Received: 3 April 2025 | Revised: 22 December 2025 | Accepted: 20 January 2026

DOI: 10.1002/ijc.70364

INTERNATIONAL JOURNAL of CANCER | UICC

RESEARCH ARTICLE

Innovative Tools and Methods

Toward a validated European instrument to measure the socioeconomic impact of cancer

Jasper Ubels[1] | Julie M. Vancoppenolle[2] | Josephine Tetteh[1] | Wim H. van Harten[2,3] | Valesca P. Retèl[2,4] | Michael Schlander[1,5] | Nora Franzen[4]

[1]Division of Health Economics, German Cancer Research Center (DKFZ), Heidelberg, Germany

[2]Department of Psychosocial Research and Epidemiology (PSOE), Netherlands Cancer Institute (NKI), Amsterdam, The Netherlands

[3]Department of Health Technology and Services Research (HTSR), University of Twente, Enschede, The Netherlands

[4]Erasmus School of Health Policy & Management (ESHPM), Erasmus University Rotterdam, Rotterdam, The Netherlands

[5]Institute for Innovation & Valuation in Health Care (InnoValHC), Wiesbaden, Germany

Correspondence

Valesca P. Retèl, Netherlands Cancer Institute (NKI), Amsterdam, The Netherlands.
Email: v.retel@nki.nl

Funding information

European Fair Pricing Network (EFPN)

Abstract

A cancer diagnosis can impose a financial burden on patients and their families, defined as socio-economic impact (SEI) within a framework of the Organization of European Cancer Institutes (OECI). The Socio-Economic Consequences of Cancer (SEC) study assessed SEI in 25 European countries using the Canadian Financial Index of Toxicity (FIT) instrument, showing substantial variation and supporting the need for a validated Europe-specific instrument. We examined the FIT instrument's validity and reliability in a secondary analysis of the SEC study, exploring whether the SEI framework supported its validation. Factor analyses were performed on the largest subgroup sharing cancer type, language, and country. The aim was to test whether the Canadian model could be replicated or a SEI-based model fit better. Reliability and construct validity were analyzed, followed by configural invariance and Differential Item Functioning (DIF) analysis for cross-country comparability. We used data from Bulgaria, France, Germany, the Netherlands, Norway, and Spain. The original FIT-instrument failed to replicate in the Spanish sample, leading to an SEI-based model with better fit (CFI = 0.975, RMSEA (90% CI) = 0.104 (0–0.278), $\chi^2 = 18$, $p = .60$). The instrument was reliable. Construct validity was partly confirmed. Configural invariance testing suggested that the SEI-based model's factor structure fits better in Europe, while DIF was identified, implying that direct score comparisons across countries should be done with care. In conclusion, the original FIT-instrument could not be fully validated in Europe whereas the SEI-framework improved score interpretation, supporting its use in developing a validated instrument tailored to the European context.

KEYWORDS

cancer, Europe, financial toxicity, instrument validation, socioeconomic impact

Abbreviations: CFA, confirmatory factor analysis; CFI, comparative fit index; CI, confidence interval; DIF, differential item functioning; EFA, exploratory factor analysis; EQ5D, EuroQol 5 Dimensions; FIT, Financial Index of Toxicity; MIMIC, multiple indicators multiple causes; OECI, Organization of European Cancer Institutes; QALY, quality-adjusted life year; RMSEA, root mean square error of approximation; SEC-study, socio-economic consequences of cancer study; SEI, socio-economic impact; SRMR, standardized root mean square residual; TLI, Turker-Lewis Index; WLSMV, weighted least squares mean and variance adjusted; χ^2, chi-square test of model fit.

Shared first authorship by Jasper Ubels and Julie M. Vancoppenolle.

Shared last authorship by Michael Schlander and Nora Franzen.

What's New?

The Financial Index of Toxicity (FIT) evaluates patients' experiences of financial strain after cancer diagnosis. Originally developed in Canada for head and neck cancer patients, the applicability of FIT to other populations remains uncertain. This study assessed the validity of FIT in Europe using patient-level data from six countries and the socioeconomic impact (SEI) framework of the Organization of European Cancer Institutes to support validation. Analyses show that the original FIT model has limited replicability, due in part to narrow response options for certain assessment items. The SEI-based framework was a better fit, supporting further development for European populations.

1 | INTRODUCTION

A cancer diagnosis often affects individuals and their household members in multiple dimensions beyond health. One adverse consequence following a cancer diagnosis is financial distress, which may arise from various sources, including out-of-pocket expenditures[1,2] or loss of income due to reduced working hours.[3,4] Existing research has primarily provided insights into aspects of financial impact, with some studies focusing on direct health-related out-of-pocket expenditures and others on the psychological impact of finance-related problems. A comprehensive approach is needed to integrate these often complementary aspects and to more effectively focus support programs for cancer patients and survivors.[5]

The health economics taskforce of the Organization of European Cancer Institutes (OECI) recently published consensus-based recommendations to facilitate comprehensive research on the socioeconomic impact (SEI) of cancer, defining SEI as the financial consequence of a cancer diagnosis. More precisely, SEI should be understood as "(…) the changes of the social and economic components of their quality of life due to cancer, attributable to out-of-pocket expenses, loss of income, or both."[5] Additionally, the taskforce provided definitions of underlying concepts related to SEI and pointed out the need for a validated pan-European instrument to measure SEI across countries.

Following the same logic, the Socio-Economic consequences of Cancer study (SEC-study) measured the financial impact of cancer on 2507 patients and survivors across 25 European countries. Next to objective and validated measures such as the EQ5D and self-reported household income, the SEC study used the Financial Index of Toxicity (FIT) to capture individuals' subjective experience of the financial impact following a cancer diagnosis.[6] The results showed that despite publicly funded healthcare systems in Europe, a considerable number of participants in all included countries reported financial toxicity and expressed a need for better financial support.[7]

In their limitations, the authors of the SEC study's noted that the FIT, originally developed and validated for Canadian patients with head and neck cancer,[6] requires further validation in the European setting.[7] Establishing construct validity is essential to ensure the instrument accurately captures the intended concepts. For example, if a tool designed to assess subjective financial distress does not correlate with actual changes in a household's financial situation, it raises questions about what is truly being measured. Moreover, differences in language and healthcare systems across Europe can influence how patients interpret and respond to questionnaires, potentially affecting comparability.[8] A lack of a validated, Europe-specific instrument leads to several disadvantages: it increases measurement error, hinders meaningful cross-country comparison, and limits the use of SEI data in health policy and planning. Validating the FIT in a Pan-European context would help ensure conceptual consistency and support a more accurate and comparable understanding of SEI. These efforts are reinforced by the recommendations from the OECI taskforce and by systematic reviews evaluating instruments used to assess the SEI.[8–10]

Therefore, the objective of the present analysis is to investigate the validity of the FIT-instrument in Europe with data from the SEC-study and to explore how the SEI conceptual model could support instrument validation and development.

2 | MATERIALS AND METHODS

In this secondary analysis of the SEC study, we tested the validity of the FIT-instrument in a European context. To assess whether the FIT provides valid and comparable results across European countries, we followed three steps:

1. *Measurement model evaluation*: We replicated the FIT's original measurement model ("original-FIT model") and tested an alternative model based on the SEI framework of the OECI health economic taskforce ("SEI-based model").
2. *Construct validity*: We assessed whether the FIT's dimensions measure what they are intended to.
3. *Cross-country measurement equivalence*: We examined whether the FIT performs equivalently across different countries.

To reduce heterogeneity in the data, steps one and two were limited to breast cancer patients from a single country and language group. In step three, we used a cross-country sample to test for differential item functioning (DIF) and examine whether patients from different European countries interpreted and responded to the FIT-instrument in a comparable way.

2.1 | Study population

This study represents a distinct secondary analysis of data from the original SEC study.[7] The SEC study collected data from 2507 patients in 25 European countries from September 2021 to September 2022. Patients were eligible if they were over 18, lived in the European region, and were undergoing or had recently completed cancer treatment (within 2 years). Recruitment was done via convenience sampling, resulting in a large but heterogeneous sample. A detailed description is available in the original article[7] and Supplementary Materials S.1 and S.2. Another secondary analysis of the SEC study investigated the experiences of young adults with cancer, a vulnerable group.[11]

For this manuscript, we focused on breast cancer patients to reduce heterogeneity. We further stratified participants by language and country, treating groups like Dutch speakers in the Netherlands and Belgium separately. Also, Spanish and Catalan speakers from Spain were treated as different groups. The largest group, which was Spanish speaking breast cancer patients from Spain, was used for the main validation analyses (steps 1 and 2). Groups with more than 100 participants were included in the cross-country comparison (step 3). This cut-off value was based on methodological recommendations for DIF and multiple-indicator, multiple-cause (MIMIC) models (described later in this section).[12]

2.2 | SEC study questionnaire and the Financial Index of Toxicity (FIT) instrument

The SEC study used a 41 item questionnaire, including the FIT-instrument, the EQ-5D-5L, and additional questions regarding socio-demographic/treatment characteristics, coping behavior, and personal finances, such as pre-diagnosis household income, out-of-pocket costs, and income loss. The full SEC-study questionnaire is presented in Supplementary Material S.2.

The FIT instrument includes nine items in three subscales:

- *Financial stress* (three items): ability to pay for daily food expenses, daily housing expenses, and medication related to treatment.
- *Financial strain* (four items): satisfaction with one's financial situation, worry about finances, perceived financial standing compared to other families, and the need to borrow money from family, friends, or financial institutions to cover treatment-related expenses.
- *Lost productivity* (two items): whether the respondent had to quit their job due to the cancer diagnosis or treatment, and whether family or friends had to stop working permanently to provide care.

The number of response options differs per item. Three items are yes/no questions; others use four- or five-point Likert scales. Responses are scored from 0 to 100 for each subscale and for a total score, with 0 indicating no financial toxicity and 100 indicating severe impact. All items are weighted equally in the scoring algorithm.[6]

The FIT was originally developed and validated in Canada[6] and was for the SEC-study translated into 16 languages using forward-backward translation. Translations were reviewed by national SEC collaborators.[7]

2.3 | The socioeconomic impact (SEI) framework

We also used the conceptual SEI framework developed by the OECI to test an alternative model ("SEI-based model"). The SEI framework outlines 25 categories of cancer-related burden, including direct medical costs, out-of-pocket expenditures, loss of income, and the broader effects on caregivers and dependents (SEI model). The framework provides a comprehensive and structured approach to identify and measure the multifaceted impacts of cancer beyond clinical outcomes.[5] Mapping between the FIT and SEI categories is provided in Supplementary Material S.2.

2.4 | Analyses

We analyzed the validity of the FIT instrument in three steps (e.g., measurement model evaluation, construct validity, and cross-country measurement equivalence). In the following, we detail the methods used in each analysis.

2.4.1 | Measurement model evaluation

We evaluated the dimensional structure of the FIT instrument to assess whether the originally proposed grouping of all nine items into three subscales (financial stress, financial strain, and lost productivity) could be statistically confirmed. To do so, we applied confirmatory factor analysis (CFA) to test whether the original three-factor model fit the data from our European sample. We also used exploratory factor analysis (EFA) to investigate whether some items might be associated with more than one subscale (known as cross-loading) or whether alternative groupings of items provided a better fit. This included testing a structure informed by the OECI's SEI model (referred to as the "SEI-based model").

Model fit was assessed using standard criteria[13,14]: Chi-square goodness of fit statistics (p >.05); comparative fit index (CFI, >0.95); Turker-Lewis Index (TLI, >0.95); root mean squared error of approximation (RMSEA, <0.08) and its 90% confidence interval; and standardized root mean square residual (SRMR <0.08).

The FIT-instrument items had different response lengths, affecting whether they could be treated as continuous or ordinal variables, and thus influencing the appropriate analysis methods.[14] Since three of the items had dichotomous response options, we used tetrachoric and polychoric correlations with a robust diagonally weighted least squares estimator to analyze the ordinal data.[14] Oblique rotation was used to allow for correlation between factors in CFA and EFA.

In the EFA, items were considered to load meaningfully onto a factor if the loading was ≥0.32. Loadings below this threshold suggest that changes in the item are not associated with changes in the dimension of interest, meaning that the item is unsuitable for scoring that specific subscale. Cross-loadings (items loading ≥0.32 on multiple dimensions) suggest that a response to an item cannot solely be linked to changes in the dimension of interest. Not accounting for this might result in unreliable or invalid subscale scores and problems interpreting the total score because it is unclear which dimension drives results.

If CFA of the original-FIT model showed poor fit, or EFA showed cross-loadings or weak item loadings, we tested the alternative SEI-based model derived from the OECI framework. Theoretical mapping of each FIT item onto the SEI domains was used to define the SEI-based structure (see Supplementary Material S.2).

We used the largest sample group, Spanish speaking breast cancer patients from Spain ($n = 187$), and split it in training ($n = 93$) and validation ($n = 94$) subsets using the SOLOMON method to balance background characteristics of groups.[15] This step prevents imbalanced subsamples, where for example, one subsample predominantly consists of individuals without financial problems, and another subsample consists of individuals with strong financial problems. The training sample was used to compare the original and SEI-based models using CFA and EFA; the best-fitting structure was then validated in the test sample using CFA. Key characteristics of the data set are presented in Supplementary Material S.3.

Finally, we conducted an additional CFA with the SEI-based model during the cross-country analysis to test its fit in a multi-country context. The SEI-based model consistently outperformed the original FIT model in terms of fit indices and interpretability. Therefore, all subsequent analyses (construct validity, reliability, and DIF) used the SEI-based factor structure.

2.4.2 | Construct validity, reliability, and floor and ceiling effects of the FIT-instrument

We assessed construct validity using the full sample of Spanish-speaking breast cancer patients from Spain. For this, we tested four hypotheses, based on expected relationships between FIT scores and other indicators of financial burden or health status, using the appropriate correlation for the variable type:

- *Household income*: The FIT-instrument total score and subscales are *negatively* correlated with pre-diagnosis household income (i.e., higher household income is associated with lower socioeconomic impact). We tested this hypothesis using Spearman correlation.
- *Out-of-pocket expenses*: The FIT-instrument total score and subscales are *positively* correlated with out-of-pocket expenditures (i.e., higher out-of-pocket expenditures result in higher socioeconomic impact). We tested this hypothesis using Pearson correlation.
- *Utility score*: The FIT-instrument total score and subscales are *negatively* correlated with the EQ-5D-5L utility score (i.e., lower preference adjusted health-related quality of life is associated with higher socioeconomic impact). We tested this hypothesis using Biserial correlation.
- *Income loss*: The FIT-instrument total score and subscales are *positively* correlated with income loss (i.e., higher income loss is associated with higher socio-economic impact). We tested this hypothesis using Biserial correlation.

A p-value <0.05 was used to establish whether an association is statistically significant. Internal reliability was assessed using Cronbach's alpha, with $\alpha \geq 0.7$ considered acceptable.[16] Floor and ceiling effects were also examined, which describe whether many respondents score at the lowest or highest possible value. Such patterns could limit an instrument's sensitivity to detect variation. For example, if two individuals both receive the maximum FIT score, it is not possible to determine who is more severely affected, as their experiences likely fall outside of the scale of the instrument.

2.4.3 | Cross-country measurement equivalence

To assess whether the FIT instrument functions equivalently across different European countries, we conducted a DIF analysis. DIF analysis examines whether response patterns to specific survey items vary systematically between countries or groups, even after accounting for differences in the underlying construct, in this case the socioeconomic impact of participants. If, even after controlling for SEI, individuals from different countries respond differently to the same items, it may indicate that the item is interpreted or understood differently, which could compromise the validity of cross-country comparisons.

DIF was assessed using multiple steps. First, we tested configural invariance, using the same fit indices as for the CFA,[17] to evaluate whether the dimensional structure of the SEI-based model was stable across countries. Configural invariance was tested to verify that items clustered into the same factors across countries, providing a necessary foundation for subsequent DIF analysis.

Next, we used MIMIC models to test for uniform and non-uniform DIF.[18] Uniform DIF occurs when individuals from different countries respond differently to an item in a consistent way across all levels of SEI. Non-uniform DIF occurs when these differences vary depending on the respondent's level of SEI, indicating an interaction.[18]

In each MIMIC model, direct paths from countries (the grouping variables) to specific items were analyzed to test for DIF, that is whether item responses were influenced by country. To ensure that any observed DIF was not due to true differences in SEI across countries, we also included paths from the grouping variables (countries) to the latent variables (i.e., the subscales of the FIT). Ideally, controlling for SEI should eliminate any significant associations between countries and item responses. If not, this suggests that items function differently across countries.

To detect non-uniform DIF, we added interaction terms between country and SEI dimensions (subscales) and modeled paths from these interactions to specific items. This allowed us to assess whether country effects on item responses varied across SEI levels. For example, when SEI is moderate, responses across countries may appear similar, but when SEI is high, individuals from one country may respond more extremely than those from another. This process was repeated iteratively for each item.

A constrained baseline approach was applied to identify DIF.[12] First, model fit was compared between a base model without paths from countries to specific items, for example, where item responses could be explained solely by SEI. Second, model fit was compared between alternative models with direct paths from all the countries to individual items. A significantly better fit in the latter (based on a chi-square difference test for nested models) indicated the presence of uniform DIF. Significant interaction terms indicated non-uniform DIF.

To ensure comparability, the Spanish sample was used as a reference group. All countries and interactions were coded as dummies. We did not adjust for multiple testing to capture any potential DIF across items. Since DIF was tested simultaneously with all countries, a significant improvement in model fit indicated that DIF was present, but not between which countries specifically. Therefore, a post-hoc test was conducted to identify patterns of DIF for each item, with a significant association ($p < .05$) indicating DIF presence.

To examine whether the factor structure of the FIT instrument was comparable across countries, we tested for configural invariance using DWLS estimators with polychoric correlations. Since robust maximum likelihood estimators and continuous variables are better studied in the context of MIMIC models, all MIMIC models were estimated using this method, treating variables as continuous.

All analyses were conducted in R 4.4.1,[19] with the packages Psych version 2.4.6[20] and Lavaan.[21] The anonymized data set is provided in the Supplementary Material S.6.

3 | RESULTS

Out the 25 countries of the SEC-study, six countries (Bulgaria, France, Germany, The Netherlands, Norway, and Spain) had a sample of participants that was larger than 100 according to the inclusion criteria of our study. We extensively tested the validity of the FIT-instrument using the largest sample, which was from Spain. Table 1 presents the characteristics of the country samples used in our analysis.

3.1 | Measurement model evaluation

CFA of the original FIT- model showed good model fit in the training dataset (Chi-square $p = 0.526$, CFI $= 1.00$, TLI $= 1.065$, RMSEA $= 0.000$, SRMR $= 0.051$). However, the EFA with three dimensions (factors) indicated that item 8 loaded on a different factor than in the original Canadian FIT-instrument development paper. In other words, it was statistically more closely related to a different subscale than originally intended (see Table 2). As such, item 8 provided information about a different dimension in the Spanish sample compared to the original Canadian sample. Due to this misalignment, an alternatively dimensional structure based on the SEI-model was considered.

Considering the mapping of items according to the SEI-based model and the factor loadings of the exploratory factor analysis, we grouped FIT items 1 to 3 (*Satisfaction/worrying about finances and comparison with others*) under one dimension and FIT items 4 to 6 (*Ability to pay for living expenses*) under another. Items 7, 8, and 9 of the FIT were kept as separate single exogenous indicators and were not linked to any specific dimensions. These items were structured to correlate with both dimensions and with each other. Supplementary Material S.4 presents the factor loadings (association between items and dimensions), factor correlations (association between dimensions), and factor-single item correlations (correlation between dimensions and items 7, 8, and 9) of the SEI-based model with the validation subsample.

Comparing the model fit between the SEI-based model and the original FIT-instrument model using the validation subsample showed no (significant) difference in RMSEA or chi-square test. Although the point estimate of the RMSEA was high, the wide confidence interval indicates that good fit cannot be rejected for either model. However, the SEI-based model showed a better fit in terms of model fit statistics: CFI, TLI, and SRMR (see Table 3). Based on these findings, we proceeded with the SEI-based model for assessments of construct validity, floor and ceiling effects, and DIF.

3.1.1 | Construct validity, reliability, and floor and ceiling effects of the FIT-instrument

Table 4 shows the results of the construct validity tests for the different hypotheses. The FIT total score as well as the two subscales from the SEI-model constructed with items 1 to 3 (*Satisfaction/worrying about finances and comparison with others*) and items 4 to 6 (*Ability to pay for living expenses*) were significantly correlated with QALY scores and income loss as hypothesized. Correlations between yearly pre-diagnosis household income and out of pocket expenses were generally not identified, except for a significant correlation between the subscale constructed with FIT items 4 to 6 (*Ability to pay for living expenses*) and yearly household income. The Cronbach's alpha of the full scale was 0.83 in the Spanish sample, which indicates a good level of reliability. The floor and ceiling effects of the FIT and its subscales are presented in Table 5.

3.2 | Cross-country measurement equivalence

Tests for DIF were conducted with data from Bulgaria ($n = 165$), France ($n = 131$), Germany ($n = 128$), the Netherlands ($n = 113$), Norway ($n = 102$), and Spain ($n = 187$). First, we further compared the fit of the SEI-based model with the FIT- model in the test for

TABLE 1 Sample characteristics of the six included countries.

	Bulgarian speakers from Bulgaria	French speakers from France	German speakers from Germany	Dutch speakers from the Netherlands	Norwegian speakers from Norway	Spanish speakers from Spain
Total	165	131	128	113	102	187
Marital status						
Divorced	29 (17.58%)	17 (12.98%)	17 (13.28%)	9 (7.96%)	10 (9.80%)	26 (13.90%)
Married or partnered	122 (73.94%)	90 (68.70%)	85 (66.41%)	87 (76.99%)	72 (70.59%)	118 (63.10%)
Single	10 (6.06%)	23 (17.56%)	24 (18.75%)	12 (10.62%)	19 (18.63%)	35 (18.72%)
Widowed	4 (2.42%)	1 (0.76%)	2 (1.56%)	5 (4.42%)	1 (0.98%)	8 (4.28%)
Education						
Doctoral studies	5 (3.03%)	11 (8.4%)	5 (3.91%)	8 (7.08%)	3 (2.94%)	7 (3.74%)
Bachelor/master	106 (64.24%)	68 (51.91%)	45 (35.16%)	48 (42.48%)	65 (63.73%)	73 (39.04%)
Vocatinal education	23 (13.94%)	29 (22.14%)	61 (47.66%)	37 (32.74%)	16 (15.69%)	53 (28.34%)
High school	28 (16.97%)	20 (15.27%)	12 (9.38%)	19 (16.81%)	16 (15.69%)	30 (16.04%)
Primary school	3 (1.82%)	2 (1.53%)	0 (0.00%)	0 (0.00%)	2 (1.96%)	24 (12.83%)
Other	0 (0.00%)	0 (0.00%)	5 (3.91%)	1 (0.88%)	0 (0.00%)	0 (0.00%)
Metastasized						
Metastasized	40 (24.24%)	42 (32.06%)	27 (21.09%)	41 (36.28%)	24 (23.53%)	50 (26.74%)
Not metastasized	125 (75.76%)	88: (67.18%)	101 (78.91%)	72 (63.72%)	78 (76.47%)	132 (70.59%)
Unknown	0 (0.00%)	1 (0.76%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	5 (2.67%)
Age	Mean: 48.56 SD: 8.7	Mean: 49.34 SD: 11.96	Mean: 55.98 SD: 10.43	Mean: 53.73 SD: 11.78	Mean: 47.87 SD: 9.48	Mean: 50.27 SD: 9.41
Occupation						
Employed	102 (80.31%)	89 (72.95%)	99 (79.84%)	66 (60%)	83 (88.3%)	108 (63.91%)
Self-employed	17 (13.39%)	15 (12.30%)	7 (5.65%)	27 (24.55%)	4 (4.26%)	20 (11.83%)
Not employed, looking	1 (0.79%)	4 (3.28%)	4 (3.23%)	3 (2.73%)	1 (1.06%)	16 (9.47%)
Not employed, not looking	1 (0.79%)	2 (1.64%)	1 (0.81%)	4 (3.64%)	0 (0.00%)	6 (3.55%)
Student	1 (0.79%)	0 (0.00%)	1 (0.81%)	1 (0.91%)	1 (1.06%)	2 (1.18%)
Retired	5 (3.94%)	12 (9.84%)	11 (8.87%)	6 (5.45%)	1 (1.06%)	10 (5.92%)
Surgery						
Yes	147 (89.09%)	114 (87.02%)	6 (4.69%)	86 (76.11%)	90 (88.24%)	153 (81.82%)
No	18 (10.91%)	17 (12.98%)	122 (95.31%)	27 (23.89%)	12 (11.76%)	34 (18.18%)
Chemotherapy						
Yes	133 (80.61%)	97 (74.05%)	9 (7.03%)	79 (69.91%)	93 (91.18%)	159 (85.03%)
No	32 (19.39%)	34 (25.95%)	119 (92.97%)	34 (30.09%)	9 (8.82%)	28 (14.97%)

(Continues)

TABLE 1 (Continued)

	Bulgarian speakers from Bulgaria	French speakers from France	German speakers from Germany	Dutch speakers from the Netherlands	Norwegian speakers from Norway	Spanish speakers from Spain
Hormonal therapy						
Yes	129 (78.18%)	51 (38.93%)	55 (42.97%)	62 (54.87%)	66 (64.71%)	93 (49.73%)
No	36 (21.82%)	80 (61.07%)	73 (57.03%)	51 (45.13%)	36 (35.29%)	94 (50.27%)
Immunotherapy						
Yes	6 (3.64%)	16 (12.21%)	9 (7.03%)	28 (24.78%)	12 (11.76%)	26 (13.9%)
No	159 (96.36%)	115 (87.79%)	119 (92.97%)	85 (75.22%)	90 (88.24%)	161 (86.1%)
Targeted therapy						
Yes	38 (23.03%)	37 (28.24%)	22 (17.19%)	16 (14.16%)	17 (16.67%)	0 (0.00%)
No	127 (76.97%)	94 (71.76%)	106 (82.81%)	97 (85.84%)	85 (83.33%)	187 (100%)
Radiation therapy						
Yes	120 (72.73%)	104 (79.39%)	3 (2.34%)	83 (73.45%)	89 (87.25%)	137 (73.26%)
No	45 (27.27%)	27 (20.61%)	125 (97.66%)	30 (26.55%)	13 (12.75%)	50 (26.74%)
Other treatments						
Yes	1 (0.61%)	2 (1.53%)	57 (44.53%)	5 (4.42%)	8 (7.84%)	7 (3.74%)
No	164 (99.39%)	129 (98.47%)	71 (55.47%)	108 (95.58%)	94 (92.16%)	180 (96.26%)

TABLE 2 Results of exploratory factor analysis with three factors with the training dataset.

FIT-instrument items	Factor 1	Factor 2	Factor 3
1. Satisfaction with finances	**0.91**	−0.03	−0.02
2. Worry about finances	**0.90**	−0.04	−0.05
3. Comparison of finances with others	**0.61**	0.11	0.09
4. Ability to pay daily living expenses: food	−0.05	**1.01**	0.00
5. Ability to pay daily living expenses: housing	0.17	**0.81**	−0.08
6. Ability to pay daily living expenses: medical expenses	−0.02	**0.91**	0.04
7. Borrow money because of cancer treatment	**0.71**	0.15	0.16
8. Quit job because of cancer	**0.53**	0.03	0.04
9. Caregiver quit job because of cancer	0.01	−0.01	**1.00**
Inter-factor correlations			
Factor 1	-	-	-
Factor 2	0.65	-	-
Factor 3	0.33	0.22	-

Note: Factor loadings >0.32 are highlighted in bold. Items 1, 2, 3, 7, and 8 load on factor 1. Items 4, 5, and 6 load on factor 2. Item 9 loads on factor 3.

TABLE 3 Fit index values of the two measurement models in the validation subsample.

Model	Chi-square (df)	*p*-value	RMSEA (90% CI)	CFI	TLI	SRMR
SEI-based model	17.82 (20)	.60	0.104 (0.000–0.278)	0.975	0.955	0.040
Original FIT-instrument model	19.49 (24)	.73	0.141 (0.000–0.272)	0.945	0.917	0.047

Note: CFI, comparative fit index; df, degrees of freedom; RMSEA, root mean square error of approximation; SRMR, standardized root mean square residual; TLI, Tucker-Lewis Index.

configural invariance. Results showed that the FIT- instrument model could not be replicated due to problems in how individuals responded to item 9. This item assesses whether caregivers quit their job because of cancer, with response options "yes" or "no." No participants from Germany reported that caregivers quit their job. Furthermore, only five participants from France, four from the Netherlands, and three from Norway indicated that caregivers quit their job. Due to this lack of variance in the responses to item 9, the third factor (dimension) of the FIT- model (consisting of items 8 and 9) could not be estimated in these countries, given that these models are based on the covariance of different items.

We therefore focused on establishing configural invariance with items 1 to 7 of the FIT-instrument according to the FIT-model and items 1 to 6 for the SEI-model. The difference between these two models is that the original FIT-model has an additional item on the dimension "financial strain" which thus consists of items 1, 2, 3, and 7. In the SEI-model, item 7 is excluded in the analysis.

A comparison of model fit between the two models indicated that SEI-model (Chi-square $p = .15$, CFI = 0.980, TLI = 0.963, RMSEA (90%CI) = 0.10 (0.027–0.158), SRMR = 0.037) showed acceptable FIT. The original FIT-model showed inadequate fit (Chi-square $p = .000$, CFI = 0.940, TLI = 0.903, RMSEA (90%CI) = 0.154 (0.110–0.197), SRMR = 0.055). Based on these findings, we rejected configural invariance for the original FIT-model and therefore decided to test for DIF with SEI-model.

Table 6 shows the results of the measurement invariance test. DIF was identified in items 1, 2, 3, and 5. Items 4 and 6 did not exhibit DIF. More detailed information about the pairwise occurrence of uniform and nonuniform DIF between different countries is presented in Supplementary Material S.5.

4 | DISCUSSION

The objective of this study was to evaluate the validity and reliability of the FIT-instrument in European breast cancer patients and to explore how the SEI-based model using the recommendations of the health economic taskforce of the OECI could support instrument validation. We first examined whether findings from the original Canadian development study were reproducible and explored how the SEI framework could be used to group items into constructs relevant for European populations, starting with Spain. Additional tests related to construct validity and reliability were also conducted in this sample. We then examined whether these findings could be generalized to other European countries (Bulgaria, France, Germany, the Netherlands, and Norway), comparing the SEI-based model to the original model of the FIT-instrument. Overall, the FIT instrument was only partially validated, while applying the SEI framework appeared to offer advantages for instrument development and validation.

TABLE 4 Construct validity of the FIT-instrument in Spain.

Hypotheses	Type of correlation	*p*-value
1. Yearly household income	Spearman	
FIT total score	0.13	.897
FIT subscale items 1–3: Satisfaction/ worrying about finances and comparison with others	0.08	.327
FIT subscale items 4–6: Ability to pay for living expenses	0.21	.008
2. Out-of-pocket expenses	Biserial	
FIT total score	0.09	.240
FIT subscale items 1–3: Satisfaction/ worrying about finances and comparison with others	0.00	.990
FIT subscale items 4–6: Ability to pay for living expenses	0.08	.278
3. Utility score	Pearson	
FIT total score	−0.44	<.001
FIT subscale items 1–3: Satisfaction/ worrying about finances and comparison with others	−0.44	<.001
FIT subscale items 4–6: Ability to pay for living expenses	−0.18	.030
4. Income loss	Biserial	
FIT total score	0.42	<.001
FIT subscale items 1–3: Satisfaction/ worrying about finances and comparison with others	0.40	<.001
FIT subscale items 4–6: Ability to pay for living expenses	0.30	<.001

TABLE 5 Floor and ceiling effects of the FIT-instrument and its subscales.

	Floor	Ceiling
FIT total score	1.60%	0.00%
FIT subscale items 1–3	1.60%	14.97%
FIT subscale items 4–6: Ability to pay for living expenses	62.57%	1.60%

Note: Floor effects refer to the percentage of participants who scored the minimum possible score. Ceiling effects refer to the percentage of participants who scored the maximum possible score.

4.1 | Replication of the FIT model and configural invariance

The first notable finding is that the original FIT model showed only partly replication in the Spanish sample of breast cancer patients. Although in the training dataset the initial CFA indicated excellent model fit, the EFA revealed that Item 8 did not load onto its original factor. This seemingly contradictory result may be explained by the relatively small Spanish sample size and the use of

TABLE 6 Differential item functioning—results of Chi-square test including all 6 countries and using Spain as a reference.

FIT items	Chi-square test results
1. Satisfaction with finances	$p = .0003$
2. Worry about finances	$p < .0001$
3. Comparison of finances with others	$p < .0001$
4. Ability to pay daily living expenses: food	$p = .6946$
5. Ability to pay daily living expenses: housing	$p = .0004$
6. Ability to pay daily living expenses: medical expenses	$p = .1802$
7. Borrow money because of cancer treatment	Not possible, not related to a construct
8. Quit job because of cancer	Not possible, not related to a construct
9. Caregiver quit job because of cancer	Not possible, not related to a construct and lack of variance

dichotomous items, both of which can make it more difficult to detect model misfit.

Additional statistical evidence against using the original FIT model was the lack of configural invariance, suggesting that European breast cancer patients interpret some items of the FIT instrument differently compared to the original Canadian head and neck cancer patients. One item (e.g., Item 9 Caregiver quit job because of cancer) showed a lack of variance in multiple samples.

Beyond this, there is a theoretical issue whether the dichotomous items of the FIT-instrument truly reflect "theoretical constructs" or merely observable occurrences–a consideration relevant for the development of factor models. Additionally, items with limited response options may not provide meaningful information, particularly when most respondents across countries did not experience caregivers having to quit their jobs (item 9). In the context of European-wide instruments, which must accommodate heterogeneous patient populations, items with a broader range of response options may be preferable to capture a wider spectrum of socioeconomic experiences.

4.2 | The SEI-based model and implications for instrument development

An alternative model was developed based on the SEI conceptual framework.[5] Tests for configural invariance showed that the SEI-based model adequately fit across six European countries, suggesting its potential for instrument validation and development.

Beyond evaluating the FIT-instrument, our study demonstrates that the SEI framework can provide a deeper understanding of how FIT-instrument items map onto key dimensions of socioeconomic impact. By aligning FIT-instrument items with SEI framework concepts, we identified separately measurable constructs, supporting the operationalization of the OECI Health Economics Taskforce recommendations into quantifiable domains. Furthermore, the establishment of configural invariance supports the cross-European applicability of these constructs across diverse settings. These results illustrate how the SEI framework of the OECI Health Economics Taskforce can be used to assess and develop patient-reported outcomes related to socioeconomic consequences that may arise after a cancer diagnosis.

4.3 | Construct validity and reliability

Additional tests of construct validity and reliability were conducted with the SEI-based model. The FIT-instrument demonstrated reliability in the Spanish subsample. However, in terms of construct validity results were mixed. Hypothesis 1 (Yearly household income) was only confirmed for one FIT-instrument subscale, while Hypothesis 2 (Out-of-pocket expenses) could not be confirmed at all. In contrast, the Canadian FIT-instrument development study found a significant correlation in the expected direction.

Rather than indicating insufficient construct validity, these differences may reflect variations in healthcare systems. In Spain, individuals with higher household income might have higher out-of-pocket costs, yet may be financially protected from the SEI of cancer due to their income level. Wagstaff, et al.[22] found that compared to Canada, individuals with lower household income in Spain spend a lower percentage of their budget on out-of-pocket expenditures. This could explain why no significant correlation was found for Hypothesis 2 (Out-of-pocket expenses). Further research is needed to understand how healthcare and broader social systems influence the experience of socioeconomic impact. Despite the lack of significance in Hypotheses 1 and 2, this does not invalidate the constructs measured by the FIT-instrument. Hypothesis 3 (Utility score) and Hypothesis 4 (Income loss) were both significant, indicating that quality of life (utility score) and income loss are associated with SEI after cancer diagnosis.

4.4 | DIF and cross-country comparability

DIF was identified in Item 1 (satisfaction with finances), Item 2 (worry about finances), Item 3 (comparison of finances with others), and Item 5 (ability to pay daily living expenses: housing) with Spain as the reference country. Post-hoc testing revealed complex patterns of uniform and non-uniform DIF across countries. With respect to uniform DIF, this means that one group consistently scores higher or lower on an item compared to another group, regardless of their financial burden. In contrast, non-uniform DIF means that the difference between groups changes depending on financial burden, for example, one group might respond more strongly to an item at certain levels of financial problems but not at others.

Still, some items did not exhibit DIF, suggesting that these could potentially be used as anchor items for further analysis of the effect of DIF on total test scores between specific countries. Such research would help to determine whether the DIF of single items leads to systematic differences in overall test scores across countries.

For items 7–9 (*Borrowing money/Patient/caregiver quit job because of cancer*), DIF could not be tested due to either limited variation in responses and/or because these items were not clearly associated with any of the specific dimensions measured in the model. Furthermore, beyond the issue of DIF, the lack of variation in some countries suggests that these items offer limited insight into the SEI experienced by breast cancer patients. As such, the comparability of the FIT instrument's scores across countries remains uncertain.

4.5 | Limitations

The RMSEA for the original FIT model could be rejected, since the lower value of the 90% CI was higher than the defined cut-off value. Still, our analyses showed wide confidence intervals around the RMSEA index, which, combined with the relatively low sample size, indicates that there may have been insufficient power to identify potential misfit in the measurement models. Further studies with higher power are required to further confirm the results presented in this study.

Additionally, in the DIF tests, we did not correct p-values for multiple testing, which may have led to an overestimation of DIF.

This manuscript presents a secondary analysis of the SEC study. The SEC study represents the most extensive collection of European data on the socioeconomic consequences of a cancer diagnosis across countries and cancer types to date. However, as recruitment relied on convenience sampling through hospitals and patient organizations, the representativeness of the sample may be limited, which also affects our study. This approach may have introduced selection bias, for instance by underrepresenting less engaged patients or those with limited access to patient networks. Therefore, the results require further validation in other subpopulations.[7] The FIT-instrument itself requires further refinement and validation for European use, addressing cut-off points, translations, and applicability to various cancer types and European nations. Additionally, the current analysis focuses on breast cancer patients. Although the alignment with the OECI SEI domains suggests wider applicability, further research is necessary to study to confirm the results of this study in other cancer populations as well.

Cultural and recall biases may also have influenced results, especially in regions with lower inclusion rates and greater hesitancy to disclose financial details, such as the United Kingdom. Addressing these factors in future studies could improve the generalizability of findings.

4.6 | Implications

One direct implication of this study concerns the calculation of subscale scores in the FIT-instrument. Among European breast cancer patients, two distinct constructs were identified as opposed to the three-factor structure in the original model. Additionally, one construct was measured with three instead of four items (i.e., one construct with items 1 to 3 (*Satisfaction/worrying about finances and comparison with others*)), a second construct with items 4 to 6 (*Ability to pay for living expenses*).

Following the methodology of the original FIT-instrument, we recommend scoring these subscales from 0 to 100. Furthermore, each item should contribute equally to the score, with adjustments based on the number of items per subscale and the number of response options, consistent with the original scoring system of the FIT-instrument.

Items 7 to 9 (*Borrowing money/Patient/caregiver quit job because of cancer*) can still be used for the calculation of the complete FIT score, following the original FIT-model. Although our findings suggest that not all items, such as item 9, are necessarily for assessing the SEI of breast cancer in some European countries, retaining this item still may facilitate international comparisons. Still, it should be noted that these items most likely yield little information in some populations, and that DIF for these items was not tested. It is therefore unclear in how far total scores that include these items can be compared internationally.

Overall, our results support the development of an SEI-based instrument to measure the SEI of cancer in the European context. Further studies with higher power are necessary to confirm the findings. Still, this approach holds promise for improving measurement accuracy and ensuring that instruments capture the financial burden of cancer across diverse healthcare systems.

AUTHOR CONTRIBUTIONS

Jasper Ubels: Conceptualization; methodology; validation; writing – original draft; data curation; formal analysis; writing – review and editing. **Julie M. Vancoppenolle:** Writing – original draft; methodology; formal analysis; validation; writing – review and editing; conceptualization. **Josephine Tetteh:** Formal analysis; validation; methodology; writing – review and editing. **Wim H. van Harten:** Conceptualization; methodology; project administration; writing – review and editing; funding acquisition; supervision. **Valesca P. Retèl:** Supervision; writing – review and editing; methodology; funding acquisition; conceptualization. **Michael Schlander:** Conceptualization; funding acquisition; validation; writing – review and editing; supervision. **Nora Franzen:** Conceptualization; writing – original draft; methodology; validation; writing – review and editing; supervision; project administration.

ACKNOWLEDGMENTS

We thank the European Fair Pricing Network (EFPN) and the Organization of European Cancer Institutes (OECI) for the support of this study.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflict of interest.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available on request from the corresponding author.

ETHICS STATEMENT

The study has received ethical approval from the board of the Netherlands Cancer Institute under the number IRBd24-050.

ORCID

Julie M. Vancoppenolle https://orcid.org/0000-0003-2193-8391
Nora Franzen https://orcid.org/0000-0002-4539-1709

REFERENCES

1. Schneider J, Hernandez D, CAESAR study group, Schlander M, Arndt V, CEASAR study group. Out-of-pocket payments and loss of income among long-term breast cancer survivors in Germany: a multi-regional population-based study. *J Cancer Surviv*. 2023;17(6):1639-1659.
2. Fitch MI, Sharp L, Hanly P, Longo CJ. Experiencing financial toxicity associated with cancer in publicly funded healthcare systems: a systematic review of qualitative studies. *J Cancer Surviv*. 2022;16(2):1-15.
3. Hernandez D, Schlander M. Income loss after a cancer diagnosis in Germany: an analysis based on the socio-economic panel survey. *Cancer Med*. 2021;10(11):3726-3740.
4. Richter-Ehrenstein C, Martinez-Pader J. Impact of breast cancer diagnosis and treatment on work-related life and financial factors. *Breast Care*. 2021;16(1):72-76.
5. Schlander M, van Harten W, Retèl VP, et al. The socioeconomic impact of cancer on patients and their relatives: organisation of European cancer institutes task force consensus recommendations on conceptual framework, taxonomy, and research directions. *Lancet Oncol*. 2024;25(4):e152-e163.
6. Hueniken K, Douglas CM, Jethwa AR, et al. Measuring financial toxicity incurred after treatment of head and neck cancer: development and validation of the financial index of toxicity questionnaire. *Cancer*. 2020;126(17):4042-4050.
7. Vancoppenolle J, Franzen N, Azarang L, et al. Financial toxicity and socio-economic impact of cancer in Europe. *ESMO Open*. 2025;10(6):105293.
8. Thomy L, Crichton M, Jones L, et al. Measures of financial toxicity in cancer survivors: a systematic review. *Support Care Cancer*. 2024;32(7):403.
9. Zhu Z, Xing W, Wen H, et al. Psychometric properties of self-reported financial toxicity measures in cancer survivors: a systematic review. *BMJ Open*. 2022;12(6):e057215.
10. Pham PD, Ubels J, Schlander M. Measuring the socioeconomic impact of cancer: a systematic review and quality appraisal of patient-reported outcome (PRO) instruments. *Pharmacoecon Open*. 2025;9(4):519-539.
11. Janssen S, Vancoppenolle J, Franzen N, et al. 1694MO socio-economic consequences among adolescent and young adult cancer patients: a European perspective. *Ann Oncol*. 2023;34:S927.
12. Kim ES, Yoon M, Lee T. Testing measurement invariance using MIMIC: likelihood ratio test with a critical value adjustment. *Educ Psychol Meas*. 2012;72(3):469-492.
13. Browne MW, Cudeck R. Alternative ways of assessing model fit. In: Bollen KA, Long JS, eds. *Testing Structural Equation Models*. Sage; 1993.
14. Hu L, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Struct Equ Model Multidiscip J*. 1999;6(1):1-55.
15. Lorenzo-Seva U. SOLOMON: a method for splitting a sample into equivalent subsamples in factor analysis. *Behav Res Methods*. 2022;54(6):2665-2677.
16. Cronbach LJ. Coefficient alpha and the internal structure of tests. *Psychometrika*. 1951;16(3):297-334.
17. Putnick DL, Bornstein MH. Measurement invariance conventions and reporting: the state of the art and future directions for psychological research. *Dev Rev*. 2016;41:71-90.

18. Chun S, Stark S, Kim ES, Chernyshenko OS. MIMIC methods for detecting DIF among multiple groups: exploring a new sequential-free baseline procedure. *Appl Psychol Measur*. 2016;40(7):486-499.
19. R Core Team. *R: A Language and Environment for Statistical Computing*. Foundation for Statistical Computing; 2024.
20. Revelle W. psych: Procedures for personality and psychological research. 2024.
21. Rosseel Y. Lavaan: an R package for structural equation modeling. *J Stat Softw*. 2012;48(2):1-36.
22. Wagstaff A, Eozenou P, Smitz M. Out-of-pocket expenditures on health: a global stocktake. *World Bank Res Obs*. 2020;35(2):123-157.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Ubels J, Vancoppenolle JM, Tetteh J, et al. Toward a validated European instrument to measure the socioeconomic impact of cancer. *Int J Cancer*. 2026;158(12):3288-3299. doi:10.1002/ijc.70364