Using the Nottingham Health Profile with Cancer Patients: Factorial Validity and Psychometric Properties of the Romanian Adapted Version

Claudiu C. Papasteri*[v], Ioana M. Neagoe*

***** Faculty of Psychology and Educational Sciences, Psychology Department, University of Bucharest, Bucharest, Romania

Abstract

Introduction: *The Nottingham Health Profile (NHP) has been portrayed as a multipurpose measure of health status, capable of being used in population surveys and in the evaluation of both medical and psychosocial interventions. In the fields of oncology and psycho-oncology, it has been widely used to assess health-related quality of life or disease activity and any changes that occur with regard to them.*

Objectives: *The aim of the present study was to adapt the NHP for the use with Romanian cancer patients and to test its psychometric properties and factorial validity.*

Methods: *The final Romanian adapted NHP was obtained thought translation and blind back-translation by a panel of bilingual experts. Afterwards, a mixed sample of 250 adult cancer patients ages 23-82 years old (M=54.38, SD=11.85) from the Bucharest Institute of Oncology completed both NHP and FACT-G along with 14 questions regarding demographics. FACT-G was used to measure the concurrent validity of NHP scales and to test its usefulness within multi-instrument assessment protocols.*

Results: *The findings show that NHP exhibits adequate internal consistencies, in strong agreement with the literature, as well as moderate inter-scale correlations. Second-order confirmatory factor analysis using WLSMV estimation strongly supported the fit of the original six-factor model. No additional changes in factor structure or items of the Romanian NHP were warranted by the results of this study.*

Conclusions: *The study has shown that the Romanian NHP questionnaire possesses adequate psychometric properties and can be used with cancer patients in both clinical and research settings. Further research on test-retest reliability and potentially on the weighting system is required to establish its complete equivalence to the original instrument.*

Keywords: nottingham health profile, nhp, validation, romanian adaptation, cancer

[v] Corresponding author: Claudiu C. Papasteri, Faculty of Psychology and Educational Sciences, University of Bucharest, Panduri 90, Sector 5, 050663, Bucharest, Romania. Email: claudiu.papasteri@gmail.com.

I. Introduction

The Nottingham Health Profile (NHP; Hunt et al. 1980; Hunt et al., 1981; Hunt et al., 1985) is a self-report instrument that measures perceived health status. The instrument was designed as a generic health status questionnaire for rapidly determining perceived physical, social and emotional health problems (Hunt et al., 1985). It has been portrayed as a multipurpose measure of health status, capable of being used in population surveys and in the evaluation of both medical and psychosocial interventions. Accordingly, it was used to assess disease activity, the impact of illness on patients and changes in health status over time, and also as an indicator of perceived distress (McDowell, 2006).

The questionnaire's items address the World Health Organization (WHO) definitions of disability and the resulting profile is most often viewed by specialists to measure health-related quality of life (hereafter HRQoL) (Wiklund, 1989; Yildiz et al., 2010).

The development of the NHP has started at Nottingham University's Public Health Department, back in 1975. Initially there were 2000 statements belonging to 700 patients who described the effects of deteriorated health, statements that covered psychological, social, physical and behavioral fields. The first key concepts of the instrument were selected later on and only 138 statements were left after the removal of redundant and ambiguous items. A series of studies conducted between 1976 and 1978 refined the number of remaining statements even further, after which the number of statements was reduced to 82 (Hunt et al., 1985). Participants with a wide range of medical conditions were selected in order to establish the instrument's future general efficacy.

After further testing, statements that fit the criteria needed in order to develop the NHP into a population survey tool were retained and this resulted in the 45 item questionnaire. The first version of the NHP was called the Nottingham Health Index and it contained only 33 questions (Martini & McDowell, 1976). The current revised version contains 38 items included in Part I and 7 optional items that comprise Part II, the latter being regarded as a brief indicator of handicap.

Part II was developed after Part I in order to assess how perceived health problems may affect daily living (Hunt, McEwen & McKenna, 1986). The original statements collected during the development stage of the NHP were reviewed to identify areas of "task performance" most often affected by ill health. After identifying the key domains relevant for Part II, interviews were conducted with patients attending a hospital out-patient clinic. Problems regarding wording and presentation were identified, and further interviews were conducted with out-patients and university employees. The wording of the items was revised by the authors making them more comprehendible and acceptable for average people with potentially limited education (Hunt, McEwen & McKenna, 1986).

Today, the NHP has been translated in 26 languages and is considered by some authors to be in the public domain (Martinez-Martin et al., 2011) although copyright is still owned by Galen Research. It is one of the most widely used instruments of its kind, especially within Europe, next to SF-36 (Short-Form-36 Health Survey), which is arguably very similar to the NHP (Prieto et al., 1997).

Items

The first part of the NHP (Part I) contains 38 items grouped into 6 subscales that assess the following general domains:

- Energy level consists of 3 items (1, 12, and 26 - e.g.: I'm tired all the time; everything is an effort).

- Pain consists of 8 items (2, 4, 8, 19, 24, 28, 36, and 38 - e.g.: I have pain at night; I have unbearable pain).

- Emotional reactions consists of 9 items (3, 6, 7, 16, 20, 23, 31, 32, and 37 - e.g.: Things are getting me down; I feel that life is not worth living).

- Sleep consists of 5 items (5, 13, 22, 29, and 33 - e.g.: I take tablets to help me sleep; I lie awake for most of the night).

- Social isolation consists of 5 items (9, 15, 21, 30, and 34 - e.g.: I feel there is nobody I am close to; I feel I am a burden to people).

- Physical abilities consists of 8 items (10, 11, 14, 17, 18, 25, 27, and 35 - e.g.: I can only walk about indoors; I find it hard to dress myself).

Part II provides a brief indicator of handicap and contains seven items that record the impact of ill health on daily life domains like occupation, work around the house, personal relationships, social life, sex life, hobbies, and holidays (i.e. vacations).

Dichotomous Yes/No responses are used throughout the whole questionnaire.

Administration

The NHP is designed to be administered to adults over 16 years of age who are affected by a specific medical condition. However, it can be utilized on general population as well. The NHP is usually thought to be a pen and paper type questionnaire, but it can also be computer administered, individually and/or collectively.

The instructions for respondents are the following: "Listed below are some problems people might have in their daily lives. Read the list carefully and put a tick in the box under Yes for any problem that applies to you at the moment. Tick the box under No for any problem that does not apply to you. Please answer every question. If you are not sure whether to answer Yes or No, tick whichever answer you think is truer at the moment."

The NHP was designed to be acceptable to and understood by a majority of people while having also a low respondent burden. Filling in the questionnaire usually takes about 5 minutes, and no additional information is needed for its completion (Ebrahim, Barer & Nouri). It should be also noted that researchers found the NHP to produce the lowest missing value rate out of four commonly used generic health status measures (Essink-Bot et al., 1997).

Part II of the NHP is independent and optional, and can be left out without affecting the results (Garcia & McCarthy, 2000). Part I is frequently used on its own.

Scoring

The items are to be scored dichotomously on a scale where YES = 1 point and NO = 0 points. The scoring for all 6 subscales included in Part I can be accomplished by using either one of three methods:

1. By summing all YES responses of every subscale. Although it is the simplest way, it has been noted that in the case of three samples tested, the composite scores calculated by this method and the ones calculated using the weight system (which will be presented third) have had an extremely high correlation (.98 and above) (Jenkinson, 1991). Thus, either scoring method is valid and usable.

2. All YES responses of every subscale are summed up. This sum is divided by the total number of items and then the result is multiplied by 100. The final score will range between 0 and 100, where 0 signifies perfect health and 100 is the most deteriorated state of health possible for one patient.

3. Every item has a different weight, depending on the severity of the symptom. The weights for the NHP were derived using Thurstone's method of paired comparisons from a sample of 215 members of the general public (McKenna, Hunt & McEwen, 1981). The sum total of the weighted scores is 100, with weights intended to reflect the perceived severity of a health state represented by the item from the point of view of the general public, rather than a specific patient population (Jenkinson, C., Fitzpatrick, R., & Argyle, 1988).

Thus, as a result of summing all weighted values of every item from the scales, the obtained values will be between 0 and 100, where 0 is the best state possible and 100 is the worst possible state. Some of the versions of the NHP that have been adapted and validated in different cultures have different weighting systems. Such is the example of the Swedish version (Hunt & Wiklund, 1987) or of the French version (Bucquet, Condon & Ritchie, 1990), however, most of other versions have kept the original weighting system. The weight values of all YES items of Part I (items 1-38) of the original English NHP can be found in the table 1.

Table 1. Item weights for NHP – Part I

Item	Weight	Item	Weight	Item	Weight
1.	39,20	*14.*	21,30	*27.*	11,20
2.	12,91	*15.*	19,36	*28.*	20,86
3.	10,47	*16.*	7,08	*29.*	16,10
4.	19,74	*17.*	10,79	*30.*	22,53
5.	22,37	*18.*	9,30	*31.*	13,95
6.	9,31	*19.*	11,22	*32.*	16,21
7.	7,22	*20.*	9,76	*33.*	21,70
8.	9,99	*21.*	20,13	*34.*	15,97
9.	22,01	*22.*	27,26	*35.*	12,69
10.	11,54	*23.*	13,99	*36.*	5,83
11	10,57	*24.*	8,96	*37.*	12,01
12.	36,80	*25.*	12,61	*38.*	10,49
13	12,57	*26.*	24,00		

Part II is scored by summing up all YES values. No weighting or special reporting method is needed.

II. Methods

1. Questionnaire

The NHP was both presented in the Romanian literature (e.g. Iustin, 2006) and used in academic research (e.g. Moldovanu, 2010). Although the Romanian translation of the NHP is also being listed in the Galen Research (www.galen-research.com), multiple translation versions seem to have been utilized in Romania thoughtout the years. Nevertheless, none of these translations where accompanied by any description of the adaptation process or reported a questionnaire validation process. Moreso, to our knowledge, the psychometric properties of these translated versions are either insufficient or completly lacking.

Thus, the NHP Romanian adaption presented in this study began with two independent translations of the original English instrument (Hunt et al., 1985; full version with scoring and interpretation can be found in McDowell, 2006) and one of the French adaptation (Bucquet and Condon, 1990a; 1990b). The method of comparing these three translation by a pannel of psychology experts was ment to control the translation transformations in order to arrive at an interlinguistic equivalent that would yield the most accurate and appropriate conceptual content in the target language.

Blind back-translations and comparisons with the original were employed in order to arrive at a preliminary final adapted version. Ultimately, 10 lay people, self-selected relatives of random hospitalized cancer patients, were recruited and presented with versions of the questionnaire containing most of the alternative expressions that originated during the translation stage. They were asked to select those that would best conform to their language usage and to make any of these modifications to the final adapted version that was handed to them last. This final version that was almost identical to the preliminary one was pilot tested on a small sample of out-patients and psychology undergraduate students and proved to be understandable and easy to complete.

Before the begining of the study, participants were informed in both oral and written form about the study‘s nature and objectives. Only the first part of the NHP was administered and no weighting system was used during scoring.

2. Sample

A mixed sample of 250 adult cancer patients ages 23-82 (M=54.38, SD=11.85) from the Bucharest Institute of Oncology completed both NHP and FACT-G along with 14 questions regarding demographics. Although 64% of the sample were women, age did not differ significantly between genders (Welch's t-test, $t(97.65)=1.04$, $p=.30$).

Of the sample, 76.4% came from Romanian urban areas and 74% had received high school education or above. The most prevalent types of cancer were breast cancer (45.6%), genital cancers (17.2%), lung cancer (11.6%) and colorectal cancer (9.2%).

III. Results

1. Reliability

Franks & Moffatt (2001) reported the following internal consistencies (Cronbach’s α) for the NHP – Part I: .63 for energy level, .81 for pain, .81 for emotional reactions, .76 for sleep, .65 for social isolation, and .80 for physical abilities. Similar results were found on different patient samples: .62, .82, .78, .77, .63, and .71 respectively (Essink-Bot et al., 1997). The internal consistencies measured by Cronbach’s α for the Romanian NHP follow closely the ones presented in the literature (table 3).

Due to alpha’s shortcomings falling under heavy scrutiny in numerous statistics articles, we decided to also report McDonald’s (1999) omega (ω_h), considered to be a more accurate approximation of a scale’s reliability (Zinbarg et al., 2005; Revelle & Zinbarg, 2009; Peters, 2014; Dunn, Baguley & Brunsden, 2014). ω_h and corresponding bootstrap 95% CIs were calculated using R according to the procedure suggested by Dunn, Baguley & Brunsden (2014). Both Cronbach’s α and McDonald’s ω indicate adequate internal consistency.

There is some indication in the literature that the NHP has a floor effect, whereby it is not sensitive to minor levels of disability and cannot distinguish between levels of good health (Franks & Moffatt, 2001; McDowell, 2006). Of course, often designing a particular instrument involves trade-offs between ceiling effects and floor effects, but the impact of such effects on assessment is determined by the characteristics of the population of interest. For our sample, the floor effects were a little larger as the ones reported by Franks & Moffatt (2001), particularly in the cases of social isolation (52%), pain (50.4%) and energy level (48.4%), and none of the participants rated themselves as being in the worst health by obtaining the highest possible score of the NHP (table 2).

This makes discrimination among subjects among the bottom end (healthiest) of the scale impossible. Also, a related consequence will be the observed skewness, with subscales like pain (1.53) and social isolation (1.45) presenting positively skewed

distributions of scores.

The four-week test-retest reliability reported by Hunt et al. (1985) ranged between .75-.88 for NHP – Part I and .55-.89 for Part II. Unfortunately, the test-retest reliability and sensitivity to change over time of the Romanian NHP could not be tested due to the cross-sectional character of the present study.

Table 2. Descriptive statistics of NHP scales

Scale	No. of Items	Score Range	N	M	SD	Floor (%)	Ceiling (%)	Skewness
Energy level	3	0-3	250	1.07	1.19	48.4	19.6	.56
Pain	8	0-8	250	1.60	2.24	50.4	2.8	1.53
Emotional reactions	9	0-9	250	2.46	2.71	34	1.6	.94
Sleep	5	0-5	250	1.78	1.56	28	4.4	.43
Social isolation	5	0-5	250	.90	1.22	52	1.6	1.45
Physical abilities	8	0-8	250	2.18	2.22	32.8	1.2	.78
NHP Global	38	0-38	250	9.98	8.86	12	0	.78

Table 3. Internal consistencies for NHP scales

Scale	No. of Items	N	Cronbach's α		McDonald's ω	
			α	95% CI	ω	95% CI
Energy level	3	250	.78	.67 - .90	.79	.73 - .83
Pain	8	250	.80	.74 - .86	.80	.64 - .87
Emotional reactions	9	250	.82	.76 - .87	.82	.70 - .86
Sleep	5	250	.70	.61 - .80	.72	.66 - .76
Social isolation	5	250	.63	.53 - .74	.64	.54 - .73
Physical abilities	8	250	.80	.74 - .86	.82	.79 - .85
NHP Global	38	250	.93	.91 - .95	.93	.91 - .94

2. Validity

Content validity seemed to be the most important consideration of the instrument's early development stage with items being based on patients' descriptions of their experience. After this initial stage patient consultation was combined with expert consultation, therefore enhancing the relevance of the questionnaire to patients and clinicians.

Although content validity of a measure is largely dependent on the concept being measured, the NHP covers a broad range of health-related functions that could be expected to be affected in many diverse chronic conditions, and therefore appears to have good content validity as a measure of general health status (Essink-Bot et al., 1997).

The NHP also appears to have good face validity, all items referring to an aspect of health.

Spearman correlations among domain scores reported by O'Brien, Buxton & Patterson (1993) ranged from .32 (sleep and social isolation) to .70 (pain and physical abilities), most of them ranging from .41 to 0.58. In another study (Franks & Moffatt, 2001), the highest Pearson correlations were between social isolation and emotional status (.60), sleep and bodily pain (.59), while least correlated were social isolation and bodily pain (.23), and social isolation and energy (.26). Internal correlations between the sub-scales and the global score were calculated for the Romanian NHP as a further measure of the instrument's validity (table 4).

Table 4. Inter-scale correlations (Spearman's rho)

	Energy	*Pain*	*Emo Re*	*Sleep*	*Soc Iz*	*Fiz Abil*	*Global*
Energy	-						
Pain	*.61****	-					
Emo Re	*.64****	*.50****	-				
Sleep	*.51****	*.42****	*.60****	-			
Soc Iz	*.54****	*.37****	*.64****	*.46****	-		
Fiz Abil	*.73****	*.73****	*.57****	*.47****	*.47****	-	
Global	*.83****	*.76****	*.84****	*.72****	*.68****	*.85****	-

*p <.05, **p <.01, ***p <.001.

The concurrent validity of the NHP was assessed through many related measures like the McGill Pain Questionnaire, General Health Questionnaire, and SF-36 to name a few (see McDowell, 2006; Essink-Bot et al., 1997). For the present study we have chosen the Functional Assessment of Cancer Therapy – General (FACT-G; Cella et al., 1993; for Romanian validation see Dégi 2011; 2013) because it is one of the most prevalently used HRQoL measures in oncology and its subscales measure similar psychosocial domains to the ones measured by NHP subscales.

However, we expected to find negative correlations between the two instruments because high scores on NHP correspond to poor QoL while high scores on FACT-G correspond to good QoL.

Sure enough, the FACT-G total score was found to be highly related to the NHP global score (-.78). More so, subscales that measure related constructs were also highly correlated. Physical abilities and energy level were strongly related to physical and functional well-being while emotional reactions presented the highest correlation towards emotional well-being (table 5). Social isolation was already expected to correlate more strongly with psychological/emotional well-being as indicated by Essink-Bot et al. (1999) because it relates to the ability to make contact with other people and can be accordingly considered to belong to a psychological rather than a social role domain. Indeed, it was more strongly related to emotional well-being (-.56) than to social well-being (-.55).

Table 5. Subscale and global score correlations (Spearman's rho) between NHP and FACT-G

	Physical Well-Being	Social/Family Well-Being	Emotional Well-Being	Functional Well-Being	FACT-G Total
Energy	-.69***	-.33***	-.54***	-.65***	-.69***
Pain	-.55***	-.21**	-.37***	-.47***	-.49***
Emo Re	-.65***	-.47***	-.70***	-.67***	-.72***
Sleep	-.43***	-.34***	-.39***	-.51***	-.51***
Soc Iz	-.47***	-.55***	-.56***	-.56***	-.65***
Fiz Abil	-.63***	-.29***	-.47***	-.60***	-.62***
Global	-.70***	-.45***	-.64***	-.74***	-.78***

*p <.05, **p <.01, ***p <.001.

3. Confirmatory factor analysis

The latent structure of the NHP questionnaire and its original theoretically conceptualized six-factor structure was evaluated using confirmatory factor analysis (CFA). To avoid complications associated with the factor analysis of dichotomous items, researchers usually form parcels by summing items from a common scale, the resulting parcel scores falling afterwards on a greater-than-dichotomous scale.

Additionally, the number of parameters to be estimated will be significantly reduced. Although item parcelling is the prevailing approach in analysing scales with many items, especially dichotomous ones, in CFA, it is almost always preferable to evaluate latent variable models based on item-level data (Marsh *et al.*, 2013).

In order to conduct CFA with binary indicators the use of a robust estimator was warranted instead of the popular maximum likelihood (ML) estimation that is still largely used in CFA irrespective of the nature of the data (Hirschfeld & von Brachel, 2014).

There seems to be growing consensus that the best approach to analysing binary indicators is the diagonal weighted least squares (DWLS). Its two robust estimators: weighted least squares- mean adjusted (WLSM) and weighted least squares - mean and variance adjusted (WLSMV), both provide accurate parameter estimates for categorical data (DiStefano & Morgan, 2014) and are also considered to be superior to the normal theory-based maximum likelihood (ML) approach when ordinal observed variables are being analysed.

The current CFA was conducted using the *cfa* function with WLSMV estimation in *lavaan* (v. 0.5-18; Rosseel, 2012; Rosseel *et al.*, 2015) package for R (v. 3.1.2; R Core Team, 2014) statistical environment.

When specifying the WLSMV in the estimator argument, *lavaan* uses diagonally weighted least squares (DWLS) to estimate the model parameters, but uses the full weight matrix to compute robust standard errors, as well as a mean- and variance-adjusted test statistic. The WLSMV is a robust estimator which does not assume the normality of the variables and provides the best option for modelling categorical or ordinal data (Brown, 2006). Furthermore, the estimator works well with sample sizes of 200 or more (Flora & Curran, 2004; Rhemtulla, Brosseau-Liard & Savalei, 2012).

In line with the recommendations of most structural equation modelling researchers (e.g. Brown, 2015), several fit indices were used to assess model fit. In the present CFA, the following goodness-of-fit indices were employed: Root Mean Square Error of Approximation (RMSEA; Steiger, 1990), Comparative Fit Index (CFI; Bentler, 1990), Tucker-Lewis Index (TLI; Tucker & Lewis, 1973), goodness of fit index (GFI; Jöreskog & Sörbom, 1981) and adjusted goodness of fit index (AGFI; Jöreskog & Sörbom, 1981). It is generally considered that an adequate model fit would yield a RMSEA<.05, CFI>.90, TLI>.90, GFI >.90 and AGFI >.90 (Schumacker & Lomax, 2010; West, Taylor & Wu, 2012; Brown, 2015).

The results of the CFA indicated that the original six-factor structure of the NHP questionnaire provided a good overall fit to the sample data: χ^2_{WLSMV} (659 df)=935, p<.001; RMSEA =.041, 90% CI for RMSEA .035-.047, p=.995; TLI = .954; CFI = .957; GFI=.948; AGFI=.941.

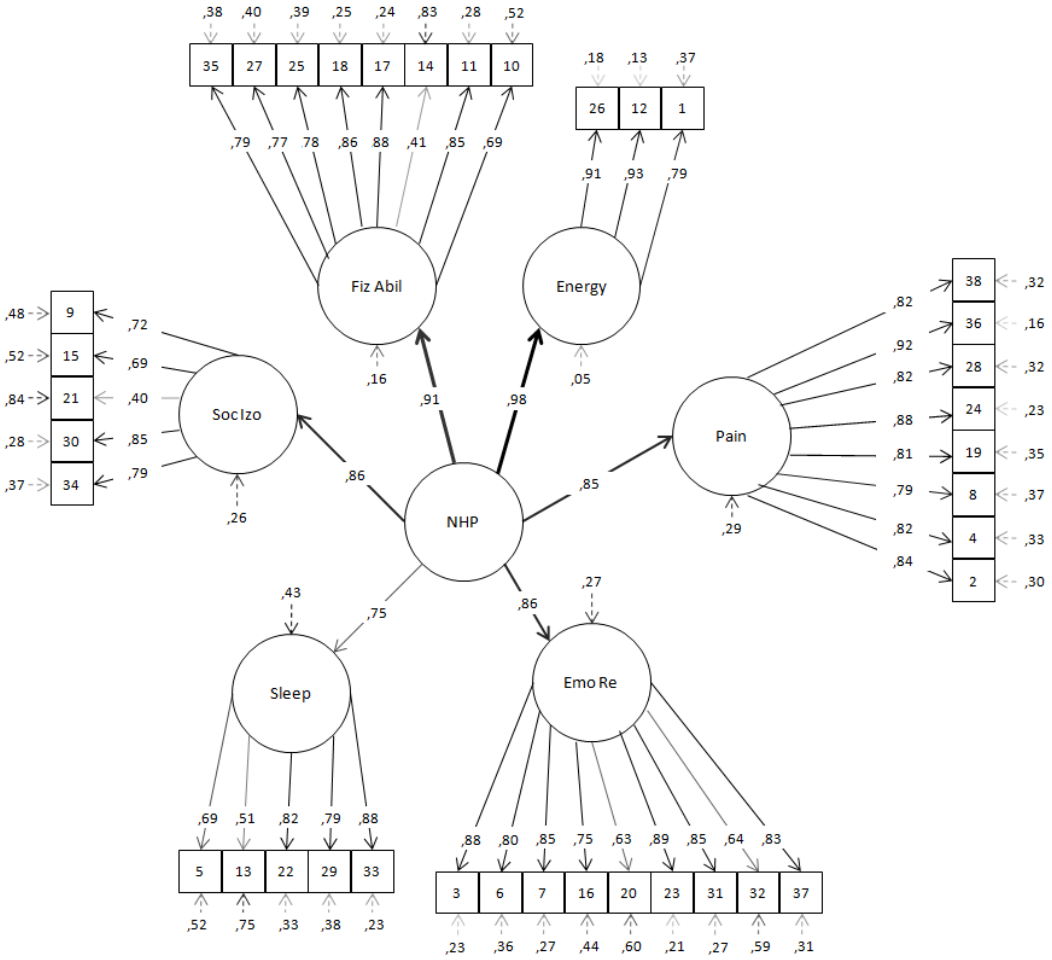

Fig 1. CFA model for the six-factor structure

IV. Conclusions

The NHP was designed to provide information about self-perceived health status and HRQoL. The purposes of this study were first to adapt a Romanian version of the NHP for use with Romanian cancer patients and then to assess its validity and psychometric properties.

Sartorius and Kuyken (1994) have described the translation of a health status instrument developed in one culture into the language of another without prior consideration of the appropriateness of the instrument in the target culture as problematic because of potential difficulties in achieving equivalent conceptual dimensions. Based on their experience in producing so many versions of health measures, Hunt and her group have proposed guidelines for the development of culturally equivalent versions of an instrument (Hunt et al., 1991).

Generally, many of the European adaptation (Swedish, French, Spanish, etc.) were conducted on the basis of this method. Conceptual equivalence rather than linguistic or semantic equivalence is the goal for the adapted versions of health measure. Having said this, the Romanian NHP adaptation was also based on the same methodology proposed by the original authors and followed most of the steps suggested by the guidelines: producing several translations, using multiple translators, producing several back-translations, using a panel review and pre-testing. Our method was informed mostly by the one used for the Spanish NHP adaptation (Alonso, Antó, & Moreno, 1990; Alonso, Prieto, & Antó, 1994).

The samples used in this study represented a good cross-section of cancer patients in Bucharest, although some cancer types were overrepresented. The heterogeneity of the sample (participants belonged to both sexes, to different age groups and to both urban and rural areas, had diverse education levels, and differing cancer locations and staging) aimed to ensure the external validity of the results. This attribute has been strengthened by the inclusion of patients pertaining to daily clinical practice. All patients in the study were users of the public health system, and there is no reason to believe that the patients included in this

study are not representative of the general population. Furthermore, the analysis of a large sample of cancer patients and the randomness in its selection also ensure the internal validity of the results.

The NHP is a valuable screening instrument in psycho-oncology and was extensively used especially before other more disease- or population- specific HRQoL measures. Of course, the original NHP was designed during the late 70s and new needs for an appropriate assessment of outcomes appeared since. These needs were more adequately addressed by disease-specific HRQOL measures that add subscales to tap common symptoms (like the FACT system that has many specific questionnaires).

Indeed, cancer-specific HRQOL measures are better able to detect changes and to discriminate between cancer types or stages. Regardless, all QoL are relative. The NHP was designed to compare patients across many disease categories and may lead to equivocal results when a single disease is studied (Fucini et al., 2008). Consequently when studying a single population or disease group one should chose the more specialized tool. However, when comparing cancer patients' QoL with that of the general population or that of other patient category or disease-free cancer survivors, a generic HRQoL like the NHP is more suitable. This is why we advocate for the use of both generic and disease-specific measures in multi-instrument screening protocols.

The present study has shown that the Romanian NHP questionnaire possesses adequate psychometric properties multiple measures of its validity and reliability approximating those of the original instrument. As such, we conclude that this culturally adapted version can be used with cancer patients in both clinical and research settings.

Further research on test-retest reliability and, potentially, on the weighting system is required to establish its complete equivalence to the original instrument. Nevertheless, the NHP appears to be a valid tool for screening HRQoL in Romanian cancer patients and allowing study results to be compared across different countries and between patients and general population. Future applications of the Romanian NHP on different patient groups would provide an important basis for definitively assessing validity and reliability of the profile.

References

Alonso, J., Antó, J.M., & Moreno, C.O.N.C.H.I. (1990). Spanish version of the Nottingham Health Profile: translation and preliminary validity. *American Journal of Public Health*, *80*(6), 704-708.

Alonso, J., Prieto, L., & Antó, J. M. (1994). The Spanish version of the Nottingham Health Profile: a review of adaptation and instrument characteristics. *Quality of life research*, *3*(6), 385-393.

Bentler, P.M. (1990). Comparative fit indexes in structural models. *Psychological bulletin*, *107*(2), 238.

Brown, T.A. (2006). *Confirmatory factor analysis for applied research.* New York: Guildford.

Brown, T.A. (2015). *Confirmatory factor analysis for applied research*. Guilford Publications.

Bucquet, D., & Condon, S. (1990a). Adaptation en français du Nottingham health profile (NHP) et caractéristiques opératoires de la version française. Montpellier: INSERM-CJF 88-12.

Bucquet, D., Condon, S., & Ritchie, K. (1990b). The French version of the Nottingham Health Profile. A comparison of items weights with those of the source version. *Social science & medicine*, *30*(7), 829-835.

Cella, D.F., Tulsky, D.S., Gray, G., Sarafian, B., Linn, E., Bonomi, A., ... & Brannon, J. (1993). The Functional Assessment of Cancer Therapy scale: development and validation of the general measure. *Journal of clinical oncology*, *11*(3), 570-579.

Dégi, L.C. (2011). Psychosocial aspects of tumoral diseases: assessment and intervention. Cluj-Napoca: Presa Universitară Clujeană.

Dégi, C.L. (2013). Psychosocial aspects of Cancer in Hospitalized Adult Patients in Romania. *Procedia-Social and Behavioral Sciences*, *82*, 32-38.

DiStefano, C., & Morgan, G.B. (2014). A Comparison of Diagonal Weighted Least Squares Robust Estimation Techniques for Ordinal Data. *Structural Equation Modeling: A Multidisciplinary Journal*, *21*(3), 425-438.

Dunn, T.J., Baguley, T., & Brunsden, V. (2014). From Alpha to Omega: A Practical Solution to the Pervasive Problem of Internal Consistency Estimation. *British Journal of Psychology*, *105*(3), 399-412.

Ebrahim, S., Barer, D., & Nouri, F. (1986). Use of the Nottingham Health Profile with patients after a stroke. *Journal of Epidemiology and Community Health*, *40*(2), 166-169.

Essink-Bot, M.L., Krabbe, P.F., Bonsel, G.J., & Aaronson, N.K. (1997). An Empirical Comparison of Four Generic Health Status Measures: the Nottingham Health Profile, the Medical Outcomes Study 36-item Short-Form Health Survey, the COOP/WONCA charts, and the EuroQol instrument. *Medical care*, *35*(5), 522-537.

Flora, D.B., & Curran, P.J. (2004). An Empirical Evaluation of Alternative Methods of Estimation for Confirmatory Factor Analysis with Ordinal Data. *Psychological methods*, *9*(4), 466.

Franks, P.J., & Moffatt, C. J. (2001). Health Related Quality of Life in Patients With Venous Ulceration: Use of the Nottingham Health Profile. *Quality of life research*, *10*(8), 693-700.

Fucini, C., Gattai, R., Urena, C., Bandettini, L., & Elbetti, C. (2008). Quality of Life Among Five-Year Survivors After Treatment for Very Low Rectal Cancer with or Without a Permanent Abdominal Stoma. *Annals of surgical oncology*, *15*(4), 1099-1106.

Garcia, P., & McCarthy, M. (2000). Measuring health: A step in the development of city health profiles. In *Measuring Health: a step in the development of city health profiles*. Denmark: WHO Regional Office for Europe.

Hirschfeld, G. & von Brachel, R. (2014). Multiple-Group Confirmatory Factor Analysis in R–A Tutorial in

Measurement Invariance with Continuous and Ordinal Indicators. *Practical Assessment, Research & Evaluation*, *19*(7), 2.

Hunt, S.M., & Wiklund, I. (1987). Cross-Cultural Variation in the Weighting of Health Statements: A Comparison of English and Swedish Valuations. *Health Policy*, *8*(2), 227-235.

Hunt, S.M., Alonso, J., Bucquet, D., Niero, M., Wiklund, I., & McKenna, S. (1991). Cross-Cultural Adaptation of Health Measures. *Health Policy*, *19*(1), 33-44.

Hunt, S.M., McEwen, J., & McKenna, S.P. (1985). Measuring Health Status: a New Tool for Clinicians and Epidemiologists. *The Journal of the Royal College of General Practitioners*, *35*(273), 185-188.

Hunt, S.M., McKenna, S.P., McEwen, J., Backett, E. M., Williams, J., & Papp, E. (1980). A Quantitative Approach to Perceived Health Status: A Validation Study. *Journal of epidemiology and community health*, *34*(4), 281-286.

Hunt, S.M., McKenna, S.P., McEwen, J., Williams, J., & Papp, E. (1981). The Nottingham Health Profile: Subjective Health Status and Medical Consultations. *Social Science & Medicine. Part A: Medical Psychology & Medical Sociology*, *15*(3), 221-229.

Hunt, S.M., McEwen, J. & McKenna, S.P. (1986). *Measuring health status*. London: Croom Helm.

Iustin, L. (2006). The Quality of Life in Health. Definitons and Assessment Instruments. *Quality of life, XVII*, 1–2.

Jenkinson, C. (1991). Why Are We Weighting? A Critical Examination of the Use of Item Weights in a Health Status Measure. *Social science & medicine*, *32*(12), 1413-1416.

Jenkinson, C., Fitzpatrick, R., & Argyle, M. (1988). The Nottingham Health Profile: An Analysis of its Sensitivity In Differentiating Illness Groups. *Social Science & Medicine*, *27*(12), 1411-1414.

Joly, F., Espie, M., Marty, M., Heron, J. F., & Henry-Amar, M. (2000). Long-Term Quality Of Life In Premenopausal Women With Node-Negative Localized Breast Cancer Treated with or Without Adjuvant Chemotherapy. *British journal of cancer*, *83*(5), 577.

Jöreskog, K.G., & Sörbom, D. (1981). *LISREL V: Analysis of linear structural relationships by maximum likelihood and least squares methods*. Chicago: International Educational Services.

Marsh, H.W., Lüdtke, O., Nagengast, B., Morin, A.J., & Von Davier, M. (2013). WHY Item Parcels Are (Almost) Never Appropriate: Two Wrongs Do Not Make A Right - Camouflaging Misspecification With Item Parcels In CFA Models. *Psychological methods*, *18*(3), 257.

Martinez-Martin, P., Jeukens-Visser, M., Lyons, K.E., Rodriguez-Blazquez, C., Selai, C., Siderowf, A., ... & Stebbins, G.T. (2011). Health-Related Quality-Of-Life Scales In Parkinson's Disease: Critique And Recommendations. *Movement Disorders*, *26*(13), 2371-2380.

Martini, C.J., & McDowell, I. (1976). Health status: Patient And Physician Judgments. *Health services research*, *11*(4), 508.

McDonald, R.P. (1999). *Test theory: A unified treatment*. Hillsdale: Erlbaum.

McDowell, I. (2006). *Measuring health: a guide to rating scales and questionnaires*. Oxford University Press.

McDowell, I.W., Martini, C.J., & Waugh, W. (1978). A Method For Self-Assessment Of Disability Before And After Hip Replacement Operations. *British medical journal*, *2*(6141), 857-859.

McKenna, S.P., Hunt, S.M., & McEwen, J. (1981). Weighting the Seriousness of Perceived Health Problems Using Thurstone's Method of Paired Comparisons. *International journal of epidemiology*, *10*(1), 93-97.

Moldovanu, C.M. (2010). *Managing the asymptomatic aortic stenosis.* Ph.D. thesis, coord. Pandele, G.I. & Cosovanu, A. Iaşi: "Gr. T. Popa" University of Medicine and Farmacy.

O'Brien, B.J., Buxton, M. J., & Patterson, D. L. (1993). Relationship Between Functional Status and Health-Related Quality-Of-Life After Myocardial Infarction. *Medical care*, *31*(10), 950-955.

Ozalevli, S., Ilgin, D., Karaali, H.K., Bulac, S., & Akkoclu, A. (2010). The Effect of In-Patient Chest Physiotherapy in Lung Cancer Patients. *Supportive care in cancer*, *18*(3), 351-358.

Peters, G.J.Y. (2014). The Alpha and the Omega of Scale Reliability and Validity: Why and How to Abandon Cronbach's Alpha and the Route Towards More Comprehensive Assessment of Scale Quality. *European Health Psychologist*, *16*(2), 56-69.

Prieto, L., Alonso, J., Ferrer, M., Antó, J.M., & Quality of Life in COPD Study Group. (1997). Are Results of the SF-36 Health Survey and the Nottingham Health Profile similar?: A comparison in COPD patients. *Journal of clinical epidemiology*, *50*(4), 463-473.

R Core Team (2014). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Retrieved from http://www.R-project.org/.

Revelle, W., & Zinbarg, R.E. (2009). Coefficients alpha, beta, omega, and the glb: Comments on Sijtsma. *Psychometrika*, *74*(1), 145-154.

Rhemtulla, M., Brosseau-Liard, P.E., & Savalei, V. (2012). When Can Categorical Variables Be Treated as Continuous? A Comparison of Robust Continuous and Categorical Sem Estimation Methods Under Suboptimal Conditions. *Psychological methods*, *17*(3), 354.

Rosseel, Y. (2012). lavaan: An R package for structural equation modeling. *Journal of Statistical Software*, *48*(2), 1-36.

Rosseel, Y., Oberski, D., Byrnes, J., Vanbrabant, L., Savalei, V., Merkle, E., ... & Rosseel, M. Y. (2015). Package 'lavaan'. Retrieved from https://cran.r-project.org/package=lavaan.

Sartorius, N, Kuyken W. (1994). Translation of Health Status Instruments. In J. Orley , W. Kuyken (Eds.), Quality of Life Assessment: International Perspectives (pp. 3-18). Berlin, Heidelberg: Springer-Verlag.

Schumacker, R.E., & Lomax, R.G. (2010). *A beginner's guide to structural equation modeling* (3rd Edition). UK: Routledge.

Steiger, J.H. (1990). Structural Model Evaluation and Modification: An Interval Estimation Approach. *Multivariate behavioral research*, *25*(2), 173-180.

Tucker, L.R., & Lewis, C. (1973). A Reliability Coefficient for Maximum Likelihood Factor Analysis. *Psychometrika*, *38*(1), 1-10.

West, S.G., Taylor, A.B., & Wu, W. (2012). Model fit and model selection in structural equation modeling. In R. H. Hoyle (Ed.), *Handbook of Structural Equation Modeling*. New York: Guilford.

Whynes, D.K., & Neilson, A. R. (1997). Symptoms Before and After Surgery for Colorectal Cancer. *Quality of Life Research*, *6*(1), 61–66.

Wiklund, I. (1989). The Nottingham Health Profile - a Measure of Health-Related Quality of Life. *Scandinavian journal of primary health care. Supplement*, *1*, 15-18.

Yildiz, N., Topuz, O., Gungen, G.O., Deniz, S., Alkan, H., & Ardic, F. (2010). Health-Related Quality of Life (Nottingham Health Profile) In Knee Osteoarthritis: Correlation with Clinical Variables and Self-Reported Disability. *Rheumatology international*, *30*(12), 1595-1600.

Zinbarg, R.E., Revelle, W., Yovel, I., & Li, W. (2005). Cronbach's α, Revelle's β, and McDonald's ωH: Their Relations with Each Other and Two Alternative Conceptualizations of Reliability. *Psychometrika*, *70*(1), 123–133.