Journal of Clinical Epidemiology 164 (2023) 65−75

ORIGINAL ARTICLE

Assessing real-world representativeness of prospective registry cohorts in oncology: insights from patients with esophagogastric cancer

Steven C. Kuijper[a,b], Joost Besseling[a,b], Thomas Klausch[c], Marije Slingerland[d], Charlène J. van der Zijden[e], Ewout A. Kouwenhoven[f], Laurens V. Beerepoot[g], Nadia Haj Mohammad[h], Bastiaan R. Klarenbeek[i], Rob H.A. Verhoeven[a,b,j], Hanneke W.M. van Laarhoven[a,b,*]

[a]Department of Medical Oncology, Amsterdam UMC Location University of Amsterdam, Meibergdreef 9, Amsterdam, The Netherlands
[b]Cancer Center Amsterdam, Cancer Treatment and Quality of Life, Amsterdam, The Netherlands
[c]Department of Epidemiology and Data Science, Amsterdam University Medical Centers, Amsterdam, The Netherlands
[d]Department of Medical Oncology, Leiden University Medical Center, Leiden, The Netherlands
[e]Department of Surgery, Erasmus Medical Center, Rotterdam, The Netherlands
[f]Department of Surgery, Ziekenhuisgroep Twente, Almelo, The Netherlands
[g]Department of Medical Oncology, Elisabeth Tweesteden Ziekenhuis and EMBRAZE Cancer Network, Tilburg, The Netherlands
[h]Department of Medical Oncology, Utrecht University Medical Center, Utrecht, The Netherlands
[i]Department of Surgery, Radboud University Medical Center, Nijmegen, The Netherlands
[j]Department of Research & Development, Netherlands Comprehensive Cancer Organisation (IKNL), Utrecht, The Netherlands

Accepted 16 October 2023; Published online 21 October 2023

Abstract

Objectives: This study aimed to explore the real-world representativeness of a prospective registry cohort with active accrual in oncology, applying a representativeness metric that is novel to health care.

Study Design and Setting: We used data from the Prospective Observational Cohort Study of Esophageal-Gastric Cancer Patients (POCOP) registry and from the population-based Netherlands Cancer Registry (NCR). We used Representativeness-indicators (R-indicators) and overall survival to investigate the degree to which the POCOP cohort and clinically relevant subgroups were a representative sample compared to the NCR database. Calibration using inverse propensity score weighting was applied to correct differences between POCOP and NCR.

Results: The R-indicator of the entire POCOP registry was 0.72 95% confidence interval [0.71, 0.73]. Representativeness of palliative patients was higher than that of potentially curable patients (R-indicator 0.88 [0.85, 0.90] and 0.70 [0.68, 0.71], respectively). Stratification to clinically relevant subgroups based on treatment resulted in higher R-indicators of the respective subgroups. Both after stratification and calibration weighting survival estimates in the POCOP registry were more similar to that in the NCR population.

Conclusion: This study demonstrated the assessment of real-world representativeness of patients who participated in a prospective registry cohort and showed that real-world representativeness improved when the variability in treatment was accounted for.

Keywords: Representativeness; R-indicators; Health-related quality of life; Survival; Esophageal cancer; Gastric cancer

1. Introduction

In oncological research, patient reported outcome measures (PROMs) are a popular method to obtain quality of life and other self-report data from patients [1−5]. When using PROMs, patients are usually requested to complete (digital) questionnaires and data collection thus relies on the active participation of patients. Consequently, some patients may be more inclined to participate than others which could result in a selection bias [6]. For example, patients

* Corresponding author. Department of Medical Oncology, Amsterdam UMC, Office D3-312, PO Box 22660, 1100 DD Amsterdam, The Netherlands. Tel.: +31-020-566-5955; fax: +31-020-444-4355.

E-mail address: h.vanlaarhoven@amsterdamumc.nl (H.W.M. van Laarhoven).

https://doi.org/10.1016/j.jclinepi.2023.10.009

What is new?

Key findings

- This study demonstrated the utility of Representativeness-indicators to explore real-world representativeness of prospective registry cohort as well as calibration techniques and stratification to correct for differences between the prospective registry cohort and the population.

What this adds to what was known?

- This is the first time that Representativeness-indicators were used in the field medical oncology.
- Using R-indicators we found that prospective registry cohorts in oncology can be representative of a nation-wide population when differences in treatment are accounted for.

What is the implication and what should change now?

- Future observational and clinical studies could be evaluated uniformly by using Representativeness-indicators.

with high age and higher levels of comorbidities in self-administered health-related quality of life questionnaires may be less likely to participate [7]. Thus, the question arises to what extent patients who are willing to fill out PROMs questionnaires accurately reflect the real-world oncological patient.

Studies on the representativeness of prospective cohort studies that collect PROMs compared to the population are still scarce. Recently, a study was published investigating the representativeness of the Prospective Dutch Colorectal Cancer cohort with respect to the Dutch population of patients with colorectal cancer [8]. In this study, standardized mean differences (SMD) between the prospective cohort and population were used to identify key differences between cohort and population. Although SMDs can provide valuable insights into differences at the variable-level, it does not provide a single intuitive metric capturing the representativeness of a sample. Moreover, because SMDs are calculated for each variable separately, the relative effects of variables on the sample's representativeness cannot be explored.

The aim of this study was to apply and demonstrate the utility of a metric called Representativeness-indicators (R-indicators) for investigating the real-world representativeness of prospective cohort studies in the field of oncology [9−11]. R-indicators have attractive properties which enable researchers to express a sample's representativeness with respect to the population in a single, intuitive metric and allows for the examination of multiple variables simultaneously which can be used to explore variables' relative effect on the representativeness. To this end, we investigated the real-world representativeness of patients included in the Prospective Observational Cohort Study of Esophageal-Gastric Cancer Patients (POCOP) with respect to the Dutch population of patients with esophagogastric cancer [12]. Additionally, we investigated if potential differences between prospective cohort registries and the population could be corrected to produce externally valid results and conclusions.

2. Methods

2.1. Study population

Data of all patients that were diagnosed between 2016 and 2021 with esophageal or gastric cancer (including gastro-esophageal junction carcinoma) and participated in the POCOP registry were used as the sample data. The methods of the POCOP registry have been reported elsewhere [12]. In short, patients are referred to the investigators of the project by someone from the medical team. POCOP investigators then contact potential participants by phone and send the questionnaire by mail or email. When patients provide written informed consent, they are contacted telephonically, via mail or email, every 3 months during the first year, twice in the second year and annually after that, to collect PROMs.

The reference population consisted of all patients diagnosed between 2016 and 2021 with primary esophageal or gastric cancer in the Netherlands. The data from these patients were obtained from the Netherlands Cancer Registry (NCR). The NCR is an annually updated nation-wide database containing all patients diagnosed with cancer. Trained data managers routinely extract information of the diagnosis, tumor stage and treatment from patients' electronic medical records and add this to the NCR. Identification is mainly based on notification from the nationwide network and registry of histopathology and cytopathology in the Netherlands [13]. Patients in the NCR diagnosed in 2016 were staged using TNM Classification of Malignant Tumours seventh edition, patients diagnosed between 2017 and 2021 were staged using the eighth edition [14]. Patients with TNM-staging based on the eighth edition were converted to the staging of the seventh edition to ensure uniformity of this variable in the analyses. Detailed clinical disease (c-) stages (e.g., IA or IIIB) were simplified to 0, I, II, III, IV, or X, which was then used as categorical variable in all analyses. Finally, all clinical patient data was gathered from the NCR database. By linking the POCOP registry to the NCR we identified which patients participated in POCOP.

2.2. Coding and classifications

Patients were defined as either potentially curable (TNM classification $cT_{x\text{-}4a}$ and cM_0) or palliative (cT_{4B} or cM_1).

Primary tumor location was defined as stomach or esophagus. Gastro-esophageal junction and cardia carcinomas was coded as esophageal carcinoma if the patient underwent a esophagectomy and coded as gastric carcinoma if the patients underwent a gastrectomy.

Treatments were defined for both the curative and palliative intent, for example, chemotherapy (either adjuvant, neoadjuvant or palliative), (definitive) chemoradiotherapy and resection (see Supplementary materials Table 1 for all treatments and frequencies). Definitive chemoradiotherapy and neoadjuvant chemoradiotherapy not followed by resection were distinguished from each other based on the dose of radiation. Patients with a dose of 41.4 Gy or less not followed by resection were assumed to have been treated with neoadjuvant chemoradiotherapy. Patients with a dose higher than 41.4 Gy not followed by resection were assumed to have been treated with definitive chemoradiotherapy [15].

2.3. Real-word representativeness

We explored the real-world representativeness of the POCOP registry using the recently developed representativeness indicators (R-indicators) [9−11]. This method was developed by the Dutch national bureau of statistics, Statistics Netherlands, and has attracted interest in survey studies [16,17]. The R-indicator quantifies sample representativeness between 0 and 1, with 0 indicating that the sample is not representative at all and one that the sample is completely representative for the reference population. The calculation of the R-indicator is based on the variation (i.e., standard deviation) of the propensity that patients participate in the prospective registry cohort, conditional on a set of covariates.

In this analysis using the POCOP registry, response propensities were estimated using a multivariable logistic regression model with all available patient and tumor characteristics as independent variables (treatment, sex, WHO performance status, stage, morphology, age, number of comorbidities, and primary tumor location). In addition to the R-indicator, the so-called partial R-indicators were calculated and reflect which variable in the model contributed the most to the lack of representativeness. The R-indicators and partial R-indicators of the entire POCOP sample with 95% confidence intervals (CI) and the R-indicators for all clinically relevant groups were calculated using the R-indicators code for R [18]. By stratifying the analyses of R-indicators to potentially curable and palliative patients, and treatment groups, we were able to observe if the real-world representativeness was consistent across strata.

2.4. Adjusting for differences

In order to correct potential selection bias in the prospective registry cohort and subgroups thereof, we used a calibration weighting (hereafter referred to as calibration) technique based on the Inverse Propensity Weight (IPW) [19]. By calibrating the prospective registry cohort to the target population, a pseudo-population is created from the prospective registry cohort with which should more accurately reflect the population data. In this analysis, we calculated the IPW of being included in the POCOP registry with the same multivariable logistic model to estimate the response propensities.

To investigate the degree to which calibration was successful in creating a pseudo population which was better reflective of the NCR population data, we performed a survival analysis in which we constructed Kaplan-Meier (KM) curves of patients in the NCR, POCOP and calibrated POCOP registry. This was performed for the POCOP registry as a whole, and for the subgroups based on treatment intent. Bias was defined as the deviation between the KM curves of POCOP vs. NCR data and calibration POCOP vs. NCR data, which was inspected visually. Additionally, median survival and 5-year overall survival was calculated for all analyzed groups. All missing data on variables used in this study were imputed using the random forest imputation implementation of the *missForest* package for R, and the accompanying out-of-bag normalized root mean squared error was reported to the imputation error [20]. Values very close to zero indicate low imputation error. All analyses were conducted in R version 4.1.0 and R studio version 4.0.3.

3. Results

3.1. Patient population

In total, 2,702 patients were available from POCOP and 16,856 from the NCR (which included all POCOP patients) (Table 1); 65% of patients were treated with curative intent, while 35% of patients were treated with palliative intent (Table 2). The covariate balance after calibrating the POCOP database to the NCR can be found in Table 1. The out-of-bag normalized root mean squared error of the imputation was 9.2×10^{-10}.

3.2. Real-world representativeness

Using the R-indicators, we observed that the R-indicator of the total, nonstratified POCOP registry was 0.72 95% CI [0.71, 0.73]. Stratified to treatment intent, the R-indicator was 0.88 [0.86, 0.90] for patients treated with palliative intent and 0.70 [0.68, 0.71] for patients treated with curative intent.

Among patients with esophageal cancer treated with curative intent representativeness of the largest group, neoadjuvant chemoradiotherapy followed by a resection, was 0.88 [0.85, 0.92]. For patients who were treated with neoadjuvant chemoradiotherapy not followed by resection representativeness was 0.80 [0.75, 0.86]. In smaller groups, representativeness was 0.90 [0.85, 0.94] for definitive

Table 1. Descriptive statistics of the NCR population database, the POCOP database and calibrated POCOP database

Variable	NCR (*N* = 16,856)	POCOP (*N* = 2,706)	Calibrated POCOP
Treatment intent			
Potentially curable	10,966 (65.1%)	2,197 (81.2%)	10,907.4 (65.0%)
Palliative	5,890 (34.9%)	509 (18.8%)	5,868.5 (35.0%)
Sex			
Male	11,883 (70.5%)	2,064 (76.3%)	11,626.0 (69.3%)
Female	4,973 (29.5%)	642 (23.7%)	5,149.9 (30.7%)
WHO performance status			
0	5,228 (31.0%)	1,330 (49.2%)	5,289.6 (31.5%)
1	5,700 (33.8%)	989 (36.5%)	5,507.7 (32.8%)
2	1,707 (10.1%)	126 (4.7%)	2,271.2 (13.5%)
>2	656 (3.9%)	21 (0.8%)	504.2 (3.0%)
Unknown	3,565 (21.1%)	240 (8.9%)	3,203.2 (19.1%)
Stage			
1	1,429 (8.5%)	139 (5.1%)	1,276.4 (7.6%)
2	2,771 (16.4%)	570 (21.1%)	2,771.9 (16.5%)
3	4,432 (26.3%)	1,145 (42.3%)	4,454.2 (26.6%)
4	6,742 (40.0%)	768 (28.4%)	6,905.7 (41.2%)
M/X	1,482 (8.8%)	84 (3.1%)	1,367.8 (8.2%)
Histology			
Adenocarcinoma	13,342 (79.2%)	2,217 (81.9%)	13,231.5 (78.9%)
Squamous cell carcinoma	2,919 (17.3%)	459 (17.0%)	2,884.0 (17.2%)
Other/not microscopically verified	595 (3.5%)	30 (1.1%)	660.4 (3.9%)
Age			
< 49	775 (4.6%)	119 (4.4%)	700.2 (4.2%)
50–59	2,199 (13.0%)	449 (16.6%)	2,245.3 (13.4%)
60–69	5,046 (29.9%)	1,069 (39.5%)	4,847.9 (28.%9)
70–79	5,880 (34.9%)	916 (33.9%)	6,129.7 (36.5%)
> 80	2,956 (17.5%)	153 (5.7%)	2,852.8 (17.0%)
Number of comorbidities			
0	8,085 (48.0%)	1,557 (57.5%)	8,181.4 (48.8)
≥1	8,771 (52.0%)	1,149 (42.5%)	8,594.5 (51.2)

Calibrated POCOP refers to the POCOP database after inverse propensity weighting calibration to the NCR population distributions

chemoradiation; 1.00 [0.89, 1.00] for neoadjuvant chemotherapy followed by resection (with or without adjuvant chemotherapy); 0.98 [0.98, 1.00] for other treatments; and 0.88 [0.85, 0.92] for endoscopic resection (Fig. 1). For patients with gastric cancer, representativeness of patients treated with neoadjuvant chemotherapy followed by resection (with or without adjuvant chemotherapy) was 0.84 [0.79, 0.89], and 0.95 [0.89, 1.00] for patients that underwent a resection only (Fig. 1).

Comparable representativeness estimates were found among patients with esophageal cancer treated with palliative intent. In the largest group, patients treated with chemotherapy or target therapy, representativeness was 0.91 [0.88, 0.94]. Patients treated with chemoradiotherapy had a representativeness of 0.84 [0.76, 0.93]; 1.00 [0.97, 1.00] for radiotherapy; and 1.00 [0.78, 1.00] for palliative resection + chemo (radio)therapy or radiotherapy. Among patient with gastric cancer, only patients treated with chemotherapy or targeted therapy was sufficiently large to compute representativeness which was 0.93 [0.88, 0.97].

Across the entire POCOP cohort and the curative and palliative subgroups, the partial R-indicators showed that treatment contributed most to the degree of nonrepresentativeness (Fig. 2).

3.3. Calibration and survival

The KM curves of all patients and conditioned on treatment intent are displayed in Fig. 3. Overall survival was higher for patients in POCOP compared to the NCR population data. After calibration, upon visual inspection survival of patients in POCOP as a whole was more alike the NCR population data. Survival curves of patients in the calibrated POCOP and NCR conditioned on type of treatment are shown in Figures 4 and 5.

Table 2. Observed frequencies of patients' treatments in the NCR and POCOP

Treatment intent	Treatment	Tumor location	NCR frequency	POCOP frequency
Potentially curable	Definitive CRT	Esophagus	1,276 (11.56%)	171 (6.7%)
	Endoscopic resection	Esophagus	664 (6.02%)	26 (1.02%)
	nCRT + resection	Esophagus	3,018 (27.34%)	1,145 (44.85%)
	nCRT no resection	Esophagus	892 (8.08%)	313 (12.26%)
	nCT + resection (+- adjuvant CT)	Esophagus	314 (2.84%)	78 (3.06%)
	nCT + resection (+- adjuvant CT)	Stomach	1,250 (11.32%)	288 (11.28%)
	Other treatment	Esophagus	1,201 (10.88%)	35 (1.37%)
	Resection only	Stomach	687 (6.22%)	52 (2.04%)
Palliative	Chemoradiotherapy	Esophagus	318 (2.88%)	56 (2.19%)
	Chemotherapy or targeted therapy	Esophagus	2,037 (18.45%)	249 (9.75%)
	Chemotherapy or targeted therapy	Stomach	876 (7.94%)	70 (2.74%)
	Palliative resection + C(R)T or RT	Esophagus	97 (0.88%)	28 (1.1%)
	Radiotherapy only	Esophagus	962 (8.71%)	42 (1.65%)

Abbreviations: CRT, neoadjuvant chemoradiotherapy; nCT, neoadjuvant chemotherapy; CT, chemotherapy; RT, radiotherapy.

Median survival of the NCR, POCOP and calibrated POCOP was 19 [18, 20], 32 [31, 36], and 23 [20, 25] months, respectively. For potentially curable patients, median survival of the NCR, POCOP and calibrated POCOP was 32 [30, 33], 43 [40, 47], and 36 [32, 42] months, respectively. For palliative patients, median survival of the NCR, POCOP and calibrated POCOP was 9 [9, 9], 13 [11, 14], and 11 [10, 12] months. Median survival of clinically relevant subgroups stratified to treatment can be found in Table 3.

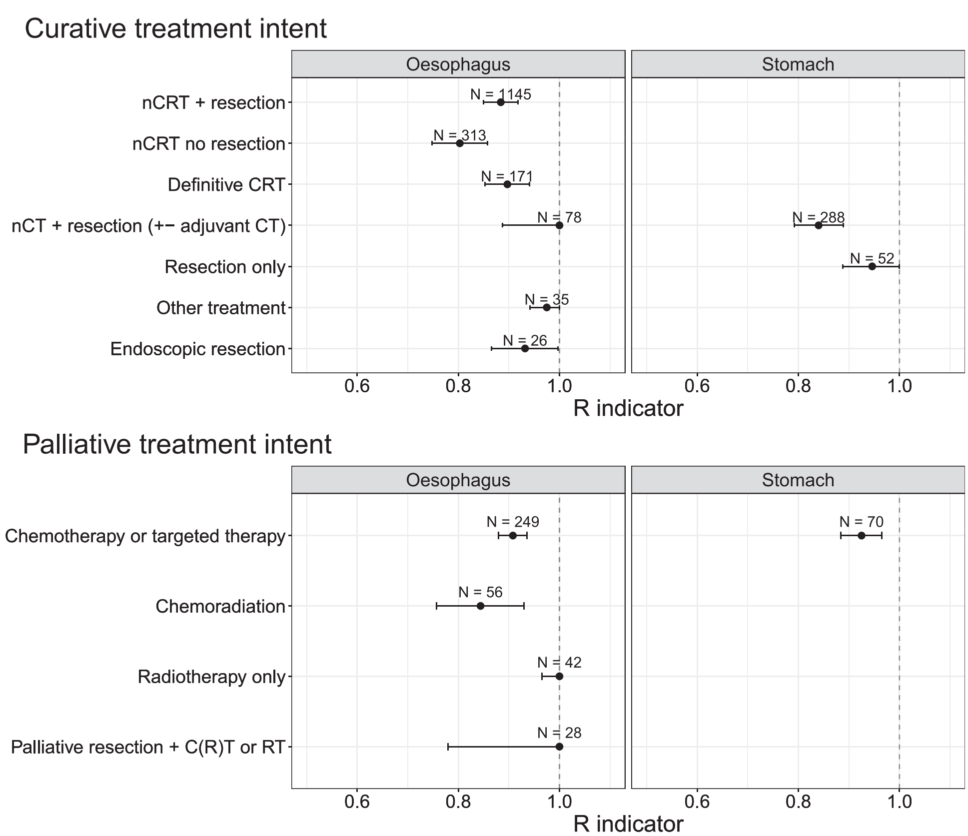

Fig. 1. R-indicators of subgroups with curative and palliative treatment intent. Sample sizes of the patients in POCOP in specific subgroups are reported. The dashed line indicates perfect representativeness. Horizontal bars represent the 95% confidence intervals of the R-indicators. nCRT = neoadjuvant chemoradiotherapy, nCT = neoadjuvant chemotherapy, CT = chemotherapy, RT = radiotherapy.

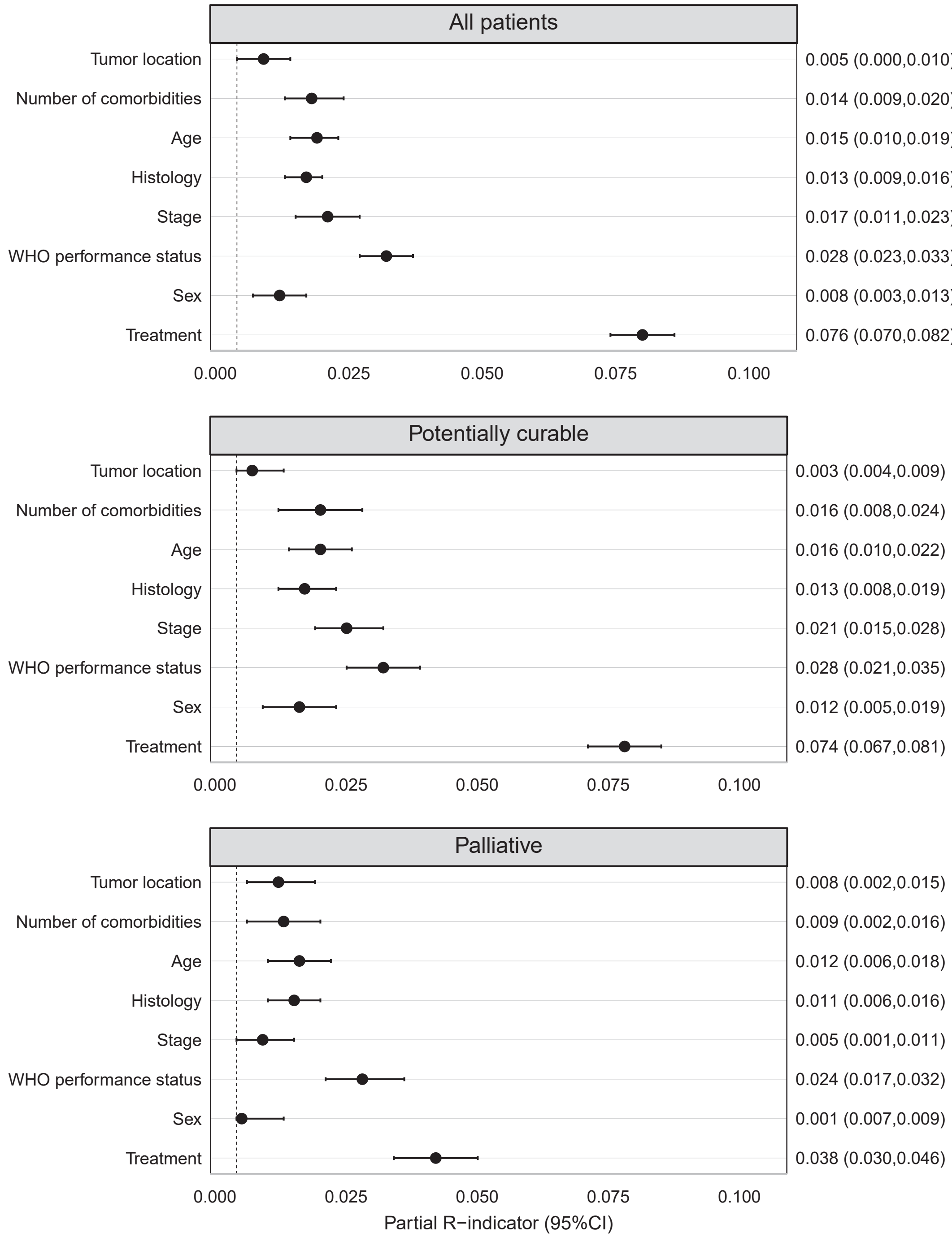

Fig. 2. Partial R-indicators for all patients and stratified to treatment intent. Values should be interpreted relative to each other. Higher values of the partial R-indicator correspond to a larger contribution to nonrepresentativeness due to that variable relative to the other variables.

The 5-year overall survival rates of patients in the NCR, POCOP and calibrated POCOP were 26%, 36%, and 27%, respectively. For potentially curable patients in the NCR, POCOP, and calibrated POCOP, the 5-year overall survival rates were 36%, 42% and 37%, respectively. For palliative treated patients in the NCR, POCOP, and calibrated POCOP, the 5-year overall survival rates were 4%, 8% and 6%, respectively. The 5-year overall survival of clinically relevant subgroups stratified to treatment can be found in Table 3.

4. Discussion

In this study we evaluated the representativeness of a prospective registry cohort for real-world data in oncology,

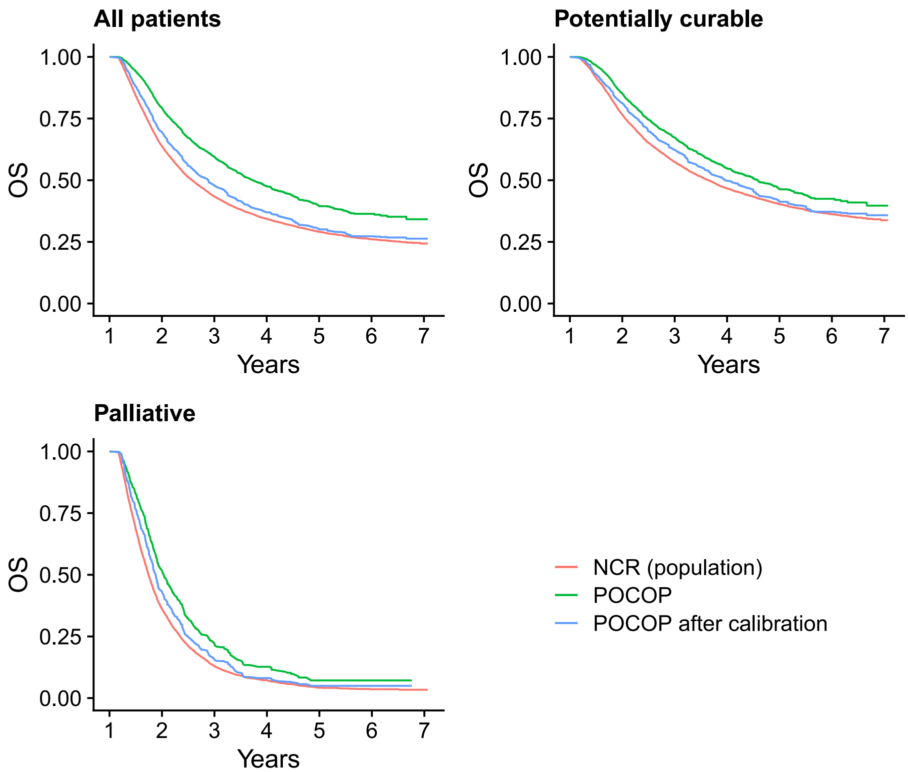

Fig. 3. Kaplan-Meier curves of the NCR (population), POCOP and calibrated POCOP for all patients. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

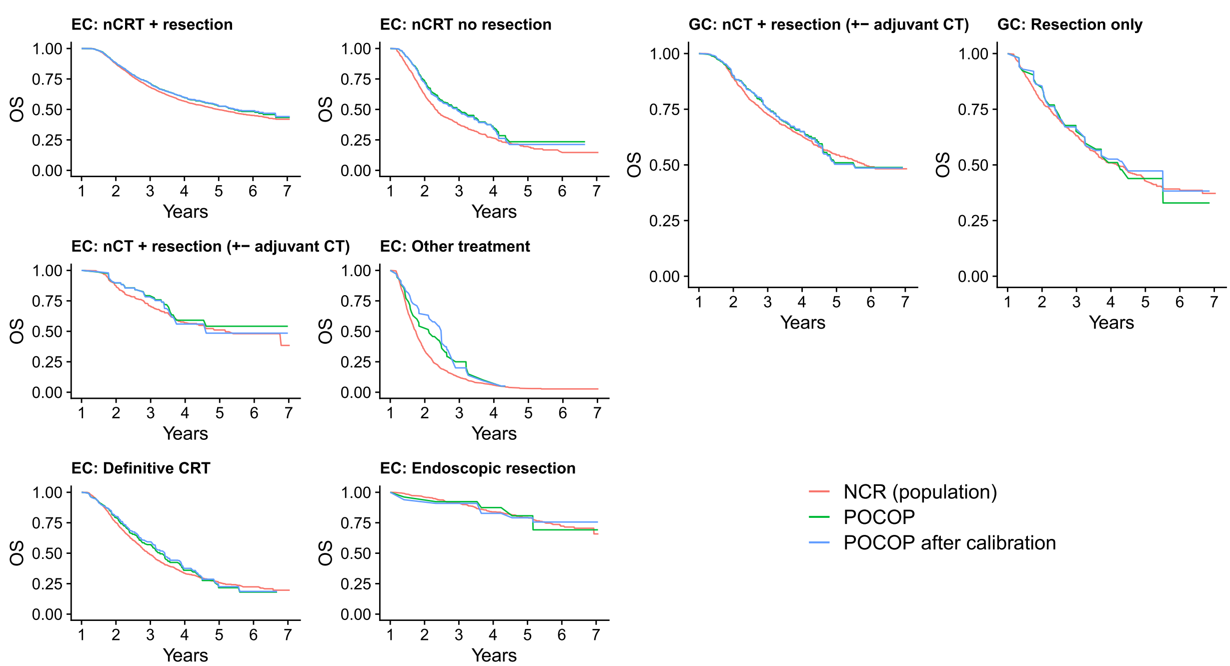

Fig. 4. Kaplan-Meier curves of the NCR (population), POCOP and calibrated POCOP, of potentially curable patients conditioned on treatment. EC = Esophageal cancer, GC: Gastric cancer, nCRT = Neoadjuvant chemoradiotherapy, CT = Chemotherapy, CRT = Chemoradiotherapy. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

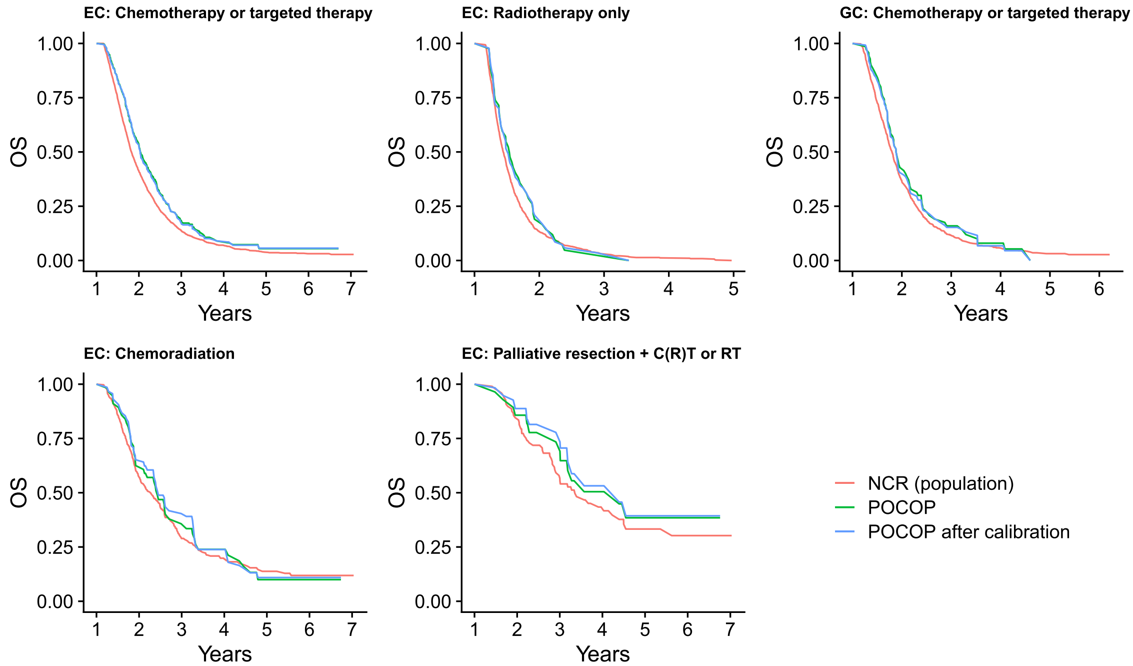

Fig. 5. Kaplan-Meier curves of the NCR (population), POCOP and calibrated POCOP, of palliative patients conditioned on treatment. EC = Esophageal cancer, GC = Gastric cancer. CRT = Chemoradiotherapy, RT = Radiotherapy. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

by applying a metric to quantify representativeness called R-indicators for the first time in medical oncology, and the extent to which differences between the patients included in the PROMs and population could be corrected to produce generalizable estimates.

We found that subgroups stratified to treatment generally had a higher real-world representativeness with respect to their respective target populations than the complete unstratified prospective registry cohort (POCOP). This implies that accounting for the variability of included treatments in the PROMs improved real-world representativeness of the prospective registry cohort. This pattern was also observed in survival analyses in which survival estimates from patients included in the prospective registry cohort were more similar to the target population in treatment-stratified samples. Therefore, although the variability of treatment may potentially introduce selection bias, the effects of selection bias can be mitigated by accounting for treatment using stratification. This is an important finding for other cancer types where treatment may also be a contributing factor in the selection mechanism that determines willingness to participate in PROMs studies.

Another important finding from stratification of the prospective registry cohort to clinically relevant groups was that lower inclusion rates of patients did not correspond with lower real-world representativeness. For example, the real-world representativeness of palliative patients was higher compared to potentially curable patients despite the inclusion rates being lower. This implied that, insofar included patients on average resemble patients from the real world, the absolute number of included patients does not matter. This confirms previous findings that lower inclusion rates do not necessarily cause biased samples and is contradictory to what is speculated in published studies using similar PROMs data [1,2,21,22].

In addition to the finding that stratification of the prospective registry cohort to clinically relevant subgroups improved representativeness, our study also showed that calibration weighting can be used as an alternative to stratification to obtain generalizable estimates from the PROMs with respect to the population to a large extend. Stratification in combination with calibration weighting only marginally improved the estimates. The advantage of using calibration weighting techniques is that it allows to control for multiple variables without the need to creating subgroups, in contrast to stratification where the number of strata increase exponentially as the number of variables to control for increase [23]. However, a disadvantage of weighting calibration is that is requires the population data in addition to the prospective cohort data to be able to perform the calibration of the PROMs data to the population, whereas stratification does not need these population data. Moreover, in some instances the calibrated survival curve did not perfectly overlap with the population curve, indicating that there still was some unobserved confounding for which we did not account. Known factors that may induce nonrepresentativeness in health related quality of life studies are physical condition and comorbidities; patients with better physical condition and fewer comorbidities are more likely to participate [7,24]. However, the WHO performance status and the number of comorbidities

Table 3. Median overall survival (months) of patients in the NCR, POCOP and calibrated POCOP

	NCR		POCOP		POCOP calibrated	
Treatment group	Median OS (months, 95%CI)	5-yr OS (%)	Median OS (months, 95%CI)	5-yr OS (%)	Median OS (months, 95%CI)	5-yr OS (%)
Potentially curable						
Esophagus						
nCRT + resection	48 [44, 53]	45	53 [48, NA]	48	54 [48, NA]	49
nCRT no resection	16 [15, 17]	15	24 [20, 29]	23	23 [18, 29]	22
Other treatment	9 [8, 9]	3	14 [8, 23]	-	18 [10, 23]	-
Definitive CRT	23 [22, 25]	22	27 [23, 34]	18	29 [25, 36]	19
Endoscopic resection	-	72	-	69	-	77
nCT + resection (+- adjuvant CT)	50 [40, NA]	48	-	54	-	51
Stomach						
nCT + resection (+- adjuvant CT)	58 [51, NA]	49	55 [43, NA]	49	55 [43, NA]	49
Resection only	38 [32, 45]	39	39 [27, NA]	33	42 [27, NA]	38
Palliative						
Esophagus						
Radiotherapy only	5 [5, 6]	-	7 [5, 10]	-	6 [5, 10]	-
Palliative resection + C(R]T or RT	28 [23, 42]	30	41 [24, NA]	38	41 [26, NA]	39
Chemoradiation	15 [12, 18]	12	17 [13, 25]	10	17 [14, 28]	12
Chemotherapy or targeted therapy	10 [9, 10]	3	12 [11, 14]	5	12 [11, 14]	5
Stomach						
Chemotherapy or targeted therapy	10 [9, 10]	3	11 [9, 14]	-	11 [9, 14]	-

- Not observed

Abbreviations: OS, overall survival; nC(R)T, neoadjuvant chemo(radio)therapy; nCT, neoadjuvant chemotherapy; CT, chemotherapy; RT, radiotherapy.

were included in the propensity model and are therefore unlikely to explain the remaining noncorrectable bias.

The R-indicator has shown potential to be able to estimate real-world representativeness of a prospective registry cohort. R-indicators can be intuitively interpreted as it expresses representativeness on the same scale (0−1) regardless of the type of variables and the set of data that is used. It provides a good alternative to more classic methods such as statistically testing observed frequencies of sample and population characteristics for significance, which is influenced by samples sizes and does not provide an overall summary statistic for the samples' representativeness [25].

4.1. Considerations

This study has a number of limitations. First, is the relatively small sample size of some of the clinically relevant subgroups that leads to large CIs of the R-indicator. Larger samples of treatment groups are needed to be able to prove a more precise estimate of the representativeness. A second limitation was that we could only estimate bias and corrected bias of survival analyses rather than health related quality of life or other patient reported outcomes. Health related quality of life measures were by definition only available for patients actively participating in the POCOP registry and not for patients in the NCR. However, built into the weighting calibration is that it increases covariate balance between the NCR and POCOP patients by creating a pseudo population from the original POCOP with properties that resemble the population from which it borrowed information. By analyzing this pseudo population (or calibrated sample) generalizability to the total population is increased [26,27]. Therefore, an outcome such as survival can still provide information on whether calibration can correct existing bias.

What is more, in the computation of the R-indicators there is an additional bias correction because the R-indicator is inherently biased to be smaller than one since the variance of the propensity scores is rarely zero. This bias is corrected through a built-in bias correction of the software and explains why some R-indicators are exactly one.

Major strength of this study was that used data from a large prospective cohort study and data from the reliable nationwide NCR which includes all diagnosed malignancies and is thus a very comprehensive population database, which enabled us to make comparisons between the prospective registry cohort and the Dutch population of patients with esophageal or gastric cancer.

4.2. Future perspectives

Given the advantages of R-indicators to express representativeness, future observational and clinical studies could be evaluated and managed more structurally and more uniformly. Additionally, real-world representativeness of PROMs registries could be improved by monitoring characteristics of included patients and adjusting inclusion strategies to reflect the total population. Adjusting inclusion strategies to target such specific patient groups could improve the real-world representativeness of PROMs registries. Finally, more research is needed in other health-care research settings to further investigate the suitability of prospective registry cohorts as real-world data.

5. Conclusion

This study demonstrated the assessment of real-world representativeness of patients who participated in a prospective registry cohort and showed that real-world representativeness improved when the variability in treatment was accounted for. Moreover, this study demonstrated the utility of representativeness indicators to explore real-world representativeness of prospective registry cohort as well as calibration techniques and stratification to correct for differences between the prospective registry cohort and the population.

Data availability

Data is available from the Netherlands Cancer Registry.

Declaration of competing interest

MS has served as a consultant for BMS and Lilly. NHM has served as a consultant for BMS, Merck, Lilly, Astra Zeneca and Servier. RV reports grants from BMS and has served as a consultant for Daiichi Sankyo. HvL has served as a consultant for BMS, Dragonfly, Lilly, Merck, Nordic Pharma and Servier and has received research funding and/or medication supply from Bayer, BMS, Celgene, Janssen, Incyte, Lilly, Merck, Nordic Pharma, Nordic, Philips, Roche and Servier. SCK, JB, TK, CJvdZ, EAK, LvB, and BRK have no disclosures.

Acknowledgments

The authors thank the registration team of the Netherlands Comprehensive Cancer Organisation (IKNL) for the collection of data for the Netherlands Cancer Registry. Furthermore, the authors thank all participating patients for filling out the questionnaires and all hospitals that included patients for the POCOP study: Admiraal de Ruyter Ziekenhuis, Amphia ziekenhuis, Amsterdam Universitair Medische Centra, Antonius Sneek ziekenhuis, Albert Schweitzer ziekenhuis, Antonie van Leeuwenhoek ziekenhuis, Bernhoven ziekenhuis, BovenIJ ziekenhuis, Bravis ziekenhuis, Catherina Ziekenhuis Eindhoven, Helmond Elkerliek ziekenhuis, Erasmus Medisch Centrum, Elisabeth-TweeSteden Ziekenhuis, Flevoziekenhuis, Gelre ziekenhuizen, HagaZiekenhuis, Ikazia ziekenhuis, Isala ziekenhuis, Jeroen Bosch ziekenhuis, Leiden University Medical Center, Maasstad ziekenhuis, Martini ziekenhuis, Medisch Centrum Leeuwarden, Meander Medisch Centrum, Máxima Medisch Centrum, Maastricht Universitair Medisch Centrum Plus, Nij Smellinghe ziekenhuis, Noordwest Ziekenhuisgroep, R. de Graaf ziekenhuis, Radboudumc, Rijnstate ziekenhuis, Rode Kruis Ziekenhuis, St Jans Gasthuis Weert, Slingeland ziekenhuis, Spaarne ziekenhuis, St. Anna ziekenhuis, St. Antonius ziekenhuis, St. Jansdal ziekenhuis, Tjongerschans ziekenhuis, Universitair Medisch Centrum Utrecht, VieCuri Medisch Centrum, Weel-Bethesda ziekenhuis, Ziekenhuisgroep Twente, Zuyderland Medisch Centrum, and ZorgSaam Zorggroep Zeeuws-Vlaanderen.

References

[1] Dijksterhuis WPM, Latenstein AEJ, Van Kleef JJ, Verhoeven RHA, De Vries JHM, Slingerland M, et al. Cachexia and dietetic interventions in patients with esophagogastric cancer: a multicenter cohort study. J Natl Compr Canc Netw 2021;19(2):144–52.

[2] van Kleef JJ, Dijksterhuis WPM, van den Boorn HG, Prins M, Verhoeven RHA, Gisbertz SS, et al. Prognostic value of patient-reported quality of life for survival in oesophagogastric cancer: analysis from the population-based POCOP study. Gastric Cancer 2021;24:1203–12.

[3] Davie A, Carter GC, Rider A, Pike J, Lewis K, Bailey A, et al. Real-world patient-reported outcomes of women receiving initial endocrine-based therapy for HR+/HER2- advanced breast cancer in five European countries. BMC Cancer 2020;20(1):1–15.

[4] Jakob A, Zahn MO, Nusch A, Werner T, Schnell R, Frank M, et al. Real-world patient-reported outcomes of breast cancer or prostate cancer patients receiving antiresorptive therapy for bone metastases: final results of the PROBone registry study. J Bone Oncol 2022;33:100420.

[5] Gotto G, Drachenberg DE, Chin J, Casey R, Fradet V, Sabbagh R, et al. Real-world evidence in patient-reported outcomes (PROs) of metastatic castrate-resistant prostate cancer (mCRPC) patients treated with abiraterone acetate + prednisone (AA+P) across Canada: final results of COSMiC. Can Urol Assoc J 2020;14(12):10–4.

[6] Hernán MA, Hernández-Díaz S, Robins JM. A structural approach to selection bias. Epidemiology 2004;15:615–25.

[7] Unruh M, Yan G, Radeva M, Hays RD, Benz R, Athienites NV, et al. Bias in assessment of health-related quality of life in a hemodialysis population: a comparison of self-administered and interviewer-administered surveys in the HEMO Study. J Am Soc Nephrol 2003;14(8):2132–41.

[8] Derksen JWG, Vink GR, Elferink MAG, Roodhart JML, Verkooijen HM, van Grevenstein WMU, et al. The Prospective Dutch Colorectal Cancer (PLCRC) cohort: real-world data facilitating research and clinical care. Sci Rep 2021;11(1):1–12.

[9] Schouten B, Bethlehem J, Beullens K, Kleven Ø, Loosveldt G, Luiten A, et al. Evaluating, comparing, monitoring, and improving representativeness of survey response through R-indicators and partial R-indicators. Int Stat Rev 2012;80(3):382–99.

[10] Schouten B, Cobben F, Bethlehem J. Indicators for the representativeness of survey response. Surv Methodol 2009;35(1):101–13.

[11] Bethlehem J, Cobben F, Schouten B. Handbook of nonresponse in household surveys. Hoboken, NJ: John Wiley and Sons Inc.; 2011.
[12] Coebergh van den Braak RRJ, van Rijssen LB, van Kleef JJ, Vink GR, Berbee M, van Berge Henegouwen MI, et al. Nationwide comprehensive gastro-intestinal cancer cohorts: the 3P initiative. Acta Oncol 2018;57(2):195−202.
[13] Casparie M, Tiebosch ATMG, Burger G, Blauwgeers H, Van De Pol A, Van Krieken JHJM, et al. Pathology databanking and biobanking in The Netherlands, a central role for PALGA, the nationwide histopathology and cytopathology data network and archive. Cell Oncol 2007;29(1):19−24.
[14] Brierley JD, Gospodarowicz MK, Wittekind C. TNM classification of malignant tumours. 8th edition. International Union Against Cancer (UICC); 2014.
[15] Hulshof MCCM, van Laarhoven HWM. Chemoradiotherapy in tumours of the oesophagus and gastro-oesophageal junction. Best Pract Res Clin Gastroenterol 2016;30(4):551−63.
[16] Luiten A, Schouten B. Tailored fieldwork design to increase representative household survey response: an experiment in the survey of consumer satisfaction. J R Stat Soc Ser A Stat Soc 2013;176(1):169−89.
[17] Moore JC, Durrant GB, Smith PWF. Data set representativeness during data collection in three UK social surveys: generalizability and the effects of auxiliary covariate choice. J R Stat Soc Ser A Stat Soc 2018;181(1):229−48.
[18] De Heij V, Schouten B, Sholomo N. RISQ manual 2.1 tools in SAS and R for the computation of R-indicators, partial R-indicators and partial coefficients of variation 2015.
[19] Cole SR, Hernán MA. Adjusted survival curves with inverse probability weights. Comput Methods Programs Biomed 2004;75(1):45−9.
[20] Stekhoven DJ, Bühlmann P. Missforest-Non-parametric missing value imputation for mixed-type data. Bioinformatics 2012;28(1):112−8.
[21] Burbach JPM, Kurk SA, Coebergh van den Braak RRJ, Dik VK, May AM, Meijer GA, et al. Prospective Dutch colorectal cancer cohort: an infrastructure for long-term observational, prognostic, predictive and (randomized) intervention research. Acta Oncol 2016;55(11):1273−80.
[22] Groves RM, Peytcheva E. The impact of nonresponse rates on nonresponse bias: a meta-analysis. Public Opin Q 2008;72(2):167−89.
[23] Deb S, Austin PC, Tu JV, Ko DT, Mazer CD, Kiss A, et al. A review of propensity-score methods and their use in cardiovascular research. Can J Cardiol 2016;32(2):259−65.
[24] Addington-Hall J, Kalra L. Measuring quality of life: who should measure quality of life? BMJ 2001;322(7299):1417−20.
[25] Westgeest HM, Uyl-de Groot CA, van Moorselaar RJA, de Wit R, van den Bergh ACM, Coenen JLLM, et al. Differences in trial and real-world populations in the Dutch castration-resistant prostate cancer registry. Eur Urol Focus 2018;4(5):694−701.
[26] Stuart EA, Cole SR, Bradshaw CP, Leaf PJ. The use of propensity scores to assess the generalizability of results from randomized trials. J R Stat Soc Ser A Stat Soc 2011;174(2):369−86.
[27] Nohr EA, Liew Z. How to investigate and adjust for selection bias in cohort studies. Acta Obstet Gynecol Scand 2018;97(4):407−16.