



2530-0164/© 2018 SEEN y SED. Published by Elsevier España, S.L.U. All rights reserved.

Results: Three of 85 items in ThyPROes and four of the 39 items in ThyPRO-39es lacked convergent validity, while lack of discriminant validity was found for in nine and 14 items of each version respectively. Cronbach's alpha was >0.7 for 12 of 13 scales in the ThyPRO and 10 of 12 scales in the ThyPRO-39es. Eight items in the ThyPROes were flagged with DIF (one with non-uniform DIF), as were two items in the ThyPRO-39es. DIF magnitude was small (explained variance in the item score $<3\%$) in most cases, with a minor impact on scale scores.

Conclusions: The Spanish versions of the ThyPRO and ThyPRO-39 show acceptable psychometric properties and good cross-lingual validity, and are suitable for use in clinical studies.

© 2018 SEEN y SED. Published by Elsevier España, S.L.U. All rights reserved.

PALABRAS CLAVE

Calidad de vida;
Enfermedad tiroidea;
Cuestionario;
Traducción;
Validez transcultural

Adaptación y validación transcultural de la versión española del cuestionario Thyroid-Related Quality-of-Life Patient-Reported Outcome

Resumen

Introducción: El cuestionario *Thyroid-Related Quality-of-Life Patient-Reported Outcome* (ThyPRO) es el instrumento más utilizado para medir la calidad de vida en pacientes con enfermedades tiroideas benignas. Este estudio tiene como objetivo adaptar y validar una traducción al español del ThyPRO y su versión abreviada (ThyPRO-39).

Material y métodos: La adaptación al español se realizó utilizando el método de traducción-retrotraducción, seguido de una prueba preliminar en 5 pacientes representativos. El cuestionario definitivo ThyPROes se administró a 155 pacientes con trastornos tiroideos en un hospital terciario en España. Las propiedades psicométricas se evaluaron mediante la matriz multirrasgo-multimétodo y la estimación de la fiabilidad de la consistencia interna (alfa de Cronbach). Se utilizaron datos previos de 902 pacientes daneses para analizar el funcionamiento diferencial de los ítems (FDI) entre la versión original danesa del cuestionario y la española, mediante regresión logística ordinal.

Resultados: Tres de 85 ítems del ThyPROes y 4 de 39 del ThyPRO-39es carecían de validez convergente, mientras que 9 y 14, respectivamente, carecían de validez discriminante. El alfa de Cronbach fue $>0,7$ para 12 de 13 escalas del ThyPROes y 10 de 12 del ThyPRO-39es. Ocho ítems del ThyPROes mostraron FDI (uno con FDI no uniforme) y 2 lo hicieron en el ThyPRO-39es. La magnitud del FDI fue pequeña (varianza explicada en la puntuación del ítem $<3\%$) en la mayoría de casos, con un impacto menor en las puntuaciones de las escalas.

Conclusiones: Las versiones españolas del ThyPRO y ThyPRO-39 muestran aceptables propiedades psicométricas y buena validez interlingüística, y son adecuadas para su uso en estudios clínicos.

© 2018 SEEN y SED. Publicado por Elsevier España, S.L.U. Todos los derechos reservados.

Introduction

Health-related quality of life (HRQoL) can be defined as the global impact that diseases and their treatment exert on all relevant dimensions of the patient's life.¹ The measurement of HRQoL is increasingly used to determine the outcome of health interventions. Since it is a subjective concept that includes physical, mental and social aspects of the well-being of the individual, it can only be assessed by the patients themselves. For its measurement it is necessary to resort to the use of standardized questionnaires. In general, there are two types of questionnaires: generic questionnaires and disease specific questionnaires. The generic questionnaires are applicable to all types of patients and populations, including the general population, while the specific ones are focused on the measurement of certain

aspects of a particular disease, and are more appropriate to detect differences between alternative treatments.

As highlighted in recent years, different reasons justify the use of HRQoL measurement in patients with thyroid diseases.² First, they are common and chronic disorders that rarely compromise the patient's life, so that treatments are mainly aimed at optimizing their quality of life. In addition, some of them can be treated by different procedures (i.e. drugs, radioiodine, surgery), often with no options having demonstrated absolute superiority over the others. Finally, health professionals frequently face patients who claim deterioration of their HRQoL, even after receiving apparently effective treatment.

In the last decade, different specific questionnaires have been designed to evaluate HRQoL in patients with thyroid diseases. Among them, questionnaires directed to

patients with hyperthyroidism,³ thyroid orbitopathy⁴ or primary hypothyroidism.^{5–7} However, most of them have not been subjected to strict validation procedures and, in some cases, they were not used after the studies carried out for their elaboration.

More recently, Watt et al.^{8–11} developed the questionnaire “Thyroid-specific quality of life patient-reported outcome” (ThyPRO), designed to measure the quality of life related to all the main benign thyroid diseases. Unlike what happened in other cases, the authors of ThyPRO have published a rigorous process of development and validation of their questionnaire, including the generation phase of the contents,⁸ the preliminary investigation of its initial version by means of cognitive interviewing techniques,⁹ the analysis of its validity, reliability and internal consistency¹⁰ and its clinical validity and reproducibility.¹¹ Additionally, they performed a confirmatory factor analysis to demonstrate that each of the questionnaire scales was attributable to a single underlying explanatory factor.¹² Finally, they also demonstrated the sensitivity to the effects of treatment on the different thyroid diseases evaluated by the questionnaire.¹³ In addition, as the ThyPRO is a rather long questionnaire, the authors have developed an alternative short version (ThyPRO-39) which has been shown to preserve its measurement properties.¹⁴

In recent years, the ThyPRO has been applied to the clinical investigation of different issues related to thyroid disorders, such as the role of autoimmunity markers in the HRQoL of patients with primary hypothyroidism¹⁵ or the impact of different treatments on HRQoL among patients with multinodular goiter, Graves’ disease and subclinical hypothyroidism.^{16–24} Furthermore, it is being used as the main variable outcome in ongoing large-scale clinical trials evaluating emergent therapies for autoimmune thyroid disorders.²⁵ Thus, it can be said that the ThyPRO has become the reference tool for measuring HRQoL in patients with benign thyroid diseases.

The ThyPRO has been translated into numerous languages, and is currently available in English, Danish, German, Dutch, Italian, Portuguese, French, Swedish, Serbian, Polish, Romanian, Bulgarian, Greek, Arabic, Simplified Chinese, Traditional Chinese, Hebrew, Hindi and Tamil, and most of the translated versions have been subjected to transcultural validation.²⁶ However, a translation to Spanish language is still not available. The present study was aimed to adapt and validate a Spanish version of the ThyPRO.

Material and methods

The questionnaire

The ThyPRO consists of 85 items, summarized in 13 scales plus an individual item concerning overall impact of thyroid disease on HRQoL. Four of the 13 scales cover physical symptoms (goiter-, hyperthyroidism-, hypothyroidism- and eye-symptoms), two scales focus on mental symptoms, three on function and well-being and four on participation and social function. The items are scored from 0 to 4, following a Likert scale (where “0” equals “not at all” and “4” equals “very much”), always considering the patient’s perception during the last four weeks.

The short version of the ThyPRO (ThyPRO-39) retains 39 of the 85 items composing the original questionnaire, organized into four scales about symptoms, seven scales about physical, psychological and social well-being, a scale about physical appearance concerns, as well as an individual item on the overall impact on quality of life. In addition, it includes a composite score summarizing the results from the scales on non-thyroid symptoms, physical, psychological and social well-being.

Translation of the questionnaire to Spanish language

The adaptation of the ThyPRO questionnaire to Spanish language was performed according to the ISPOR Task Force recommendations.²⁷

The English version of the original questionnaire was translated into Spanish by two independent translators with English proficiency, one with and one without medical background. Both versions were discussed to agree on a reconciled preliminary Spanish questionnaire. This draft was sent to an English native expert translator in medical and scientific texts, but not familiar with the original version of the questionnaire, for back-translation. The main researcher of this project and the developer of ThyPRO evaluated and compared the wording, grammatical structure and meaning of each of the items of this back-translation with the source English version. In case of discrepancy, modifications were made and a second back-translation of the new version was performed by a second independent English native translator. This procedure was repeated until a definitive translated Spanish version was finally reached.

The translated questionnaire was pretested on five representative individuals with different thyroid diseases, using cognitive interviewing: a 49-year-old man with Graves’ disease, active hyperthyroidism and thyroid orbitopathy, a 55-year-old woman with hypothyroidism, a 62-year-old woman with non-toxic nodular goiter, a 22-year-old woman with euthyroid Graves’ disease and goiter, and a 61-year-old woman with “Hashitoxicosis” (autoimmune thyroid disease with spontaneous phases of hypo-, hyper- and euthyroidism). Interviews were conducted by a clinical psychologist with wide experience in health-related questionnaires.

Subjects

The definitive version of the Spanish questionnaire (ThyPROes) was applied in a self-administered manner to a convenience sample of 155 adult Spanish subjects with different benign thyroid diseases. Participants were recruited successively in the Department of Endocrinology of a single tertiary hospital, when they attended a scheduled appointment for medical visit or thyroid ultrasonography. For the present investigation they were grouped in six major mutually excluding categories: non-toxic goiter, toxic nodular goiter, Graves’ disease, thyroid orbitopathy, autoimmune hypothyroidism and other benign disorders. The main clinical and sociodemographic characteristics of the participants are shown in Table 1. All patients signed informed consent

Table 1 Main clinical and sociodemographic characteristics of the participants in the present study and in the Danish reference population used for transcultural validation.

	Spanish subjects (N = 155)	Danish subjects (N = 902)
Age (years)	52 (44–61)	51 (40–61)
Women/men (%)	84.5/15.5	87/13
Education level (%)		
Primary	41.3	29
Secondary	43.2	29
Tertiary	15.5	38
Not available	–	4
Clinical diagnosis (%)		
Non-toxic goiter	38.7	29
Toxic nodular goiter	7.7	16
Graves' hyperthyroidism	14.8	18
Graves' orbitopathy	7.1	10
Autoimmune hypothyroidism	31	22
Other	0.6	4
Comorbidity (%)	52.3	56
Current treatment (%)		
Levothyroxine	48.4	32
Antithyroid drugs	9	18
Previous treatment (%)		
Radioiodine	11	13
Thyroid surgery	12.9	14
TSH (mU/l)	1.99 (0.97–3.74)	1.12 (0.36–2.7)

Data are percentages and medians (interquartile range).

before recruitment. The study was approved by the local Ethics Committee.

Psychometric validation

The internal consistency of both the ThyPROes and its short version (ThyPRO-39es) was assessed using multitrait scaling analysis. To evaluate convergent validity, correlation coefficients between each item and the total score of the remaining items of the own scale were calculated, and to evaluate discriminant validity, correlations were calculated between each item and the total score of the rest of the scales. Correlations greater than 0.4 between the individual items and their own scales were considered indicative of adequate convergent validity (all items in a scale appear to measure the same construct). If the correlations between each item and other scales were lower than those observed with their own scale, discriminant validity was considered adequate (items in different scales measure different constructs). Validity was additionally assessed by calculation of inter-scale correlations.

The reliability of the internal consistency was evaluated by the Cronbach's alpha coefficient, based on the average of the correlations between the items of each of the scales. Cronbach's alpha ranges from 0 to 1, with values >0.7 usually considered as indicative of satisfactory internal consistency.²⁸ Pearson's correlations were used in all cases. All analyses were performed for both the ThyPROes and the ThyPRO-39es.

Transcultural validation

Equivalence between the Spanish version and the original Danish questionnaire was assessed using tests for differential item functioning (DIF). An item presents DIF when the probability of a correct response by individuals with the same ability level differs across groups with different characteristics (in this case different language). Responses to the ThyPROes and to the ThyPRO-39es were compared pairwise to the Danish version using data from a previously reported sample of 902 subjects recruited from two endocrine clinics in Denmark.²⁶ The characteristics of this reference population are also shown in Table 1. DIF was assessed by means of ordinal logistic regression, controlling for specific diagnosis. The independent variables were language group and scale score, and an interaction term between scale score and language group. DIF was considered uniform or non-uniform depending on whether the interaction term was significant or not. DIF magnitude was considered substantial if it could explain more than 2% of the variance in the item score (R^2 difference >0.02).

Results

Translation of the questionnaire to Spanish language

The comparison of the first back-translation and the English master questionnaire identified discrepancies that required minor modifications in the introductory paragraph to explain the contents of the questionnaire, in six of its 85 items and in one of the five response options to the questions. Alternative wording was considered for other two items, but was eventually discarded. Item 7b, which evaluates feelings of depersonalization ('felt "not like yourself"? in the English questionnaire), which became the most complex for translation, required a new modification after the second back-translation, and was definitively written as '¿se ha sentido extraño consigo mismo?'.

In the pretesting phase of cognitive interviews, the five interviewees reported that the questionnaire was easy to understand and that the language was adequate. Nevertheless, four of them reported some doubts in the response process. According to the question answering model, in most cases the problems referred to the phase of elaboration of a judgment on the processed information. In particular they concerned attribution, i.e. the patients showed their doubts about whether the problems referred to in some questions were due to their thyroid disease or to other reasons. These issues were considered inherent to the ThyPRO itself rather

Table 2 Cronbach's alpha coefficients for scales of the ThyPROes and ThyPRO-39es questionnaires.

	ThyPROes	ThyPRO-39es
Goiter symptoms	0.92	0.84
Hyperthyroidism symptoms	0.84	0.75
Hypothyroidism symptoms	0.76	0.77
Eye symptoms	0.91	0.81
Tiredness	0.90	0.80
Cognitive complaints	0.94	0.90
Anxiety	0.93	0.90
Depression	0.93	0.79
Emotional susceptibility	0.91	0.72
Impaired social life	0.69	0.67
Impaired daily life	0.90	0.79
Impaired sexual life	0.84	–
Cosmetics complaints	0.78	0.55
Composite scale	–	0.88

than to the translation, and did not result in any additional modification of the questionnaire.

Psychometric validation

Three items of the ThyPROes and four items of the ThyPRO-39es lacked convergent validity (two of them were common to both versions of the questionnaire). On the whole, three of the five items lacking convergent validity in either of the two questionnaires were part of the scales that assess the impact of thyroid disease on impaired social life and cosmetic complaints (Tables 1 and 2 in Supplementary Material). Nine items in the ThyPROes and 14 in the ThyPRO-39es correlated more strongly with one or more other scales than with its own scale (lack of discriminant validity). Scales that evaluate symptoms related to goiter and the eyes, and those that evaluate cognitive complaints, anxiety and cosmetic problems showed complete discriminant validity in both the ThyPROes and the ThyPRO-39es (Tables 1 and 2 in Supplementary Material).

The correlation coefficients between the different scales of the ThyPROes were low or moderate in most cases. Scales dealing with tiredness and mental health (particularly depression, anxiety and emotional susceptibility) were the most strongly correlated ones (Table 3 in Supplementary Material). The analyses of the reliability of the internal consistency showed Cronbach's alpha values higher than 0.7 in all the scales of ThyPROes, except for the one measuring "Impaired Social Life", whose Cronbach's alpha coefficient was 0.69, just below the widely accepted cut-off of 0.7 (Table 2). As expected by the reduction in the number of items of each scale, internal consistency reliability estimates were lower in the ThyPRO-39es than in the ThyPROes. Even so, all the ThyPRO-39es scales, except "Impaired Social Life" and "Cosmetic Complaints" had Cronbach's alpha coefficients greater than 0.7 (Table 2).

Trans-cultural validation

Eight items in the ThyPROes (one in the "Goiter Symptoms" scale, one in the "Eye Symptoms" scale, two in

the "Tiredness" scale, three in the "Anxiety" scale and one in the "Emotional susceptibility" scale) and two items in the ThyPRO-39es (both in the "Anxiety" scale) were flagged with DIF. As a result, eight scales of ThyPROes and nine scales of ThyPRO-39es were free of DIF. DIF was found to be non-uniform in question 3a of the ThyPROes "Tiredness" scale (*¿se ha sentido lleno de vida?*). As shown in Tables 3 and 4, the magnitude of DIF was small in most of cases (explained variance in the item score <3% in seven items of the ThyPROes and in one of the ThyPRO-39es) and direction of DIF in items belonging to the same scales was variable.

Discussion

The measurement of HRQoL has become one of the main outcomes in clinical research on thyroid diseases. However, until now there have been no specific tools to assess the impact of thyroid diseases on HRQoL that have been translated and validated in Spanish language. This represents a considerable limitation in the evaluation of the results of clinical studies of thyroid disorders in patients from Spanish-speaking countries, and notably complicates their participation in multinational surveys.

The present report presents an adaptation and validation of the ThyPRO, a questionnaire whose use has been adopted in the latest years by numerous thyroid researchers, and that has been recently recommended for patients with benign thyroid diseases due to the quality of its measurement properties.²⁹

Using the forward-backward translation technique, we obtained a readily comprehensible and easy to fill out Spanish version of the ThyPRO, in which major problems were not identified during a pretesting pilot survey of five representative patients with different thyroid diseases.

Like the original questionnaire¹⁰ and other translated versions,³⁰ the ThyPROes demonstrated good quality measurement using classic procedures for reliability assessment. Additionally, following the methodology used for cross-cultural validation of other language versions,²⁶ we also evaluated DIF between the Spanish translated version and the Danish master questionnaire to investigate potential language bias. Although our results could also be explained by some differences between the two samples studied, such as the greater proportion of patients with toxic multinodular goiter and the higher level of academic education among Danish patients, a total of only eight out of 85 questionnaire items were flagged with DIF. These findings are in line with those observed for other linguistic versions, which, using as reference the English master, showed language-related DIF in a number of items ranging from a minimum of one (in Dutch and Danish) to a maximum of 12 in Serbian, 13 in Italian and 16 in Hindi.²⁶ Interestingly, seven of the eight items identified in the present study were also affected by DIF in one or more languages in the previous multicultural validation study of the ThyPRO,²⁶ suggesting difficult translatability or poor equality of these items. In any case, the magnitude of DIF in our study was small in all cases, and the direction of DIF was discordant among items that belonged to the same scale, implying that the global impact on the scale scores will be minor.

Table 3 Differential item functioning of the ThyPROes.

Abbreviated item wording	Variance explained by DIF (%)	Direction of DIF
<i>Goiter symptoms scale</i>		
Sensación de ahogamiento	2.7	Spanish patients report less sensation of suffocating
<i>Eye symptoms scale</i>		
Empeoramiento en la visión	2.9	Spanish patients report more impaired vision
<i>Tiredness scale</i>		
Sentirse lleno de vida ^a	2.5	Spanish patients report less vitality with more tiredness and more vitality with less tiredness
Afrontar las demandas de la vida cotidiana	5.6	Spanish patients report better coping
<i>Anxiety scale</i>		
Sentirse nervioso	2.7	Spanish patients report more nervousness
Preocupación por estar muy enfermo	2.4	Spanish patients report less concern about being seriously ill
Sentirse incómodo	2.1	Spanish patients report less uneasiness
<i>Emotional Susceptibility scale</i>		
Sentirse extraño con uno mismo	2.5	Spanish patients report lower levels of self-unrecognizability

^a Non-uniform DIF.**Table 4** Differential item functioning of the ThyPRO-39es.

Abbreviated items wording	Variance explained by DIF (%)	Direction of DIF
<i>Anxiety scale</i>		
Sentir miedo o ansiedad	5.8	Spanish patients report being more afraid
Sentirse incómodo	2.1	Spanish patients report less uneasiness

Unlike other ThyPRO translations, which were implemented before the development of the ThyPRO-39, we also had the opportunity to validate a Spanish version of this short version of the questionnaire. The measures of construct validity and internal consistency reliability were lower for the ThyPRO-39es than for the complete questionnaire. This finding was expected, since the scales of ThyPRO-39 only keep a reduced number of the items of the original tool, and the internal consistency reliability measures of a questionnaire increase as the number of items that make it up.³¹ Even so, 10 of the 12 scales of the ThyPRO-39es, as well as the overall Composite score summarizing the seven well-being and function scales, showed a Cronbach's alpha coefficient greater than 0.7. Regarding the DIF analyses, only two items of the ThyPRO-39 showed different functioning between the Spanish and the Danish populations. This low rate of DIF probably reflects that the item selection strategy for development of the abbreviated version of the ThyPRO gave preference to items with transcultural validity,¹⁴ and reaffirms the measurement properties of ThyPRO-39 for use in studies of multiethnic populations.

The present research has several limitations. As mentioned above, the study subjects and the reference population used for transcultural validation were not

identical regarding diagnostic groups and some sociodemographic variables that could influence the response process, such as the level of academic education. In addition, we have not addressed other validation methods for patient-reported outcome instruments, which otherwise could be subject of further investigations. Among them, test-retest reliability, which assesses the ability to provide consistent measures over time by collecting two responses separated by 2–3 weeks in the same individuals, could provide a more adequate evaluation of the precision of the questionnaire. Confirmatory factor analysis, on the other hand, could be used to support the construct validity of the ThyPROes by demonstrating that its factorial structure corresponds to the original scales of the questionnaire. Finally, the questionnaire has been adapted to the Spanish language used in Spain, and it is possible that it may contain words or linguistic uses with different connotations in other Spanish-speaking cultures. Therefore, it cannot be ruled out that this can generate biases in some items if used without prior validation in other countries.³²

In conclusion, the herein reported Spanish versions of the ThyPRO and the ThyPRO-39 show adequate psychometric measures and good cross-lingual validity and are thus suitable for use in clinical studies involving Spanish-speaking subjects.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Conflict of interest

The authors declare that they have no conflict of interest.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at [doi:10.1016/j.endinu.2018.06.010](https://doi.org/10.1016/j.endinu.2018.06.010).

References

- Ebrahim S. Clinical and public health perspectives and applications of health-related quality of life measurement. *Soc Sci Med*. 1995;41:1383–94.
- Watt T, Groenvold M, Rasmussen AK, Bonnema SJ, Hegedüs L, Bjorner JB, et al. Quality of life in patients with benign thyroid disorders. A review. *Eur J Endocrinol*. 2006;154:501–10.
- Fahrenfort JJ, Wilterdink AM, van der Veen EA. Long-term residual complaints and psychosocial sequelae after remission of hyperthyroidism. *Psychoneuroendocrinology*. 2000;25:201–11.
- Terwee CB, Gerding MN, Dekker FW, Prummel MF, Wiersinga WM. Development of a disease-specific quality of life questionnaire for patients with Graves' ophthalmopathy: the GO-QOL. *Br J Ophthalmol*. 1998;82:773–9.
- Saravanan P, Chau W-F, Roberts N, Vedhara K, Greenwood R, Dayan CM. Psychological well-being in patients on «adequate» doses of l-thyroxine: results of a large, controlled community-based questionnaire study. *Clin Endocrinol (Oxf)*. 2002;57:577–85.
- Jaeschke R, Guyatt G, Cook D, Harper S, Gerstein HC. Spectrum of quality of life impairment in hypothyroidism. *Qual Life Res*. 1994;3:323–7.
- McMillan CV, Bradley C, Woodcock A, Razvi S, Weaver JU. Design of new questionnaires to measure quality of life and treatment satisfaction in hypothyroidism. *Thyroid*. 2004;14:916–25.
- Watt T, Hegedüs L, Rasmussen AK, Groenvold M, Bonnema SJ, Bjorner JB, et al. Which domains of thyroid-related quality of life are most relevant? Patients and clinicians provide complementary perspectives. *Thyroid*. 2007;17:647–54.
- Watt T, Rasmussen AK, Groenvold M, Bjorner JB, Watt SH, Bonnema SJ, et al. Improving a newly developed patient-reported outcome for thyroid patients, using cognitive interviewing. *Qual Life Res*. 2008;17:1009–17.
- Watt T, Bjorner JB, Groenvold M, Rasmussen AK, Bonnema SJ, Hegedüs L, et al. Establishing construct validity for the thyroid-specific patient reported outcome measure (ThyPRO): an initial examination. *Qual Life Res*. 2009;18:483–96.
- Watt T, Hegedüs L, Groenvold M, Bjorner JB, Rasmussen AK, Bonnema SJ, et al. Validity and reliability of the novel thyroid-specific quality of life questionnaire, ThyPRO. *Eur J Endocrinol*. 2010;162:161–7.
- Watt T, Groenvold M, Deng N, Gandek B, Feldt-Rasmussen U, Rasmussen AK, et al. Confirmatory factor analysis of the thyroid-related quality of life questionnaire ThyPRO. *Health Qual Life Outcomes*. 2014;12:126.
- Watt T, Cramon P, Hegedüs L, Bjorner JB, Bonnema SJ, Rasmussen AK, et al. The thyroid-related quality of life measure ThyPRO has good responsiveness and ability to detect relevant treatment effects. *J Clin Endocrinol Metab*. 2014;99:3708–17.
- Watt T, Bjorner JB, Groenvold M, Cramon P, Winther KH, Hegedüs L, et al. Development of a short version of the thyroid-related patient-reported outcome ThyPRO. *Thyroid*. 2015;25:1069–79.
- Watt T, Hegedüs L, Bjorner JB, Groenvold M, Bonnema SJ, Rasmussen AK, et al. Is thyroid autoimmunity per se a determinant of quality of life in patients with autoimmune hypothyroidism? *Eur Thyroid J*. 2012;1:186–92.
- Mishra A, Sabaretnam M, Chand G, Agarwal G, Agarwal A, Verma AK, et al. Quality of life (QoL) in patients with benign thyroid goiters (pre- and post-thyroidectomy): a prospective study. *World J Surg*. 2013;37:2322–9.
- Bukvic BR, Zivaljevic VR, Sipetic SB, Diklic AD, Tausanovic KM, Paunovic IR. Improvement of quality of life in patients with benign goiter after surgical treatment. *Langenbecks Arch Surg*. 2014;399:755–64.
- Graf H, Fast S, Pacini F, Pinchera A, Leung A, Vaisman M, et al. Modified-release recombinant human TSH (MRhTSH) augments the effect of 131I therapy in benign multinodular goiter: results from a multicenter international, randomized, placebo-controlled study. *J Clin Endocrinol Metab*. 2011;96:1368–76.
- Zivaljevic VR, Bukvic Bacotic BR, Sipetic SB, Stanisavljevic DM, Maksimovic JM, Diklic AD, et al. Quality of life improvement in patients with Hashimoto thyroiditis and other goiters after surgery: a prospective cohort study. *Int J Surg*. 2015;21:150–5.
- Cramon P, Bonnema SJ, Bjorner JB, Ekholm O, Feldt-Rasmussen U, Frenzl DM, et al. Quality of life in patients with benign nontoxic goiter: impact of disease and treatment response, and comparison with the general population. *Thyroid*. 2015;25:284–91.
- Bukvic B, Zivaljevic V, Sipetic S, Diklic A, Tausanovic K, Stojanovic D, et al. Improved quality of life in hyperthyroidism patients after surgery. *J Surg Res*. 2015;193:724–30.
- Taieb D, Bournaud C, Eberle MC, Catargi B, Schwartz C, Cavarec MB, et al. Quality of life, clinical outcomes and safety of early prophylactic levothyroxine administration in patients with Graves' hyperthyroidism undergoing radioiodine therapy: a randomized controlled study. *Eur J Endocrinol*. 2016;174:491–502.
- Stott DJ, Rodondi N, Kearney PM, Ford I, Westendorp RGJ, Mooijaart SP, et al. Thyroid Hormone therapy for older adults with subclinical hypothyroidism. *N Engl J Med*. 2017;376:2534–44.
- Sorensen JR, Watt T, Cramon P, Døssing H, Hegedüs L, Bonnema SJ, et al. Quality of life after thyroidectomy in patients with nontoxic nodular goiter: a prospective cohort study. *Head Neck*. 2017;39:2232–40.
- Winther KH, Watt T, Bjorner JB, Cramon P, Feldt-Rasmussen U, Glud C, et al. The chronic autoimmune thyroiditis quality of life selenium trial (CATALYST): study protocol for a randomized controlled trial. *Trials*. 2014;15:115.
- Watt T, Barbesino G, Bjorner JB, Bonnema SJ, Bukvic B, Drummond R, et al. Cross-cultural validity of the thyroid-specific quality-of-life patient-reported outcome measure, ThyPRO. *Qual Life Res*. 2015;24:769–80.
- Wild D, Grove A, Martin M, Eremenco S, McElroy S, Verjee-Lorenz A. Principles of good practice for the translation and cultural adaptation process for patient-reported outcomes (PRO) measures: report of the ISPOR task force for translation and cultural adaptation. *Value Health*. 2005;8:94–104.
- Nunnally JC, Bernstein IH. *Psychometric theory*. 3rd ed New York: McGraw-Hill; 1994.
- Wong CK, Lang BH, Lam CL. A systematic review of quality of thyroid-specific health-related quality-of-life instruments recommends ThyPRO for patients with benign thyroid diseases. *J Clin Epidemiol*. 2016;78:63–72.
- Bukvić B, Živaljević V, Šipetić S, Diklić A, Taušanović K, Paunović I. Validation and cross-cultural adaptation of the

- questionnaire ThyPRO in thyroid patients in Serbia. *Vojnosanit Pregl.* 2015;72:583–8.
31. Tavakol M, Dennick R. Making sense of Cronbach's alpha. *Int J Med Educ.* 2011;2:53–5.
32. Vallejo-Medina P, Gómez-Lugo M, Marchal-Bertrand L, Saavedra-Roa A, Soler F, Morales A. Developing guidelines for adapting questionnaires into the same language in another culture. *Ter Psicol.* 2017;35:159–72.