494.
4. Cavaletti G, Frigeni B, Lanzani F et al. The Total Neuropathy Score as an assessment tool for grading the course of chemotherapy-induced peripheral neurotoxicity: comparison with the National Cancer Institute-Common Toxicity Scale. *J Peripher Nerv Syst* 2007; 12: 210–215.
5. Cavaletti G, Bogliun G, Marzorati L et al. Grading of chemotherapy-induced peripheral neurotoxicity using the Total Neuropathy Scale. *Neurology* 2003; 61: 1297–1300.
6. Cornblath DR, Chaudhry V, Carter K et al. Total neuropathy score: validation and reliability study. *Neurology* 1999; 53: 1660–1664.
7. Merkies IS, Schmitz PI. Getting closer to patients: the INCAT Overall Disability Sum Score relates better to patients' own clinical judgement in immune-mediated polyneuropathies. *J Neurol Neurosurg Psychiatry* 2006; 77: 970–972.

8. Cavaletti G, Cornblath DR, Merkies IS et al. The chemotherapy-induced peripheral neuropathy outcome measures standardization study: from consensus to the first validity and reliability findings. *Ann Oncol* 2013; 24: 454–462.
9. Aaronson NK, Ahmedzai S, Bergman B 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: 365–376.
10. Hershman DL, Weimer LH, Wang A et al. Association between patient reported outcomes and quantitative sensory tests for measuring long-term neurotoxicity in breast cancer survivors treated with adjuvant paclitaxel chemotherapy. *Breast Cancer Res Treat* 2011; 125: 767–774.
11. Bennett BK, Park SB, Lin CS et al. Impact of oxaliplatin-induced neuropathy: a patient perspective. *Support Care Cancer* 2012; 20: 2959–2967.
12. Inoue N, Ishida H, Sano M et al. Discrepancy between the NCI-CTCAE and DEB-NTC scales in the evaluation of oxaliplatin-related neurotoxicity in patients with metastatic colorectal cancer. *Int J Clin Oncol* 2012; 17: 341–347.
13. Martina IS, van Koningsveld R, Schmitz PI et al. Measuring vibration threshold with a graduated tuning fork in normal aging and in patients with polyneuropathy. European Inflammatory Neuropathy Cause and Treatment (INCAT) group. *J Neurol Neurosurg Psychiatry* 1998; 65: 743–747.
14. van Nes SI, Faber CG, Hamers RM et al. Revising two-point discrimination assessment in normal aging and in patients with polyneuropathies. *J Neurol Neurosurg Psychiatry* 2008; 79: 832–834.
15. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977; 33: 159–174.
16. Luckett T, King MT, Butow PN et al. Choosing between the EORTC QLQ-C30 and FACT-G for measuring health-related quality of life in cancer clinical research: issues, evidence and recommendations. *Ann Oncol* 2011; 22: 2179–2190.
17. Postma TJ, Aaronson NK, Heimans JJ et al. The development of an EORTC quality of life questionnaire to assess chemotherapy-induced peripheral neuropathy: the QLQ-CIPN20. *Eur J Cancer* 2005; 41: 1135–1139.
18. Wolf SL, Barton DL, Qin R et al. The relationship between numbness, tingling, and shooting/burning pain in patients with chemotherapy-induced peripheral neuropathy (CIPN) as measured by the EORTC QLQ-CIPN20 instrument, N06CA. *Support Care Cancer* 2011; 20: 625–632.
19. Lavoie Smith EM, Barton DL, Qin R et al. Assessing patient-reported peripheral neuropathy: the reliability and validity of the European Organization for Research and Treatment of Cancer QLQ-CIPN20 Questionnaire. *Qual Life Res* 2013; [Epub ahead of print].
20. Postma TJ. Prevention of chemotherapy-induced peripheral neuropathy: a matter of personalized treatment? *Ann Oncol* 2013; 24: 1424–1426.

appendix 1 The CI-PeriNomS study group

Steering Committee

- G. Cavaletti, Department of Surgery and Translational Medicine, University of Milano-Bicocca, Monza, Italy;
- D.R. Cornblath, Department of Neurology, Johns Hopkins University School of Medicine, Baltimore, MD, USA;
- I.S.J. Merkies, Department of Neurology, Spaarne Hospital, Hoofddorp & Maastricht University Medical Center, Maastricht, The Netherlands;
- T.J. Postma, Department of Neurology, VU University Medical Center, Amsterdam, The Netherlands.

Statistical analysis

- M.G Valsecchi, S. Galimberti, E. Rossi, Clinical Epidemiology and Biostatistic Research Center, Department of Health Sciences, University of Milano-Bicocca, Monza, Italy

Participating centers and researchers

- Department of Surgery and Translational Medicine, University of Milano-Bicocca and Department of Oncology, S. Gerardo Hospital, Monza, Italy (Cavaletti G, Frigeni B, Lanzani F, Mattavelli L, Piatti ML, Alberti P, Binda D, Bidoli P., Cazzaniga M, Cortinovis D)
- Unit of Neuro-Oncology, Department of Neurology, University Hospital of Bellvitge-ICO Duran I Reynals, L'Hospitalet, Spain. (Bruna J, Velasco R)
- Division of Clinical Oncology-Department of Medicine, University Hospital of Patras, Patras, Greece (Argyriou AA, Kalofonos HP)
- Service de Neurologie Mazarin, Hôpital Pitié-Salpêtrière, Paris, France (Psimaras D)
- Service de Neurologie, Hôpital du Val-de-Grâce, Paris, France (Ricard D)
- Neurology Unit, National Cancer Institute Regina Elena, Rome, Italy (Pace A, Galiè E)
- Department of Neurological, Psychiatric, Sensorial, Reconstructive and Rehabilitative Sciences, University of Padova, Padova, Italy (Briani C, Lucchetta M, Campagnolo M, Dalla Torre C)
- Department of Neurology, Spaarne Hospital, Hoofddorp (Merkies ISJ)
- Department of Neurology, Maastricht University Medical Center, Maastricht, The Netherlands (Faber CG, Merkies ISJ, Vanhoutte EK, Bakkers M, Brouwer B)
- Division of Medical Oncology, Department of Internal Medicine, GROW-School of Oncology and Developmental Biology, Maastricht University Medical Centre, The Netherlands (Lalisang RI)
- Department of Neuro-oncology, Netherlands Cancer Institute, Amsterdam, the Netherlands (Boogerd W, Brandsma D)
- Department of Neurology, University of Essen, Germany (Koeppen S)
- West German Cancer Center, University of Essen, Germany (Hense J)
- Edinburgh Centre for Neuro-Oncology and Edinburgh Cancer Research Centre, Western General Hospital, Edinburgh, UK (Grant R, Storey D, Kerrigan S)
- Department of Neurosciences, Ophthalmology and Genetic, Center of Excellence for Biomedical Research, University of Genova, Genova, Italy (Schenone A, Reni L, Piras B)
- Department of Neurosciences, Ophthalmology and Genetic, University of Genova, Genova, Italy (Fabbri S)
- Department of Neurosciences Cattolica University, Rome, Don C. Gnocchi Foundation, Italy (Padua L, Granata G)
- Interuniversity Centre for Pain Neurophysiology at the University of Genova (Leandri M, Ghignotti I)
- I.R.C.C.S. of Neurological Sciences, Bellaria Hospital, Bologna, Italy (Plasmati R., Pastorelli F)
- Department of Neurology, VU University Medical Center, Amsterdam, the Netherlands (Postma TJ, Heimans JJ).
- Department of Neurology, Spaarne Hospital, Hoofddorp, the Netherlands (Eurelings M, Meijer RJ)
- Department of Neurology, Kaiser Franz Josef Hospital, Vienna, Austria (Grisold W,

- Lindeck Pozza E)
- Department of Neurosciences, AOU "G. Martino", Messina, Italy (Mazzeo A, Toscano A)
- Department of Medical Oncology, University of Messina, Messina, Italy (Tomasello C, Altavilla G)
- Servicio de Neurologia, Hospital Universitario Doce de Octubre, Madrid, Spain (Penas Prado M, Dominguez Gonzalez C)
- University of Maryland School of Nursing and Marlene and Stewart Greenebaum Cancer Center, Baltimore, MD, USA (Dorsey SG)
- Division of Cancer Prevention, National Cancer Institute, Rockville, MD, USA (Brell JM)

appendix 2 THE WHOLE CI-PeriNomS Study: focus on

AIMs

Selecting outcome measures for CIPN evaluation and establishing their validity and reproducibility in a cross-sectional multicenter study.

PATIENT SELECTION

Inclusion Criteria

- Signed and dated informed consent form before study entry.
- 18 years or older.
- Karnofsky performance score ≥ 70 .
- The presence of CIPN that evolved after standard chemotherapy defined as having typical symptoms, signs, and/or test results that were not present before chemotherapy.
- Stable clinical condition defined as either an unchanged clinical functionality as declared by the subject over 1 month before the study or no clear objective changes at neurological examination by the researcher when compared with recorded findings over 2 months before study entry.

EXCLUSION CRITERIA

- Active underlying malignancy and poor prognosis.
- Chemotherapy planned during the study period.
- Concomitant diseases, e.g. diabetes, renal insufficiency, alcohol abuse (>5 IU/day).
- Concomitant neurologic conditions that would complicate interpretation.
- Treatment with antiepileptic drugs, antidepressants and major analgesics, unless stable dosing and conditions have been reached.
- The presence of peripheral nerve damage due to another illness or medication.
- Currently potentially neurotoxic medication use.
- Any other condition, which, in the investigator's judgment might prevent to achieve the study objectives.

OUTCOME MEASURES APPLIED

- at the impairment level

- National Cancer Institute-Common Toxicity Criteria (NCI-CTC) v3, sensory and motor neuropathy score.
 - Clinical version of the Total Neuropathy Score clinical version (TNSc). Vibration sense was assessed using the graduated Rydel-Seiffer. Pinprick was assessed using the calibrated Neuropen with its disposable tips (Owens Mumford, Woodstock) for both the TNSc and the mISS.
 - mISS that grades the presence and severity of a sensory deficit. It assesses pinprick, vibration, light touch, and joint position in the limbs and static two-point discrimination at the index finger, in a predefined manner.
 - Light touch using the 10 g monofilament and vibration sense with the RS tuning fork tests were also evaluated separately.
 - NCS in the sural and common peroneal nerves (unilaterally, nondominant side) were measured with the standard methods of each center and compared with its own reference values.
 - Visual analog pain scale (VAS) and the 11-point pain-intensity numerical rating scale (PI-NRS) [34] were assessed to determine the amount of pain perceived by the patients, who were explicitly asked to score the pain considered to be CIPN related.
 - at the quality-of-life level
- The disease-specific European Organization for Research and Treatment of Cancer (EORTC) QLQ-C30 and CIPN20 questionnaires.

STUDY PROCEDURES. Patients were evaluated twice to verify selected outcome measures reproducibility. At visit 1, the medical/oncologic history was collected. Subsequently, two investigators examined each patient separately and completed the TNSc and the NCI-CTC-v3 subscales and mISS independently and consecutively within 2 hours (interobserver measures). Subjects were also requested to complete the VAS, PI-NRS, EORTC QLQ-C30, and QLQ-CIPN20 in a random order. Within 2–3 weeks, subjects returned for visit 2 and both investigators re-examined each subject (for intra-observer comparison) without having access to the previous data. The patient reported outcome measures were also completed by each subject for a second time (test-retest study). NCS were carried out once at visit 1 in those centers participating in the extended study.

STUDY POPULATION

The complete study population consisted of 281 subjects (males = 146, females = 135; median age = 63.9; range = 29–85). Colorectal, breast, ovarian, non-small cell lung cancer and multiple myeloma accounted for 82.2% of the total malignancies. Accordingly, 56.9% of patients were treated exclusively with platinum drugs, 13.2% with taxanes, 3.9% with vincristine, 3.2% with thalidomide, 2.8% with bortezomib, while 20.0% of patients were treated with a combination of two or more neurotoxic drugs.