Mapping the Patient-Oriented Prostate Utility Scale from the EPIC and the SF Health Surveys

Víctor Zamora, MPH[1,2,3]; Olatz Garin, PhD[1,3,4]; Yolanda Pardo, PhD[3,1,5]; Àngels Pont, MSc[1,3]; Cristina Gutiérrez, MD, PhD[6]; Patricia Cabrera, MD, PhD[7]; Francisco Gómez-Veiga, MD, PhD[8]; José Ignacio Pijoan, MD, PhD[3,9]; Mark S Litwin, MD, PhD[10]; Montse Ferrer, MD, PhD[1,2,3] on behalf of the Multicentric Spanish Group of Clinically Localized Prostate Cancer

The Multicentric Spanish Group of Clinically Localized Prostate Cancer: Montse Ferrer, Àngels Pont, Olatz Garin, Yolanda Pardo, Víctor Zamora, Cristina Gutierrez, Montse Ventura, Ferran Guedea, Ferran Ferrer, Ana Boladeras, Andrea Slocker, José Francisco Suárez, Manuel Castells, Xavier Bonet, Patricia Cabrera, David B Delgado, MªJosé Ortiz, Ismael Herruzo, José López-Torrecilla, Jorge Pastor, Víctor Muñoz, Patricia Willsich, Marisa Vázquez, Àlvar Roselló, Arantxa Eraso, Carlos Ferrer, Ángel Sánchez, Francisco Gómez-Veiga, Víctor Macías, Lluís Fumadó, Josep Jové, Moisés Mira, Elena Villafranca, Juan Morote, Ana Celma, Pilar Samper, Luís A Glaría, MªÁngeles Cabeza, Germán Juan, Samuel Méndez Ramírez, Amalia Palacios, Amelia Béjar, Sonia Garcia, Sebastà Sabater.

1 Health Services Research Group, IMIM (Hospital del Mar Medical Research Institute), Barcelona, Spain.

2 Department of Paediatrics, Obstetrics and Gynaecology and Preventive Medicine, Universitat Autònoma de Barcelona (UAB), Bellaterra, Barcelona, Spain.

3 CIBER en Epidemiología y Salud Pública, CIBERESP, Spain.

4 Universitat Pompeu Fabra, Barcelona, Spain.

5 Department of Psychiatry and Legal Medicine, Universitat Autònoma de Barcelona, Barcelona, Spain.

6 Institut Català d'Oncologia, IDIBELL, L'Hospitalet de Llobregat, Spain.

7 Department of Radiation Oncology, Hospital Universitario Virgen del Rocío, Sevilla, Spain.

8 Complejo Hospitalario Universitario de Salamanca (CAUSA), Grupo de Investigación Translacional de Urología, Instituto de Investigación de Salamanca, Salamanca, Spain.

9 Department of Clinical Epidemiology, Hospital Universitario de Cruces, Barakaldo, Vizcaya, Spain.

10 Schools of Medicine, Public Health and Nursing, University of California, Los Angeles, USA.

Contact information for corresponding author:

Montse Ferrer and Olatz Garin

Health Services Research Group, IMIM (Hospital del Mar Medical Research Institute)

Doctor Aiguader 88, 08003 Barcelona, Spain.

Phone: +34 93 3160 740; Fax: +34 93 3160 797; e-mail address: mferrer@imim.es

Phone: +34 933 160 758; Fax: +34 933 160 797; e-mail address: ogarin@imim.es

Précis:

An algorithm to estimate PORPUS utilities from EPIC and SF Health Surveys scores has been developed showing excellent predictive capacity.

Conflict of Interest Disclosures: The authors reported no conflicts of interest.

Funding support: Mr. Zamora reported receiving a personal grant from the ISCIII-Fondo de Investigación Sanitaria (FI19/00229). Mr. Zamora, Drs Garin and Pardo, Ms. Pont and Dr Ferrer reported receiving grants from ISCIII- Fondo Europeo de Desarrollo Regional (PI13/00412) and Generalitat de Catalunya (2017 SGR 452). Mr. Zamora, Drs Garin and Pardo, Ms. Pont and Drs Gutiérrez, Cabrera, Gómez-Veiga, Pijoan and Ferrer reported receiving funding for the project from the Movember Foundation's TrueNTH Global Registry. Mr. Zamora, Drs Garin and Pardo, Ms. Pont and Drs Pijoan and Ferrer reported receiving grants from CIBER de Epidemiología y Salud Pública (CIBERESP).

Role of the Funder/Sponsor: The funders had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

Acknowledgments: Authors would like to acknowledge the collaboration of the other members of the Multicentric Spanish Group of Clinically Localized Prostate Cancer for their contributions throughout the project. All authors have read the manuscript before submission. Moreover, the authors would like to thank Aurea Martín for her support in English editing, proofreading, and preparing this manuscript for submission.

ABSTRACT (250 / 250 words)

OBJECTIVES: To develop mapping algorithms from the Expanded Prostate cancer Index Composite (EPIC) and the Short-Form Health Surveys (SF) to the Patient-Oriented Prostate Utility Scale (PORPUS), an econometric instrument specifically developed for patients with prostate cancer.

METHODS: Data were drawn from two cohorts concurrently administering PORPUS, EPIC-50 and SF-36v2. The development cohort included patients diagnosed with localized or locally advanced prostate cancer in 2017-2019. The validation cohort included men diagnosed with localized prostate cancer in 2014-2016. Linear regression models were constructed with ln(1 - PORPUS utility) as the dependent variable and scores from the original and brief versions of the EPIC and SF as independent variables. The predictive capacity of mapping models constructed with all possible combinations of these two instruments was assessed through the proportion of variance explained (R^2), and the agreement between predicted and observed values. Validation was based on the comparison between estimated and observed utility values in the validation cohort.

RESULTS: Models constructed with EPIC-50 with and without SF yielded the highest predictive capacity (R^2=0.884, 0.871 and 0.842) in comparison with models constructed with EPIC-26 (R^2=0.844, 0.827 and 0.776). The intraclass correlation coefficient was excellent in the four models (>0.9) with EPIC and SF. In the validation cohort, predicted PORPUS utilities were slightly higher than those observed, but differences were not statistically significant.

CONCLUSIONS: Mapping algorithms from both the original and the abbreviated versions of the EPIC and the SF Health Surveys allow estimating PORPUS utilities for economic evaluations with cost-utility analyses in prostate cancer patients.

.

INTRODUCTION

Prostate cancer is the most frequently diagnosed cancer among men in USA and Europe, and the third cause of death from cancer [1]. The prostate-specific antigen (PSA) has led to a rapid increase in incidence rates and an early detection of the disease among asymptomatic and younger men. Most prostate cancer patients are diagnosed in localized stages [2], which allows evidence-based shared decision making regarding their disease management, and becoming long-term survivors [3].

The optimal management of men with localized prostate cancer is controversial. There are numerous treatment alternatives available (including surgery, radiotherapy, and active surveillance, among others), and the rapid adoption of newer modalities has introduced additional uncertainty to the decision-making process [4]. Results from the ProtecT trial in localized prostate cancer patients indicate similar survival rates regardless of treatment, but relevant differences in the side effects patterns, measured with Patient-Reported Outcomes (PROs) [5, 6].

PROs can be classified as psychometric profiles or econometric indexes [7]. Psychometric measures generate scores on different health dimensions (profiles). Econometric measures provide a single global score (index) which incorporates societal or patients' preferences for health states (utilities) that can be used to calculate quality-adjusted life years (QALYs) in economic evaluations.

The Patient-Oriented Prostate Utility Scale (PORPUS) [8] is the only econometric instrument developed for patients with prostate cancer, so it presents the advantages of disease-specific PROs: it is more responsive in detecting changes and differences between clinical groups than generic instruments [9], since it measures prostate cancer

symptoms and treatments' side effects (i.e. urinary, sexual or hormonal), most of which are not covered by generic econometric instruments. The PORPUS applies patient-based preferences to obtain utilities [10]. Although PROs have been widely used in patients with localized prostate cancer [5, 6, 11, 12], most studies do not include econometric measures [13], preventing the estimation of QALYs and cost-utility analyses.

Having a mapping algorithm from the most widely used psychometric PROs in this area to the PORPUS, therefore, would be a solution of high relevance to calculate utilities for prostate cancer patients (both in existing registries or newly gathered data). Mapping involves the equating (or linking) of values from a source instrument to equivalent values on a target instrument [13]. Most prostate cancer mapping studies have linked a disease-specific psychometric instrument to a generic econometric one, even though they do not measure the same constructs [14-16]. The differences in the constructs measured between the source and the target instruments are likely the reason why prostate cancer mapping studies have not found high predictive capacities [17]. Only one study [18] mapped the PORPUS from a prostate cancer-specific instrument (the University of California-Los Angeles-Prostate Cancer Index, UCLA-PCI).

The currently most used and recommended prostate cancer-specific PRO is the Expanded Prostate cancer Index Composite (EPIC) [19], which was in fact derived from UCLA-PCI by enhancing the urinary domain. As far as we know, no study has mapped the PORPUS from the EPIC. Therefore, our aim was to develop a mapping algorithm from EPIC to PORPUS, and to evaluate whether adding any version of the most used and well validated generic instrument, the Short-Form (SF) Health Survey, improves its predictive capacity.

METHODS

Patient data

Data used for this study were drawn from two different prostate cancer cohorts in which PORPUS, EPIC-50 and SF-36v2 were administered concurrently through telephone interviews before and after treatment.

The algorithm was developed using data of Spanish patients diagnosed with localized or locally advanced prostate cancer (any T, any N and M0) between January 2017 and May 2019. This cohort, part of the Movember Foundation's TrueNTH Global Registry [20], was selected for the development of the algorithm due to its high heterogeneity [13] and, from now on, will be called the 'development cohort'. The 12 months after treatment evaluation was selected as the one that may better reflect more perdurable side effects from treatments.

Data for the validation of the algorithm came from another cohort of Spanish men diagnosed with localized prostate cancer between February 2014 and August 2016, from now on called the 'validation cohort'. Patients' inclusion criteria of this cohort were: 50–75 years old; clinical stage T1 or T2, N0/Nx and M0/Mx; Gleason ≤ 6 or 7 (if 3+4 with T1c); and PSA ≤ 10 ng/ml. These restricted criteria were defined with the primary aim of assessing the effectiveness of new treatment modalities for localized prostate cancer of low-intermediate risk, but make the sample too homogeneous to be considered for algorithm development. The evaluations at 12, 24 and 36 months after treatment were selected to validate the algorithm at mid- and long-term follow-up.

Patient-Reported Outcome Measures

The Patient-Oriented Prostate Utility Scale (PORPUS) [8] is a 10-attribute health state classification system that includes 5 broad items of health-related quality of life (pain, energy, social support, communication with doctor, and emotional well-being) and 5 prostate cancer-specific items describing sexual function and desire, urinary frequency and incontinence, and bowel function. The items in the PORPUS have a Likert scale format with four to six levels, resulting in 6,000,000 potential health states related to prostate cancer. A multiattribute utility function was constructed using the patient-weighted utilities elicited with standard gamble [10] from three cohorts of patients with prostate cancer: localized (n=91; mean age = 64.2), metastatic (n=53; mean age = 69.9) and nonmetastatic survivors (n=90; mean age = 68.4), who had undergone hormonotherapy (42%), radical prostatectomy (32%) or radiotherapy (31%) [10]. The PORPUS utility (PORPUS-U) index ranges between 0 (dead) and 1 (perfect health). The original version has shown to have good reliability, validity [21] and responsiveness [22]. The Spanish version has also demonstrated good validity and responsiveness in localized prostate cancer patients [23].

The Expanded Prostate cancer Index Composite [24] contains 50 items (EPIC-50), and its abbreviated version has 26 items (EPIC-26) [25], measuring in both cases four domains: urinary (with 11 and 9 items, respectively), sexual (13 and 6 items), bowel (14 and 6 items), and hormonal (11 and 5 items). Response options for each EPIC item are on a 4-, 5-, or 6-level Likert scale. Items are grouped in summary scores for sexual, bowel and hormonal domains, as well as in two scores for the urinary domain (incontinence and irritative/obstructive symptoms), which are transformed linearly to a scale from 0 to 100, where higher scores indicate better outcomes. Furthermore, the score proposal for EPIC-50 also includes subscales to discern function and bother for

each domain. The Spanish version of the EPIC-50 has been shown to be reliable and valid, and to have excellent sensitivity to change [26].

The 36-item Short-Form Health Survey (version 2) contains 36 items (SF-36v2) [27, 28], and its abbreviated version has 12 items (SF-12v2) [29], both of them covering eight dimensions: physical functioning (with 10 and 2 items, respectively), role physical (4 and 2 items), bodily pains (2 and 1 items), general health (5 and 1 items), vitality (4 and 1 items), social functioning (2 and 1 items), role emotional (3 and 2 items) and mental health (6 and 2 items). Both versions can generate dimension scores and physical and mental component summaries. Scores were constructed using the developers' algorithms, and were standardized to have a mean of 50 and a standard deviation of 10 in the US general population. Moreover, an utility algorithm can be derived from both versions using societal preferences elicited with standard gamble: the SF-6D index [30]. The Spanish version has been widely validated [31, 32].

Sample size calculation

Assuming that we expected to surpass the R^2 provided by the mapping model of Bremner et al (R^2=0.72) [18] and α=0.05, the statistical power is higher than 0.80, considering that we had data available from 266 patients with prostate cancer evaluated at 12 months of follow-up in the development cohort.

Statistical analyses

To describe demographic and clinical characteristics, either means and standard deviations (SD) or frequencies and percentages were calculated, according to the nature of each variable. We examined the distribution of the PORPUS-U index through

histogram, and a logarithmic transformation was applied to deal with the skewness to the left observed at the upper bound: ln(1 - PORPUS-U).

The mapping's predictive models were developed by using patient responses to the EPIC and the SF Health Survey at 12 months after treatment in the development cohort. As the brief versions of EPIC and SF Health Survey are increasingly adopted, we constructed PORPUS-U mapping models with all their possible combinations.

Linear regression models were constructed with ln(1 - PORPUS-U) as dependent variable, following a previous modelling approach [18]. As independent variables, we considered all scores of the EPIC domains (urinary, bowel, sexual and hormonal) and/or the eight dimensions of the SF Health Surveys (physical functioning, role physical, bodily pains, general health, vitality, social functioning, role emotional and mental health). We tested all possible first-level interactions between independent variables.

After transforming the distribution of PORPUS-U values, we cannot assume that the error distribution is normal. Therefore, without knowing the error distribution for a reliable parametric distribution, we applied the recommended retransformation with the Smearing estimator [33].

Predictive capacity of models was assessed with R^2, Root Mean Square Error (RMSE) and Intraclass Correlation Coefficient (ICC) between the index estimated by the model and the one observed. R^2 is the proportion of the variance of the dependent variable that is predictable from the independent variables. It ranges from 0 to 1, and the lowest acceptable threshold suggested is 0.70 [17]. RMSE is measured in the same units as the dependent variable and is representative of the size of a 'typical' error. The lower the RMSE, the better the model is performing [34]. ICC values lower than 0.50, between

0.50 and 0.75, between 0.75 and 0.90, and greater than 0.90 are indicative of poor, moderate, good, and excellent agreement, respectively [35].

The validation of the models was based on estimating the utility values with the mapping algorithm from the EPIC and/or the SF Health Survey in the validation cohort at 12, 24 and 36 months after treatment, and comparing them with those measured directly from patients with the PORPUS-U. Central tendency and dispersion statistics of predicted and observed values were calculated, and differences were examined with Wilcoxon test. The agreement between the predicted values of the models and the ones directly measured with PORPUS was also assessed with RMSE, ICC and Bland-Altman plots.

To explore the predictive capacity according to different known groups, the distribution of the observed utilities (with PORPUS-U and SF-6D) and the predicted ones was described for: tumour risk according to D'Amico's classification [36] (low, intermediate and high); treatment (radical prostatectomy, intensity-modulated radiation therapy, brachytherapy); and ranges of utility values (≤ 0.75, 0.75 -0.89, 0.89 - 0.94, 0.94 - 0.98, 0.98 – 1.00), following the Bremner et al strategy [18].

In addition, the whole study was conducted using good practices guidelines for mapping studies [13], and a tool for reporting a full assessment of the quality and relevance of the mapping study was completed: the MApping onto Preference-based measures reporting Standards (MAPS statement) [37]. The MAPS checklist is shown in Supplementary Table 1.

RESULTS

Demographic and clinical characteristics of both the development cohort (n=266) and the validation cohort (n=441) are summarized in Table 1. Mean (SD) age at baseline was 67.4 (8.1) and 65.5 (6.3) years, most participants were married (75.6% and 82.7%) and retired (64.3% and 57.4%), respectively. In the development cohort, intensity-modulated radiation therapy (IMRT) was the most frequent treatment (46.2%), followed by radical prostatectomy (30.8%) and brachytherapy (16.9%). In the validation cohort the pattern was different, being brachytherapy the most frequent treatment modality (37.7%) and active surveillance the least frequent one (17.0%). Differences between cohorts were statistically significant for all clinical variables (PSA, Gleason, tumour risk, TNM stage and treatment).

Figure 1 presents the histogram of the results from PORPUS-U index in the development cohort at 12 months after treatment, which was skewed to the left, and of their logarithmic transformation.

Table 2 shows mean (SD) of EPIC and SF Health Surveys. EPIC scores indicated, in general, more impairment in the development than in the validation cohort, specially in the sexual function means (26.8 vs 41.7) at 12 months, without great differences between the original and the brief version. In the validation cohort, EPIC scores broadly improved with each follow-up, except for the urinary incontinence (from 85.2 to 83.9) and sexual summary and function scores (from 52.7 to 50.8 and from 41.7 to 37.1, respectively). SF Health Surveys mean scores also indicated slightly more impairment in the development cohort than in the validation cohort, where they remained quite similar throughout the follow-ups.

Table 3 shows six models constructed with data collected at 12 months after treatment in the development cohort. Full linear regression models with all β coefficients are detailed in Supplementary Tables 2A and 2B. Models constructed with scores from EPIC-50 and SF Health Surveys had higher R^2 (0.884 and 0.871) than models constructed with EPIC-26 and SF Health Surveys (0.844 and 0.827). Models constructed only with EPIC had a slightly low R^2 (0.842 for EPIC-50 and 0.776 for EPIC-26). Models constructed only with SF Health Surveys are not reported due to the low R^2 (0.439 for SF-36v2 and 0.457 for SF-12v2). The retransformed RMSE of the models contructed with EPIC-50 were the lowest values of all the mapping models (0.027 and 0.033). The ICC was excellent (ICC $\geq$ 0.90) in all models obtained with EPIC and SF Health Surveys, and slightly low for models constructed without the latter (0.875 and 0.883). No statistically significant differences were found between predicted and observed values.

Table 4 shows the agreement between the PORPUS-U index estimated through mapping algorithms and the index obtained directly with PORPUS-U in the validation cohort at 12, 24 and 36 months after treatment. In all follow-ups, models obtained with data from EPIC-50 and SF Health Surveys presented better predictive capacities than models obtained with EPIC-26 and SF Health Surveys. Specifically, the model 'EPIC-50&SF-36v2' presented consistently the best results while models constructed only with EPIC presented the worst (higher RMSE and lower ICC). Almost all models showed their worst indicators at month 12 and the best ones at month 36. Predicted PORPUS-U means were slightly higher than those observed in all mapping predictive models and follow-ups, and the predicted SD was smaller than what was observed. Nonetheless, there was no statistical significance found in models with EPIC-26 alone nor combined

with SF Health Surveys throughout the follow-ups at 12 and 24 months. At 36 months after treatment, differences between observed and predicted medians were not significant either, except for the model 'EPIC-50&SF-12v2'.

Supplementary Figure 1 presents the Bland-Altman plots with the observed and predicted PORPUS-U values in the development and the validation cohorts. The agreement found between them was high (the difference between observed and predicted values is close to 0 in almost all values) in both cohorts with all models, as almost every value is within the 95% agreement band.

Figure 2 shows the mean of (observed and predicted) PORPUS-U and the SF-6D index per known groups with model 'EPIC-50&SF-36v2' in the development cohort. Observed and predicted means of PORPUS-U index were very similar for all subgroups, except for the subgroup with the lowest range of utility values (0, 0.75). SF-6D index means were lower than PORPUS-U index for all subgroups, but with wider 95% CI. Trends of SF-6D among known groups were similar to those presented by PORPUS-U index: lowest utilities in high tumour risk, in IMRT treatment, and in the lowest range of PORPUS utilities.

DISCUSSION

This study has mapped the PORPUS-U index from the EPIC with and without the SF Health Surveys, which are the most widely used PRO instruments in prostate cancer patients, showing good estimators, and validating the models proposed in a cohort that differed with the development cohort in clinical and tumour variables. The estimation of the PORPUS-U index seemed to be better with those models constructed using EPIC-50 and SF Health Surveys scores, which explained more than 87% of the variance.

Nevertheless, the differences with the other two models constructed with EPIC-26 (also combined with the SF Health Surveys) are small, explaining 84% and 83% of the variance, respectively. Models constructed with only the EPIC-50 or EPIC-26 explained less variance (84% and 77%, respectively) than those which included the SF Health Surveys.

All the models constructed presented better predictive capacities than the prostate cancer mapping studies previously published [14-16, 18]. The first study mapped two disease-specific instruments (FACT-P and EORTC QLQ-C30) to the generic instrument EQ-5D [14], and explained 58% of the variance. Two other studies aimed at mapping only the FACT-P to the EQ-5D [15, 16], obtaining also an unacceptable predictive capacity (lower than 0.70). In all three cases, the low variance explained could be due to two reasons. First, the constructs measured by the source instruments and those measured by the target instrument are quite different due to their scope (disease-specific vs generic). It is important to highlight that the validity of a mapping study depends on the assumption that the two instruments assess the same or closely similar constructs [17]. The second reason may be because the study population was too specific: metastatic hormone-refractory [14] and metastatic castration-resistant prostate cancer patients [15, 16]. These patients usually present poorer health status [38], and mapping models fail to predict low utilities [18, 39].

The only previous study which mapped the PORPUS-U index explained 72% of the variance, indicating good predictive capacity [18]. This model was constructed with the UCLA-PCI, a prostate cancer-specific instrument, in a sample of patients at different disease stages. However, since PORPUS has generic and disease-specific domains, a

mapping to this instrument should also include generic measures, in order to give sturdiness to a predictive model and not undermine it [39]. On the one hand, algorithms constructed with only the EPIC-50 or EPIC-26 in our study explained some more variance (>77%) than the previous one constructed with UCLA-PCI [18]. On the other hand, low variance was explained by models constructed only with SF Health Surveys, supporting the need of disease-specific PROs as source instrument. Models constructed with EPIC and SF Health Surveys seem to overlap well with the PORPUS domains, covering all the important dimensions of the target instrument, which is necessary so as not to undermine the predictive model [39]. The variances explained by these models in our study (from 82.7% to 88.4%) support this good overlap, on which the strength of the mapping function is based [40].

Regarding the predictive capacity of the models in the validation cohort, models with EPIC-50 and SF Health Surveys presented the best results at 24 and 36 months of follow-up after treatment (RMSE<0.026 and ICC≥0.90). Models constructed with EPIC-26 and SF Health Surveys also presented good predictive capacities (RMSE <0.037 and ICC >0.80), but always slightly lower compared with those obtained with EPIC-50. Models constructed only with EPIC presented the lowest agreement (ICC>0.72), but validity results are good enough to predict PORPUS-U values. Furthermore, results at 12 months, even when being the worst in this study, are quite better than other estimations when mapping generic utility instruments from disease-specific profiles [40]. Moreover, statistics of central tendency for the observed and predicted values are quite similar, though the predicted mean is slightly higher for all models and for almost all time periods. The levels of variance are lower than the original observed values, in accordance with the literature [40], as an effect of the

regression to the mean [17]. It is important to highlight this good predictive capacity in the validation cohort, despite the differences in clinical characteristics with the development cohort.

Finally, observed and predicted utility means were very similar for subgroups of tumour risk and treatment. Nonetheless, the subgroups analysis defined by utility values confirmed the overestimation for patients with low observed utilities [40]. Overestimation by predicting very low utilities has consequences for decision analyses, since patients in poor health will be assumed to be in a better state than they actually are [40]. However, only 5 patients (1.9%) in our study had very low observed PORPUS-U values, and the upward bias in prediction of low utilities had a small effect on the overall mean utilities obtained by the present mapping models. In fact, a small number of subjects with low utility value is expected according to the current epidemiology of prostate cancer, mainly diagnosed in early stages and at younger ages.

The SF-6D also distinguished among subgroups of tumour risk, treatment and ranges of utility values, but with a lower precision than PORPUS-U. SF-6D utility values were always clearly lower than those estimated with PORPUS-U in this study. It is congruent with the available literature, because different preference-based measures have been shown to generate different values in the same sample of patients [40, 41]. Both indexes can contribute substantially to the task of improving utility measurement in patients with prostate cancer, but patient preferences obtained with PORPUS-U provide a better understanding of the burden of the disease and the treatment side effects [8].

Bremner et al. suggested constructing a mapping predictive model with longitudinal data of EPIC [18], though the results of their previous mapping to PORPUS-U index

were acceptable, because UCLA-PCI is less frequently used than EPIC. This suggestion is now reinforced by the International Consortium for Health Outcomes Measurement's selection of EPIC-26 together with a generic instrument as part of their prostate cancer standard set of outcomes [19]. Given this intrument's widespread and common use, the mapping algorithm constructed in the present study could be very useful to estimate utility values from EPIC scores. An accurate prediction of PORPUS-U values should include scores from EPIC and SF Health Surveys, whenever possible. These mapping algorithms could improve the assessment of prostate cancer symptoms, disease stages and the treatment comparisons in terms of economic evaluations, especially when no utility instrument has been administered. Nonetheless, administering the utility instrument itself is the best way to obtain PORPUS utilities when the slight extra burden is assumable.

Nevertheless, the present study had some limitations. Firstly, data used for both the development and the validation cohort came from Spanish patients who answered Spanish versions of PRO instruments, and results may not be generalizable to other countries. Therefore, further external validation of these mapping models should be carried out using data of cohorts from other countries (including other country/language versions), as well as with prostate cancer patients at more advanced stages. Secondly, the observed range of PORPUS-U values were near the upper limit (with a high ceiling effect), and low utilities could be overestimated. However, since patients included in these cohorts have undergone a variety of widely used treatment modalities, the observed range of utilities could indeed be representative of most patients with localized or locally advanced prostate cancer.

CONCLUSIONS

The models constructed with different versions of EPIC and SF Health Surveys presented a high predictive capacity to estimate the PORPUS-U values. The algorithms developed allowed us to obtain prostate cancer-specific preference values, which will facilitate cost-utility analyses of prostate cancer treatments, and therefore provide clinicians and health care planners with clear information to make evidence-based decisions.

ACKNOWLEDGMENTS

We acknowledge the collaboration of the other members of the Multicentric Spanish Group of Clinically Localized Prostate Cancer for their contributions. Moreover, the authors would like to thank Aurea Martín for her support in English editing, proofreading, and preparing this manuscript for submission.

REFERENCES

1. Bray F, Ferlay J, Soerjomataram I, et al. Global Cancer Statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin. 2018; 68(6):394-424.
2. Shao YH, Demissie K, Shih W, et al. Contemporary risk profile of prostate cancer in the United States. J Natl Cancer Inst. 2009; 101:1280–3.
3. Chou R, Croswell JM, Dana T, et al. Screening for prostate cancer: a review of the evidence for the U.S. Preventive Services Task Force. Ann Intern Med 2011; 155(11):762–71.
4. Ávila M, Patel L, López S, et al. Patient-reported Outcomes After Treatment for Clinically Localized Prostate Cancer: A Systematic Review and Meta-Analysis. Cancer Treat Rev. 2018; 66:23-44.
5. Hamdy FC, Donovan JL, Lane JA, Mason M, Metcalfe C, Holding P et al. 10-Year Outcomes after Monitoring, Surgery, or Radiotherapy for Localized Prostate Cancer. N Engl J Med. 2016; 375(15):1415–24.
6. Donovan JL, Hamdy FC, Lane JA, et al. Patient-Reported Outcomes after Monitoring, Surgery, or Radiotherapy for Prostate Cancer. N Engl J Med. 2016; 375(15):1425–37.
7. Fitzpatrick R, Fletcher A, Gore S, et al. Quality of life measures in health care. I: Applications and issues in assessment. BMJ 1992; 305(6861):1074–7.
8. Krahn M, Ritvo P, Irvine J, et al. Construction of the Patient-Oriented Prostate Utility Scale (PORPUS): a multiattribute health state classification system for prostate cancer. J Clin Epidemiol. 2000; 53(9):920–30.

9. Volk RJ et al. Preferences of husbands and wives for outcomes of prostate cancer screening and treatment. J Gen Intern Med. 2004;19(4):339–48.
10. Tomlinson G, Bremner KE, Ritvo P, et al. Development and validation of a utility weighting function for the patient-oriented prostate utility scale (PORPUS). Med Decis Making. 2012; 32(1):11–30.
11. Ferrer M, Guedea F, Suárez JF, et al. Quality of life impact of treatments for localized prostate cancer: cohort study with a 5 year follow-up. Radiother. Oncol. 2013; 108(2):306-13.
12. Roeloffzen EM, Lips IM, van Gellekom MP, et al. Health-related quality of life up to six years after (125)I brachytherapy for early-stage prostate cancer. Int. J. Radiat. Oncol. Biol. Phys. 2010; 76(4):1054-60.
13. Wailoo AJ, Hernández-Álava M, Manca A, et al. Mapipng to Estimate Health-State Utility from Non-Preference-Based Outcome Measures: An ISPOR Good Practices for Outcomes Research Task Force Report. Value Health 2017; 20(1):18-27.
14. Wu EQ, Mulani P, Farrel MH, Sleep D. Mapping FACT-P and EORTC QLQ-C30 to patient health status measured by EQ-5D in metastatic hormone-refractory prostate cancer patients. Value Health. 2007; 10(5):408-14.
15. Skaltsa K, Longworth L, Ivanescu C, et al. Mapping the FACT-P to the preference-based EQ-5D questionnaire in metastatic castration-resistant prostate cancer. Value Health. 2014; 17(2):238-44.
16. Diels J, Hamberg P, Ford D, et al. Mapping FACT-P to EQ-5D in a large cross-sectional study of metastatic castration-resistant prostate cancer patients. Qual Life Res. 2015; 24(3):591-8.

17. Fayers PM, Hays RD. Should linking should replace regression when mapping from profile to preference-based measures? Value Health. 2014; 17(2):261–5.
18. Bremner KE, Mitsakakis N, Wilson L, Krahn MD. Predicting utility scores for prostate cancer: mapping the Prostate Cancer Index to the Patient-Oriented Prostate Utility Scale (PORPUS). Prostate Cancer Prostatic Dis. 2014; 17(1):47–56.
19. Martin NE, Massey L, Stowell C, et al. Defining a standard set of patient-centered outcomes for men with localized prostate cancer. Eur Urol. 2015; 67(3):460-7.
20. Evans SM, Millar JL, Moore CM, et al. Cohort profile: the TrueNTH Global Registry - an international registry to monitor and improve localised prostate cancer health outcomes. BMJ Open 2017;7:e017006.
21. Ritvo P, Irvine J, Naglie G, et al. Reliability and validity of the PORPUS, a combined psychometric and utility-based quality-of-life instrument for prostate cancer. J Clin Epidemiol. 2005; 58(5):466-74.
22. Krahn M, Bremner KE, Tomlinson G. Responsiveness of disease-specific and generic utility instruments in prostate cancer patients. Qual Life Res. 2007;16(3):509-22.
23. Ávila M, Pardo Y, Castells M, et al. Adaptation and validation of the Spanish version of the Patient-Oriented Prostate Utility Scale (PORPUS). Qual Life Res. 2014; 23(9):2481–7.
24. Wei JT, Dunn RL, Litwin MS, et al. Development and validation of the expanded prostate cancer index composite (EPIC) for comprehensive

assessment of health-related quality of life in men with prostate cancer. Urology 2000; 56(6):899–905.

25. Szymanski KM, Wei JT, Dunn RL, Sanda MG. Development and validation of an abbreviated version of the Expanded Prostate Cancer Index composite instrument for measuring health-related quality of life among prostate cancer survivors. Urol. 2010; 76(5):1245–50.
26. Ferrer M, Garin O, Pera J, et al. Evaluation of the quality of life of patients with localized prostate cancer: validation of the Spanish version of the EPIC. Med Clin (Barc) 2009; 132(4):128–35.
27. McHorney CA, Ware JE Jr, Lu JF, Sherbourne CD. The MOS 36-item Short-Form Health Survey (SF-36): III. Tests of data quality, scaling assumptions, and reliability across diverse patient groups. Med Care. 1994; 32(1):40–66.
28. Ferrans CE, Powers MJ. Quality of life index: development and psychometric properties. ANS Adv Nurs Sci. 1985; 8(1):15–24.
29. Ware J Jr, Kosinski M, Keller SD. A 12-Item Short-Form Health Survey: construction of scales and preliminary tests of reliability and validity. Med Care. 1996; 34(3):220–33.
30. Brazier J, Roberts J, Deverill M. The estimation of a preference-based measure of health from the SF-36. J Health Econ. 2002; 21(2):271-92.
31. Alonso J, Prieto L, Antó JM. [The Spanish version of the SF-36 Health Survey (the SF-36 health questionnaire): an instrument for measuring clinical results]. Med Clin (Barc). 1995; 104(20):771-6.

32. Vilagut G, Ferrer M, Rajmil L, et al. [The Spanish Version of the Short Form 36 Health Survey: A Decade of Experience and New Developments]. Gac Sanit 2005, 19(2):135-50.
33. Duan N. Smearing estimate: a nonparametric retransformation method. J Am Stat Assoc 1983; 78:605-610.
34. Browne C, Brazier J, Carlton J, et al. Estimating quality-adjusted life years from patient-reported visual functioning. Eye (Lond). 2012; 26(10):1295-301.
35. Portney LG, Watkins MP. Foundations of clinical research: applications to practice. New Jersey: Prentice Hall; 2000.
36. D'Amico AV, Whittington R, Malkowicz SB, et al. Biochemical outcome after radical prostatectomy, external beam radiation therapy, or interstitial radiation therapy for clinically localized prostate cancer. JAMA 1998; 280(11):969-974.
37. Petrou S, Rivero-Arias O, Dakin H, et al. Preferred reporting items for studies mapping onto preference-based outcome measure: The MAPS statement. Health Qual Life Out 2015; 13(1):106.
38. Downing A, Wright P, Hounsome L, et al. Quality of life in men living with advanced and localised prostate cancer in the UK: a population-based study. Lancet Oncol 2019; 20(3):436-47.
39. Barton GR, Sach TH, Jenkinson C, et al. Do estimates of cost-utility based on the EQ-5D differ from those based on the mapping of utility scores? Health Qual Life Outcomes 2008; 6:51.
40. Brazier JE, Yang Y, Tsuchiya A, Rowen DL. A review of studies mapping (or cross walking) non-preference based measures of health to generic preference-based measures. Eur J Health Econ 2010; 11(2):215–25.

41. Feeny D, Wu L, Eng K. Comparing short form 6D, standard gamble, and health utilities index Mark 2 and Mark 3 utility scores: results from total hip arthroplasty patients. Qual Life Res 2004; 13(10):1659-70.

Table 1. Demographic and clinical characteristics of the participants in each cohort.

	Development cohort (n=266)	Validation cohort (n=441)	p-value
Follow-up, n			
12 months	266	429	---
24 months		427	
36 months		364	
Age (years), mean (SD)	67.4 (8.1)	65.5 (6.3)	0.001
Marital status, n (%)			
Single	16 (6.0%)	18 (4.1%)	0.137
Married	201 (75.6%)	365 (82.7%)	
Widowed	16 (6.0%)	14 (3.2%)	
Separated/Divorced	18 (6.8%)	28 (6.3%)	
Missing	*15 (5.6%)*	*16 (3.6%)*	
Employment, n (%)			
Active	67 (25.2%)	138 (31.3%)	0.088
Unemployed	4 (1.5%)	17 (3.8%)	
Retired	171 (64.3%)	253 (57.4%)	
Other	9 (3.4%)	16 (3.6%)	
Missing	*15 (5.6%)*	*17 (3.9%)*	
Body mass index, n (%)			
Underweight	3 (1.1%)	3 (0.7%)	0.109
Normal weight	107 (40.2%)	209 (47.4%)	
Overweight	76 (28.6%)	135 (30.6%)	

Obesity	66 (24.8%)	80 (18.1%)	
Missing	*14 (5.3%)*	*14 (3.2%)*	
PSA, mean (SD)	9.6 (10.5)	6.1 (1.9)	< 0.001
Gleason, mean (SD)	6.8 (0.9)	6.1 (0.5)	< 0.001
Tumour risk, n (%)			
Low	84 (31.7%)	369 (85.0%)	< 0.001
Intermediate	110 (41.5%)	65 (15.0%)	
High	71 (26.8%)	0 (0.0%)	
Tumoral stage, n (%)			
T0	1 (0.3%)	0 (0.0%)	< 0.001
T1	128 (48.1%)	372 (84.4%)	
T2	90 (33.8%)	69 (15.6%)	
T3	46 (17.3%)	0 (0.0%)	
Tx	0 (0.0%)	0 (0.0%)	
Missing	*1 (0.4%)*	*0 (0.0%)*	
N, n (%)			
N0	223 (83.8%)	256 (58.0%)	< 0.001
N1	4 (1.5%)	1 (0.2%)	
Nx	39 (14.7%)	184 (41.7%)	
M, n (%)			
M0	186 (69.9%)	245 (55.6%)	< 0.001
M1	0 (0.0%)	0 (0.0%)	
Mx	80 (30.1%)	196 (44.4%)	
Treatment, n(%)			

Radical prostatectomy	82 (30.8%)	93 (21.1%)	< 0.001
IMRT	123 (46.2%)	82 (18.6%)	
Brachytherapy	45 (16.9%)	183 (41.5%)	
Active surveillance	3 (1.1%)	75 (17.0%)	
Other radiotherapy	13 (4.9%)	7 (1.6%)	
Missing	*0 (0.0%)*	*1 (0.2%)*	

p-value was calculated using χ^2 test for categorical variables and t-Student for numerical variables.

IMRT: Intensity Modulated Radiation Therapy; **PSA:** Prostate-Specific Antigen; **SD:** Standard Deviation.

Figure 1. Distribution of PORPUS-U index and its logarithmic transformation in the development cohort at 12 months after treatment.

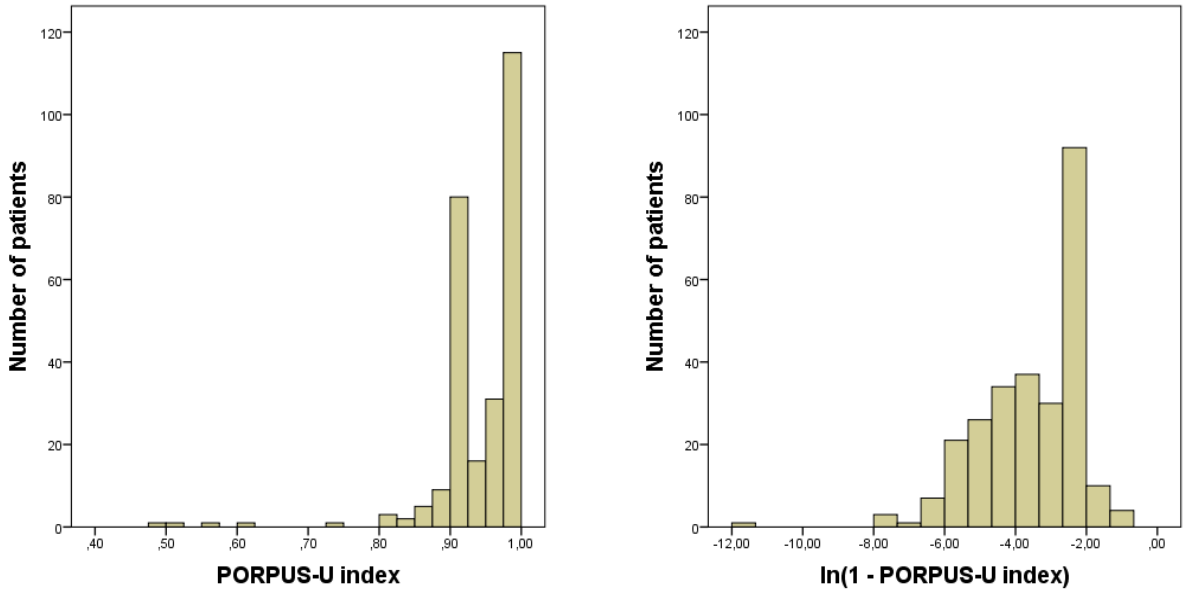


Table 2. Distribution of EPIC-50, EPIC-26, SF-36v2 and SF-12v2 scores in both cohorts: mean (SD).

	Development Cohort		Validation Cohort					
	12 months (n=266)		12 months (n=429)		24 months (n=427)		36 months (n=364)	
	EPIC-50	**EPIC-26**	**EPIC-50**	**EPIC-26**	**EPIC-50**	**EPIC-26**	**EPIC-50**	**EPIC-26**
Urinary Incontinence	78.9 (28.9)	78.9 (28.9)	85.2 (24.2)	85.2 (24.2)	85.7 (23.5)	85.7 (23.5)	83.9 (24.9)	83.9 (24.9)
Urinary Irritative/Obstructive	84.0 (18.1)	82.4 (21.0)	88.7 (18.3)	87.5 (20.1)	91.8 (14.1)	90.8 (16.2)	92.1 (13.2)	91.0 (15.7)
Sexual Summary	43.7 (20.5)	38.2 (28.5)	52.7 (21.4)	51.8 (28.5)	51.3 (21.7)	50.1 (29.7)	50.8 (20.7)	49.2 (29.1)
Sexual Bother	81.8 (26.6)		77.5 (27.7)		79.4 (26.8)		81.8 (26.1)	
Sexual Function	26.8 (28.4)		41.7 (26.4)		38.8 (26.9)		37.1 (26.7)	
Bowel Summary	92.0 (14.3)	88.6 (20.5)	95.3 (10.2)	93.3 (15.0)	97.2 (7.6)	96.3 (11.1)	97.4 (7.9)	96.2 (12.0)
Bowel Bother	89.1 (20.0)		93.7 (14.3)		96.5 (10.4)		96.4 (11.3)	
Bowel Function	94.9 (9.4)		96.8 (7.2)		98.0 (5.6)		98.3 (5.3)	
Hormonal Summary	83.0 (19.3)	80.5 (23.4)	91.3 (12.7)	89.7 (15.9)	92.0 (12.3)	90.4 (15.5)	92.0 (12.2)	91.2 (15.4)
Hormonal Bother	83.6 (19.8)		91.4 (13.3)		92.0 (12.9)		92.6 (12.9)	

Hormonal Function	82.3 (19.3)		91.1 (12.7)		91.9 (12.1)		91.3 (12.5)	
	SF-36v2	**SF-12v2**	**SF-36v2**	**SF-12v2**	**SF-36v2**	**SF-12v2**	**SF-36v2**	**SF-12v2**
Physical Function	43.4 (10.9)	44.3 (13.8)	47.3 (9.0)	48.8 (11.8)	46.4 (9.3)	47.9 (12.2)	46.4 (9.1)	48.5 (12.2)
Role Physical	49.9 (6.8)	51.8 (9.0)	51.7 (5.8)	53.4 (7.5)	51.7 (6.0)	53.5 (7.9)	51.8 (5.6)	53.8 (7.3)
Bodily Pain	47.4 (7.6)	53.0 (9.2)	49.5 (7.1)	54.9 (7.4)	49.2 (7.2)	54.6 (7.9)	49.0 (6.5)	55.0 (7.4)
General Health	46.8 (6.4)	43.7 (9.5)	49.2 (7.3)	45.7 (8.5)	48.8 (7.3)	45.8 (9.2)	49.0 (7.0)	46.4 (9.1)
Vitality	46.7 (9.6)	42.1 (10.6)	49.2 (8.3)	44.4 (9.7)	49.2 (8.8)	44.6 (10.0)	48.9 (8.8)	44.1 (9.6)
Social Functioning	51.3 (10.3)	53.8 (8.0)	52.4 (8.3)	54.5 (6.6)	52.2 (8.8)	54.6 (6.7)	52.4 (8.3)	54.6 (6.7)
Role Emotional	49.3 (7.4)	49.5 (10.2)	50.2 (6.7)	51.1 (9.1)	50.4 (6.9)	51.3 (9.4)	50.7 (6.3)	51.6 (8.7)
Mental Health	54.1 (9.2)	53.9 (11.5)	55.0 (9.0)	55.3 (10.6)	55.5 (8.9)	55.6 (10.7)	55.8 (9.0)	56.0 (10.8)

EPIC-26: Expanded Prostate cancer Index Composite-26 items; **EPIC-50:** Expanded Prostate cancer Index Composite-50 items; **SD:** Standard Deviation; **SF-12v2:** 12-items Short Form Health Survey; **SF-36v2:** 36-items Short Form Health Survey.

Table 3. Predictive capacity of models constructed with EPIC with or without SF Health Surveys.

Models	EPIC-50&SF-36v2	EPIC-50&SF-12v2	EPIC-50	EPIC-26&SF-36v2	EPIC-26&SF-12v2	EPIC-26
R^2	0.884	0.871	0.842	0.844	0.827	0.776
RMSE (model)	0.496	0.525	0.579	0.577	0.606	0.690
RMSE (retransformation)	0.027	0.033	0.039	0.034	0.037	0.039
ICC	0.953	0.917	0.875	0.922	0.911	0.883
95% Confidence Interval	0.940 ; 0.963	0.894 ; 0.935	0.841 ; 0.902	0.901 ; 0.939	0.887 ; 0.930	0.851 ; 0.908
Observed mean (SD)	0.944 (0.059)	0.945 (0.054)	0.944 (0.066)	0.946 (0.061)	0.946 (0.063)	0.944 (0.066)
Predicted mean (SD)	0.944 (0.059)	0.945 (0.054)	0.946 (0.049)	0.946 (0.061)	0.946 (0.063)	0.949 (0.053)
Observed median [IQR]	0.961 [0.916 ; 0.990]	0.961 [0.916 ; 0.990]	0.961 [0.916 ; 0.990]	0.961 [0.916 ; 0.990]	0.961 [0.916 ; 0.990]	0.961 [0.916 ; 0.990]
Predicted median [IQR]	0.958 [0.917 ; 0.991]	0.959 [0.916 ; 0.990]	0.958 [0.912 ; 0.990]	0.960 [0.921 ; 0.989]	0.956 [0.928 ; 0.990]	0.956 [0.929 ; 0.990]

Wilcoxon test value	0.151	0.304	0.103	0.717	0.508	0.452

EPIC-26: Expanded Prostate cancer Index Composite-26 items; **EPIC-50:** Expanded Prostate cancer Index Composite-50 items; **ICC:** Intraclass Correlation Coefficient; **IQR:** Interquartile Range; **RMSE:** Root Mean Square Error; **SD:** Standard Deviation; **SF-12v2:** 12-items Short Form Health Survey; **SF-36v2:** 36-items Short Form Health Survey.

Table 4. Agreement between observed PORPUS-U values and those predicted by the mapping models in the validation cohort.

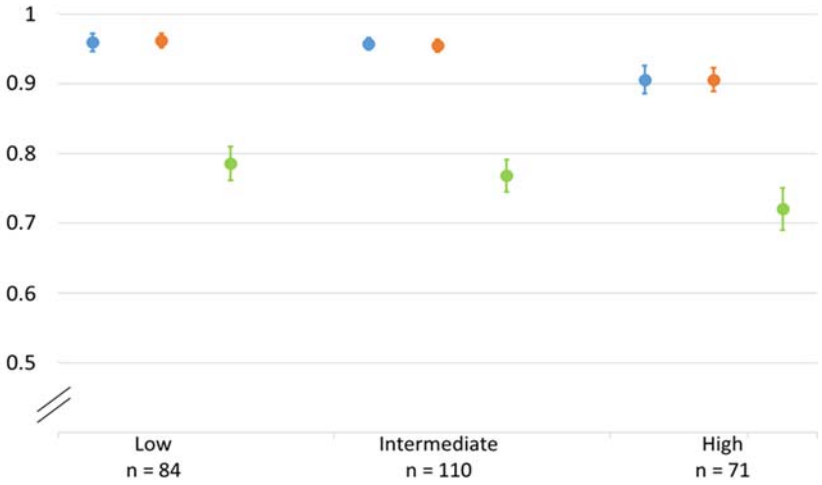
Models	EPIC-50&SF-36v2	EPIC-50&SF-12v2	EPIC-50	EPIC-26&SF-36v2	EPIC-26&SF-12v2	EPIC-26
12 months						
RMSE	0.026	0.030	0.036	0.039	0.036	0.040
ICC	0.923	0.871	0.744	0.815	0.809	0.733
95% confidence interval	0.907 ; 0.937	0.844 ; 0.894	0.691 ; 0.789	0.776 ; 0.847	0.769 ; 0.842	0.678 ; 0.779
Observed mean (SD)	0.968 (0.052)	0.968 (0.052)	0.968 (0.052)	0.968 (0.052)	0.968 (0.052)	0.968 (0.052)
Predicted mean (SD)	0.972 (0.042)	0.973 (0.035)	0.974 (0.033)	0.973 (0.032)	0.970 (0.039)	0.972 (0.033)
Observed median [IQR]	0.988 [0.958 ; 0.995]	0.988 [0.958 ; 0.995]	0.988 [0.958 ; 0.995]	0.988 [0.958 ; 0.995]	0.988 [0.958 ; 0.995]	0.988 [0.958 ; 0.995]
Predicted median [IQR]	0.989 [0.963 ; 0.997]	0.989 [0.963 ; 0.996]	0.990 [0.963 ; 0.997]	0.988 [0.958 ; 0.997]	0.986 [0.956 ; 0.996]	0.987 [0.954 ; 0.996]
Wilcoxon test value	0.001	0.001	0.001	0.145	0.684	0.849
24 months						
RMSE	0.026	0.026	0.032	0.037	0.033	0.036
ICC	0.915	0.899	0.767	0.843	0.839	0.772
95% confidence interval	0.897 ; 0.930	0.877 ; 0.916	0.718 ; 0.807	0.810 ; 0.870	0.805 ; 0.867	0.725 ; 0.812
Observed mean (SD)	0.966 (0.050)	0.966 (0.050)	0.966 (0.050)	0.966 (0.050)	0.966 (0.050)	0.966 (0.050)
Predicted mean (SD)	0.969 (0.045)	0.970 (0.037)	0.973 (0.032)	0.971 (0.036)	0.969 (0.038)	0.972 (0.033)

Observed median [IQR]	0.988 [0.953 ; 0.995]	0.988 [0.953 ; 0.995]	0.988 [0.953 ; 0.995]	0.988 [0.953 ; 0.995]	0.988 [0.953 ; 0.995]	0.988 [0.953 ; 0.995]
Predicted median [IQR]	0.988 [0.957 ; 0.996]	0.988 [0.954 ; 0.996]	0.989 [0.959 ; 0.996]	0.985 [0.954 ; 0.996]	0.986 [0.953 ; 0.996]	0.986 [0.955 ; 0.996]
Wilcoxon test value	0.015	0.001	0.000	0.061	0.269	0.049
36 months						
RMSE	0.022	0.024	0.028	0.037	0.035	0.037
ICC	0.940	0.906	0.723	0.801	0.872	0.736
95% confidence interval	0.926 ; 0.951	0.885 ; 0.924	0.659 ; 0.774	0.756 ; 0.838	0.842 ; 0.896	0.676 ; 0.785
Observed mean (SD)	0.966 (0.047)	0.966 (0.047)	0.966 (0.047)	0.966 (0.047)	0.966 (0.047)	0.966 (0.047)
Predicted mean (SD)	0.967 (0.047)	0.969 (0.038)	0.970 (0.034)	0.970 (0.032)	0.969 (0.038)	0.971 (0.032)
Observed median [IQR]	0.987 [0.947 ; 0.995]	0.987 [0.947 ; 0.995]	0.987 [0.947 ; 0.995]	0.987 [0.947 ; 0.995]	0.987 [0.947 ; 0.995]	0.987 [0.947 ; 0.995]
Predicted median [IQR]	0.986 [0.949 ; 0.996]	0.986 [0.949 ; 0.996]	0.988 [0.950 ; 0.995]	0.984 [0.953 ; 0.996]	0.983 [0.953 ; 0.996]	0.983 [0.955 ; 0.996]
Wilcoxon test value	0.750	0.023	0.076	0.243	0.287	0.118

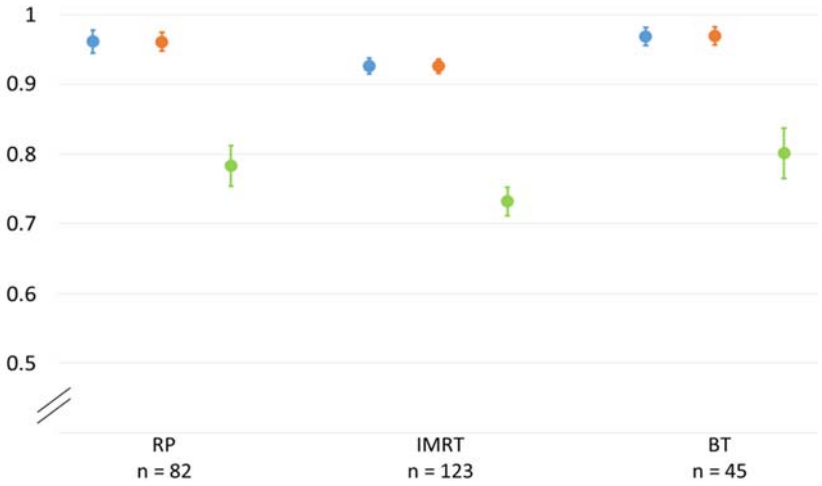
EPIC-50: Expanded Prostate cancer Index Composite-50 items; **EPIC-26:** Expanded Prostate cancer Index Composite-26 items; **SF-36v2:** 36-items Short Form Health Survey; **SF-12v2:** 12-items Short Form Health Survey; **RMSE:** Root Mean Square Error; **ICC:** Intraclass Correlation Coefficient; **SD:** Standard Deviation; **IQR:** Interquartile Range.

Figure 2. Mean (95% confidence interval) of observed PORPUS-U (blue) and predicted PORPUS-U index (orange) by tumour risk, treatment and utility ranges with model 'EPIC-50&SF-36v2' in the development cohort. Results of SF-6D index are shown in green color.

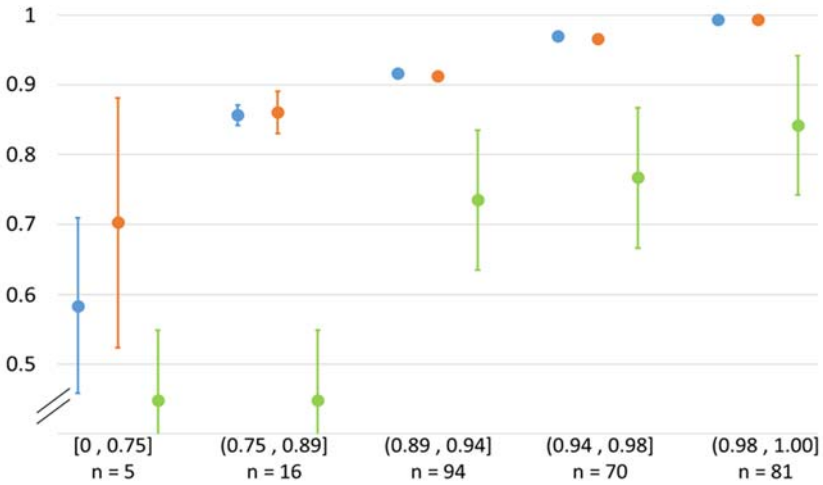
Tumour risk



Treatment



Utility value ranges



BT: Brachytherapy; IMRT: Intensity-Modulated Radiation Therapy; RP: Radical Prostatectomy.

SUPPLEMENTARY MATERIAL

Supplementary Figure 1. Bland-Altman plots between observed and predicted PORPUS-U values in the development and the validation cohorts.

Supplementary Figure 2. Mean (95% confidence interval) of observed PORPUS-U (blue) and predicted PORPUS-U index (orange) by tumour risk, treatment and utility ranges with model 'EPIC-50&SF-36v2' in the validation cohort. Results of SF-6D index are shown in green color. BT: Brachytherapy; IMRT: Intensity-Modulated Radiation Therapy; RP: Radical Prostatectomy

Supplementary Table 1. Mapping onto Preference-based measures reporting Standards (MAPS) statement for the mapping' study.

Supplementary Table 2A. Linear regression models constructed with log(1-PORPUS-U) as dependent variable and scores from the EPIC-50 and original or brief versions of the SF Health Surveys as independent variables: β coefficients (Standard Error).

Supplementary Table 2B. Linear regression models constructed with log(1-PORPUS-U) as dependent variable and scores from the EPIC-26 and original or brief versions of the SF Health Surveys as independent variables: β coefficients (Standard Error).