


Development of a normal tissue complication probability model for late unfavourable aesthetic outcome after breast-conserving therapy

Isabelle Kindts^{a,b} , Gilles Defraene^a, Annouschka Laenen^c, Saskia Petillion^b, Erik Van Limbergen^{a,b}, Tom Depuydt^{a,b} and Caroline Weltens^{a,b}

^aDepartment of Oncology, Experimental Radiation Oncology, KU Leuven – University of Leuven, Leuven, Belgium; ^bDepartment of Radiation Oncology, University Hospitals Leuven, Leuven, Belgium; ^cLeuven Biostatistics and Statistical Bioinformatics Centre (L-Biostat), KU Leuven University, Leuven, Belgium

ABSTRACT

Purpose/Objectives: To develop a normal tissue complication probability (NTCP) model for late unfavourable aesthetic outcome (AO) after breast-conserving therapy.

Material and Methods: The BCCT.core software evaluated the AO using standardized photographs of patients treated between 2009 and 2014. Dose maps in 2 Gy equivalents were calculated assuming $\alpha/\beta = 3.6$ Gy. Uni- and multivariable logistic regression analysis was performed to study the predictive value of clinicopathological and dosimetric variables for unfavourable AO. The Lyman Kutcher Burman (LKB) model was fit to the data with dose modifying factors (dmf). Model performance was assessed with the area under the curve (AUC) of the receiver operating characteristic curve and bootstrap sampling.

Results: Forty-four of the 121 analysed patients (36%) developed unfavourable AO. In the optimal multivariable logistic regression model, a larger breast volume receiving ≥ 55 Gy (V55), a seroma and an axillary lymph node dissection (ALND) were independently associated with an unfavourable AO, AUC = 0.75 (95%CI 0.64;0.85). Beta-estimates were -2.68 for β_0 , 0.057 for V55, 1.55 for seroma and 1.20 for ALND. The optimal LKB model parameters were $EUD_{3.6}(50) = 63.3$ Gy, $n = 1.00$, $m = 0.23$, $dmf(\text{seroma}) = 0.83$ and $dmf(\text{ALND}) = 0.84$, AUC = 0.74 (95%CI 0.61;0.83).

Conclusions: An NTCP model for late unfavourable AO after breast-conserving therapy was developed including seroma, axillary lymphadenectomy and V55.

ARTICLE HISTORY

Received 8 January 2018
Accepted 30 March 2018

Introduction

Breast cancer is the most common cancer in women and breast-conserving therapy (BCT), i.e., breast-conserving surgery followed by radiotherapy (RT), is the standard therapeutic option [1–3]. By adding RT more than 75 percent of the early stage breast cancer patients survive long term, making treatment-induced side-effects important outcome measures [4]. Over the last two decades, reducing incidental irradiation on heart and lung tissue was the main research focus since a significant excess of non-breast-cancer mortality has been shown (rate ratio 1.30 for cardiac death and 2.10 for secondary lung cancer after more than ten years) [5]. Currently, aesthetic outcome (AO) is becoming increasingly important because it influences the quality of life and psychosocial recovery after breast cancer treatment [6].

Whilst mortality rates can be measured objectively, several subjective and objective measurements were utilized for AO, often examining different sub-outcomes such as fibrosis, teleangiectasia or oedema. In 2012, recommendations for the evaluation of AO were established, and the BCCT.core software was suggested as standard to objectively measure AO [7]. The software scores AO using standardized photographs

resulting in an ‘excellent’, ‘good’, ‘fair’ or ‘poor’ outcome as introduced by Harris et al. [8,9]. It was shown that there was no measurable change in late AO measured by BCCT.core between two and six years after surgery [10]. Furthermore, Hennigs et al. have defined clinicopathological predictors for an unfavourable (‘fair’ or ‘poor’) AO: axillary lymphadenectomy, a tumour in the 12 o’clock position, a tumour with minimal (y)pT2 stage and a surgical specimen weight more than 75 grams [10]. To the best of our knowledge, however, the dose-volume effect of breast RT on AO measured by BCCT.core has not yet been studied. Ultimately, by taking clinicopathological and dose-volume metrics into account in a prediction model, the dose distribution of individual treatment plans could be optimized and patients could be informed on the risk of late unfavourable AO.

The purpose of the present study was to develop a prediction model for late unfavourable AO after BCT. We have fitted two normal tissue complication probability (NTCP) models. Firstly, we aimed to investigate which clinicopathological factors and dose-volume metrics were associated with an unfavourable AO and to develop a multivariable logistic regression model. Secondly, we wanted to estimate the

parameters for a Lyman Kutcher Burman (LKB) model in which the dose-volume histogram (DVH) is reduced to a single parameter: the equivalent uniform dose (EUD).

Material and methods

The TRIPOD (Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis) recommendations for a type 1B study were followed [11]. The study was approved by the Clinical Trial Centre and the Ethical Committee of our institution.

Patient cohort

The population of this study consisted of patients treated with BCT between December 2009 and July 2014, who consulted the multidisciplinary breast cancer group at the University Hospitals Leuven, Belgium between 1 December 2015 and 31 October 2016. Patients were invited to participate during their prescheduled follow-up appointment; which is, according to the recurrence risk and time since diagnosis four- or six-monthly to yearly. Informed consent was obtained. All RT treatments had to be conducted at the University Hospitals Leuven between two and six years ago.

BCT included whole-breast irradiation followed by a boost to the tumour bed. Prescription doses to the whole breast varied between 42.56 Gy in 16 fractions and 50 Gy in 25 fractions, and calculation models used were pencil beam and AAA from Eclipse (Varian Medical Systems, Palo Alto, CA). For the selection of the boost technique, an in-house developed flowchart based on the depth of the tumour bed was used. For a clinical target volume lying more than 28 mm beneath the epidermis, an interstitial or photon boost was chosen over an electron boost because of skin doses [12]. Boost doses to the tumour bed varied between 12.5 Gy in 5 fractions and 20 Gy in 10 fractions with external beam radiation therapy and between 8.5 Gy and 10 Gy with high dose rate brachytherapy. Treatment regimens were discussed at the multidisciplinary tumour board. Hormonal therapy included tamoxifen and aromatase-inhibitors and chemotherapy involved 5 fluorouracil, epirubicin, cyclophosphamide, methotrexate, and taxanes.

Toxicity analysis

An anterior view picture of the patient was taken with the patient's hands on the hips. The BCCT.core software was used to evaluate the AO. Based on the semi-automatic localization of fiducial points, the software firstly measures asymmetry, skin colour changes and surgical scar appearance. Secondly, the set of measures is automatically converted into an overall objective AO: excellent, good, fair, or poor. AO was dichotomized in excellent/good ('favourable') versus fair/poor ('unfavourable').

Data management

Clinicopathological variables known to be associated with AO were collected [10,13–16]. Dose-volume metrics for the breast were assessed from the individual patient RT plans. The breast volume was delineated according to the guidelines from the EORTC, but included the 5 mm of breast skin [17]. Summed dose maps in 2 Gy equivalents (EQD2) were calculated assuming $\alpha/\beta = 3.6$ Gy [18]. In case of a brachytherapy boost, a Deformable Image Registration mapped the brachytherapy CT to the external beam planning CT in MIM 6.1.7 (MIM software, Cleveland, OH) using a free-form intensity-based registration algorithm. After a warping of the brachytherapy dose map, a dose summation was performed. For the logistic regression analysis, the considered dose-volume metrics included the relative volumes of the breast receiving 5 Gy (V5) to 75 Gy (V75) in steps of 5 Gy and the mean breast dose. For the LKB model, the EUD was calculated from the corresponding DVH [19]. In the LKB model, 3 parameters need to be estimated to fit the sigmoidal curve of normal tissue toxicity: EUD50, m and n . EUD50 is the homogeneous dose at which 50% of the patients experience an unfavourable AO and determines the position of the curve along the dose-axis. m determines the steepness of the dose response curve. The third parameter, n , describes the magnitude of the volume-effect which is the modification of normal tissue tolerance with a change in irradiated volume [20]. Dose modifying factors (dmf) can be added to take other parameters into account. Multiplying the baseline D50 with the concrete dmf then results in the D50 for that specific parameter [21].

Statistical analysis

Summary statistics were presented as means and range/medians and interquartile range for continuous variables and as frequencies and percentages for categorical variables. Firstly, univariable logistic regression analysis was performed to study the predictive value of clinicopathological factors and dose-volume metrics for an unfavourable AO. Results were reported as odds ratios (OR) with 95% confidence intervals (CI). All patient data were used for the clinicopathological factors. Patients in whom no dosimetrical data were available were excluded for the dose-volume metric analysis. A multivariable forward stepwise model was built with covariate inclusion at p value $< .05$ and removal at p value $\geq .05$. All tests were two-sided, a 5% significance level was assumed for all tests. In individual cases, the probability of unfavourable AO can be calculated using the following formula: $NTCP = 1/(1 + e^{-t})$, where $t = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \dots$. Secondly, we estimated the parameters of the LKB model [22,23]. The maximum likelihood procedure was used to fit the LKB model to the data, and the likelihood ratio test to identify significant differences between nested models. The discriminating ability of the model was described by the area under the curve (AUC) of the receiver operating characteristic curve [24]. The reported models were optimized on the complete dataset. A 95% CI of AUC was calculated for the models using 10,000 bootstrap replications. Calibration was evaluated

using a Hosmer–Lemeshow test for the multivariable logistic regression model, and a comparison with the LKB model was done with the loglikelihood [25].

For uni- and multivariable logistic regression, Statistica version 13 (Dell Inc, Tulsa, OK, USA) and SAS software (version 9.4 of the SAS System for Windows) were used. NTCP model fits were performed in Matlab (The Mathworks Inc., Natick, MA, USA).

Results

One-hundred twenty-one patients were included in the study. No patient refused inclusion. Median follow-up was 3.42 years. Forty-four patients (36%) developed unfavourable AO. Patient, tumour and treatment characteristics are available in Table 1. The median age was 60 years old and 29 patients (24%) underwent an ALND. A seroma was present at the simulation CT in 27 patients (22%). In total, 62 patients received a brachytherapeutic boost and 59 received an external beam RT boost. Primary breast-conserving surgery was performed at the University Hospitals Leuven in all but one patient. Surgery was conducted by five surgeons from the same multidisciplinary breast cancer team for 120 patients; three surgeons operated 111 patients (92%).

Treatment plans and dosimetric data were available for 99 of the 121 patients, 22 brachytherapy plans were missing due to software problems. Calculation models used were pencil beam in 5 patients and AAA in 94 patients. 67 patients out of 99 had favourable AO and 32 unfavourable AO. Wedges were used in 38 three-dimensional (3D) conformal treatment plans and field-in-field RT in 61 3D conformal treatment plans.

Table 2 shows the results of the univariable analyses for the clinicopathological variables (tested on 121 patients) and dose-volume metrics (tested on 99 patients). Univariable clinicopathological and dosimetric factors significantly associated with an unfavourable AO were the administration of chemotherapy (OR 2.64, p .02), ALND versus a sentinel lymph node procedure (OR 3.17, p .01), a seroma present at simulation CT (OR 2.80, p .02), a higher body mass index (BMI) (OR 1.09, p .04), surgical specimen weight (OR 1.02, p .01), mean dose to the breast (OR 1.18, p .03) and most relative volumes of breast tissue from V30 to V65. V55 was the most significant dose-volume metric (OR 1.05, p .01).

The univariable significant clinicopathological variables and the most significant dose-volume metrics were included in the multivariable model building: administration of chemotherapy, ALND, seroma, BMI, specimen weight and V55. Improvements in likelihood were achieved by adding seroma (-62.3 to -58.7 , p =.008), ALND (-58.7 to -55.6 , p =.01) and V55 (-55.6 to -51.8 , p =.006). In the optimal multivariable logistic regression model, a higher V55 (x_1), a seroma (x_2) and an ALND (x_3) were significantly associated with an unfavourable AO. Beta-estimates were -2.68 for β_0 (95% CI -3.89 ; -1.49), 0.057 for β_1 (95% CI 0.01 ; 0.10), 1.55 for β_2 (95% CI 0.45 ; 2.66) and 1.20 for β_3 (95% CI 0.19 ; 2.22). Figure 1 shows a graphical representation of the model. A patient treated with ALND and V55 to the breast of 20% for example

has a 42% chance of late unfavourable AO. If V55 is reduced to 10%, the resulting NTCP is 29%. In case of no seroma or ALND, implementing a V55 constraint of 8% for breast tissue in the treatment planning process would result in an NTCP of 10%. Concerning model performance, AUC was 0.76. Bootstrap sampling resulted in AUC 0.75 (95% CI 0.64; 0.85). The loglikelihood was -51.8 . The Hosmer–Lemeshow test was not significant (chi square 2.87, p .94) indicating good agreement between expected and observed unfavourable AO rates.

The best fit NTCP parameters were $EUD_{3,6}(50) = 55.80$ Gy, $n = 1.00$ and $m = 0.18$ for the LKB model. Adding seroma or ALND or both to the LKB model significantly increased the likelihood (-59.76 to -57.20 , -57.19 and -54.33 , respectively, p =.02 in all three cases). The optimal LKB model thus included two dmfs. Taking into account seroma and ALND, the LKB model parameters were $EUD_{3,6}(50) = 63.3$ Gy (95% CI 59.2; 68.1), $n = 1.00$ (95% CI 0.62; 1), $m = 0.23$ (95% CI 0.16; 0.37), dmf (seroma) = 0.83 (95% CI 0.73; 0.95) and dmf (ALND) = 0.84 (95% CI 0.75; 0.94). Supplementary Appendix 1 shows a graphical representation of the model. To assess discrimination, an AUC of 0.74 was obtained. Bootstrap sampling resulted in AUC 0.74 (95% CI 0.61; 0.83). The loglikelihood was -54.3 , this is a lower value than in the multivariable logistic regression model and indicates a better fit of the multivariable logistic regression model.

Discussion

The purpose of the current study was to develop an NTCP model for late unfavourable AO after BCT taking into account clinicopathological and dosimetric variables. We aimed to develop a clinically useful NTCP model that can be applied to inform patients and to optimize dose distributions of individual patients.

Hennigs et al. have shown there was no measurable change in AO measured by BCCT.core between two and six years after surgery [10]. One-hundred twenty-one patients treated with BCT between two and six years ago were evaluated for the present study and forty-four patients (36%) developed unfavourable AO. Univariable clinicopathological predictors for an unfavourable AO were the administration of chemotherapy, ALND, seroma and a higher BMI or surgical specimen weight. An increasing mean dose to the breast and most relative volumes of breast tissue from V30 to V65 were dose-volume metrics associated with unfavourable AO. In multivariable logistic regression analysis, ALND, seroma and the relative volume of breast tissue receiving 55 Gy (V55) were withheld.

The different instruments measuring AO or sub-outcomes such as fibrosis, teleangiectasia or oedema make it difficult to compare studies. The unfavourable AO rate of 36% is in line with limited reports available for overall late AO measured by BCCT.core in literature (31.8% toxicity rate) [10]. Late AO measured with BCCT.core has been associated with ALND, a tumour in the 12 o'clock position, a tumour with minimal (y)pT2 stage and a surgical specimen weight more than 75 grams [10]. In the present study, ALND and an

Table 1. Patient, tumour and treatment characteristics.

Characteristic	Study population (n = 121)		Excellent/Good (n = 77)		Fair/Poor (n = 44)	
	Median		IQR ^a		Median	IQR ^a
Age at diagnosis (years)	60.0		52.8; 65.4		60.8	52.9; 66.4
Characteristic	n		%		n	%
Breast cup						
A	4		3.31		3	3.90
B	27		22.31		21	27.27
C	40		33.06		23	29.87
D	31		26.62		20	25.97
E	15		12.40		7	9.09
F	2		1.65		1	1.30
G	2		1.65		2	2.60
Breast band size (EU)						
70	15		12.40		10	12.99
75	26		21.49		19	24.68
80	27		22.31		18	23.38
85	20		16.53		11	14.29
90	14		11.57		9	11.69
95	10		8.26		5	6.49
100	5		4.13		2	2.60
≥105	4		3.31		3	3.90
Chemotherapy						
No	87		71.90		61	79.22
Yes	34		28.10		16	20.78
Hormonal therapy						
No	11		9.09		8	10.39
Yes	110		90.91		69	89.61
Trastuzumab						
No	110		90.91		73	94.81
Yes	11		9.09		4	5.19
Boost modality						
Brachytherapy	62		51.24		39	50.65
EBRT ^a	59		48.76		38	49.35
Lymphadenectomy						
None	11		9.09		9	11.69
Sentinel	81		66.94		56	72.73
ALND ^a	29		23.97		12	15.58
Position of tumor						
6 h	9		7.44		5	6.49
12 h	13		10.74		6	7.79
retro-areolar	2		1.65		2	2.60
upper inner	19		15.70		14	18.18
upper outer	46		38.02		32	41.56
lower inner	5		4.13		3	3.90
lower outer	8		6.61		3	3.90
3/9 h inner	5		4.13		3	3.90
3/9 h outer	14		11.57		9	11.69
Pathologic T-stage						
is	12		9.92		10	12.99
0	3		2.48		0	0.00
1	70		57.85		49	63.64
2	36		29.75		18	23.38
Re-resection						
No	114		94.21		73	94.81
Yes	7		5.79		4	5.19
Seroma						
No	94		77.69		65	84.42
Yes	27		22.31		12	15.58
Characteristic	Mean		Range		Mean	Range
BMI (kg/m ²)	26.8		17.7; 41.5		26.2	19.5; 41.5
Surgical specimen weight (g)	46.7		5.0; 248.0		39	5.0; 138.0
Volume breast (cm ³)	902.8		210.9; 2125.4		854.3	244.8; 2125.4
Dmean breast (Gy EQD2)	51.1		46.1; 64.1		50.6	46.1; 57.5
V30 EQD2 (%)	96.4		93.4; 99.7		96.1	93.4; 99.7
V35 EQD2 (%)	93.2		88.9; 98.5		92.9	88.9; 98.0
V40 EQD2 (%)	89		82.4; 96.0		88.6	82.4; 94.6
V45 EQD2 (%)	82.3		69.6; 91.8		81.6	69.6; 90.0
V50 EQD2 (%)	52.1		10.0; 84.9		49.9	10.0; 76.8
V55 EQD2 (%)	20.4		4.8; 65.4		18	4.8; 50.5
V60 EQD2 (%)	12.8		2.0; 52.0		10.9	2.0; 37.3
V65 EQD2 (%)	6.8		0.0; 40.6		5.3	0.0; 19.4
V70 EQD2 (%)	2.9		0.0; 25.4		2.6	0.0; 15.6
V75 EQD2 (%)	2		0.0; 19.2		1.9	0.0; 12.7

^aALND: axillary lymph node dissection; EBRT: external beam radiotherapy; EQD2: equivalent dose in 2 Gy fractions using an alpha/beta of 3.6 Gy [18]; IQR: interquartile range.

Table 2. Univariable analysis for unfavourable AO.

Variable		Odds ratio	95% CI ^a	p value
Age at diagnosis (years)	+1 year	0.99	0.96; 1.03	.64
Breast cup	+1 unit	1.21	0.88; 1.65	.24
Breast band size (EU)	+1 unit	1.13	0.92; 1.39	.24
Chemotherapy	Yes vs No	2.64	1.17; 5.96	.02
Hormonal therapy	Yes vs No	1.58	0.40; 6.31	.51
Trastuzumab	Yes vs No	3.45	0.95; 12.54	.06
Boost modality	Brachy vs EBRT	1.07	0.51; 2.24	.86
Lymphadenectomy Ref. Sentinel	None	0.50	0.10; 2.47	.39
	Axillary lymphadenectomy	3.17	1.32; 7.62	.01
Position of tumor Ref. Outer quadrants ^b	Inner quadrants	0.83	0.33; 2.09	.69
	12 o'clock	2.14	0.65; 7.09	.21
Pathologic T-stage Ref. [1] ^c	is	0.47	0.09; 2.32	.35
	2	2.33	1.02; 5.35	.05
Re-resection	Yes vs No	1.34	0.28; 6.26	.71
Seroma	Yes vs No	2.80	1.17; 6.73	.02
BMI (kg/m ²)	+1 unit	1.09	1.00; 1.18	.04
Surgical specimen weight (g)	+1 unit	1.02	1.00; 1.03	.01
Surgical specimen weight >75 g	Yes vs No	2.54	0.92; 7.01	.07
Volume breast (cm ³)	+100 units	1.09	0.99; 1.19	.08
Dmean breast (EQD2 Gy)	+1 Gy	1.18	1.02; 1.37	.03
V30 EQD2 (%)	+1%	1.42	1.04; 1.94	.03
V35 EQD2 (%)	+1%	1.28	1.03; 1.60	.02
V40 EQD2 (%)	+1%	1.20	1.02; 1.42	.03
V45 EQD2 (%)	+1%	1.12	1.01; 1.25	.03
V50 EQD2 (%)	+1%	1.02	1.00; 1.05	.10
V55 EQD2 (%)	+1%	1.05	1.01; 1.10	.008
V60 EQD2 (%)	+1%	1.07	1.01; 1.12	.012
V65 EQD2 (%)	+1%	1.09	1.02; 1.16	.02
V70 EQD2 (%)	+1%	1.05	0.96; 1.15	.32
V75 EQD2 (%)	+1%	1.01	0.88; 1.16	.86

^a95% CI: 95% confidence interval.

^bNipple-areolar complex: only 2 patients; none with unfavourable AO.

^c0: only 3 patients; all with unfavourable AO; EQD2: equivalent dose in 2 Gy fractions using an alpha/beta of 3.6 Gy [18]; significant parameters are highlighted in bold.

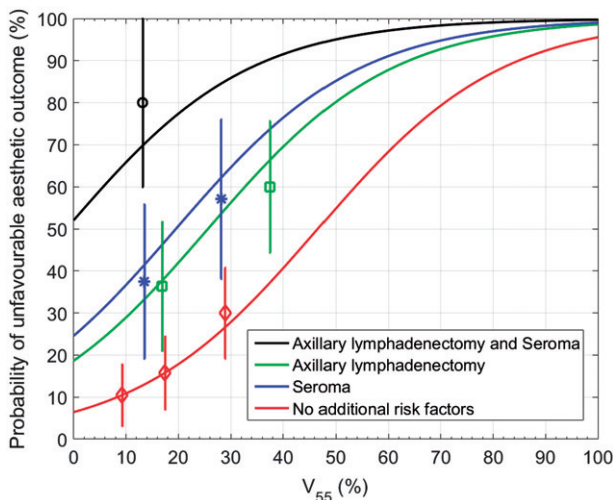


Figure 1. Graphical representation of the multivariable logistic regression model.

increasing surgical specimen weight were confirmed as significant variables. Factors associated with worse AO, regardless of the specific (sub)outcome and of the instrument applied, can be divided into patient, tumour and treatment-related factors. Age, BMI and breast size are reported as patient-related factors. The position of the tumour and the tumour stage are described as tumour-related factors. Concerning treatment-related factors: systemic therapy, oncoplastic surgery, surgical specimen weight, lymphadenectomy, re-excision, postoperative seroma or oedema, a sequential or

simultaneous boost, administration of the boost dose, the boost modality, dose inhomogeneity, (electron) energy and total RT dose are described [10,13–16,26]. As reported by Hammer, most of the prognostic clinicopathological variables highly correlate with the DVH parameters derived from the dose distribution: for example, administration of boost dose, boost modality, cup size and also BMI [15]. In our analysis, Spearman correlation tests showed no correlation between V55 and surgical specimen weight and a possible correlation between V55 and BMI with $r = -0.17$ and p value .10.

Despite the large and rather inhomogeneous amount of data available on AO after BCT, limited research has been conducted so far on NTCP models predicting the incidence of unfavourable AO in individual patients, taking into account dosimetric data. We have conducted a forward stepwise logistical regression including univariable significant clinicopathological variables and obtained a model with seroma and ALND. Adding the dose-volume metric V55 to the model significantly increased the likelihood ratio. Based on the present study, we propose the following formula: $NTCP = 1/(1 + e^{-t})$, where $t = -2.68 + (0.057 \times V55) + 1.55$ (if seroma) $+ 1.20$ (if ALND). The likelihood ratio test showed that besides V55, also a second parameter would improve the likelihood ($p = .02$). Adding a third variable would be borderline acceptable with $p = .057$. Estimating three NTCP parameters (EUD, n and m) in the LKB model therefore does not imply a high risk for overfitting. The optimal LKB model also included seroma and ALND as dmf. The LKB model parameters were $EUD_{3,6}(50) = 63.3$ Gy (95% CI 59.2;68.1),

$n = 1.00$ (95% CI 0.62;1), $m = 0.23$ (95% CI 0.16;0.37), dmf (seroma) = 0.83 (95% CI 0.73;0.95) and dmf (ALND) = 0.84 (95% CI 0.75;0.94). Our data thus suggest that a seroma, an axillary lymphadenectomy and the V55 are the most important factors and that the effect of the maximum dose is small ($n = 1.00$, this is a large n -value and corresponds to a large volume-effect [20]). As can be seen in Figure 1 and as reflected by the magnitude of the odds ratios, the presence of a seroma or an ALND causes a large increase in the probability of unfavourable AO and rises in V55 result in a smaller increase. ALND and seroma are thus the most important factors in the model and adding V55 to the model has a fine-tuning effect.

Classical prescription doses to the clinical target volume (CTV) of the breast are 50 Gy EQD2 and to the CTV boost 66 Gy EQD2. Important to remark is that all patients in our cohort have received a boost dose, as is reflected in the range of V55 that does not reach 0%. The present model could thus show a suboptimal performance in patients not undergoing a boost. However, a boost is advised in all young patients according to current guidelines and we can presume AO is mainly important for those younger patients [27]. V55 thereby is a clinically relevant parameter and strategies to lower the value have already been reported in literature [14,15]. Van der Laan et al. have shown that by using the simultaneous integrated boost (SIB) technique, the mean volume of the breast receiving $\geq 107\%$ of the breast dose (this is 53.5 Gy) was reduced by 20%. Furtheron, the mean volume outside the boost planning target volume (PTV) receiving $\geq 95\%$ of the boost dose (this is 62.7 Gy) was reduced by 54% [28]. In another study, the group of Groningen has recently developed an NTCP model for late physician-rated fibrosis in the boost area and V55 thereby is one of the 4 withheld parameters [15]. Hurkmans et al. have compared treatment plans with tangential field intensity-modulated RT followed by a sequential boost with a SIB technique using inverse optimization. Prescription doses to the boost volume were high, 26 Gy. The volume of the PTV breast, excluding the PTV boost, receiving a dose higher than 95% of the boost dose could be reduced considerably using SIB from 129 cc to 58 cc [29].

To the best of our knowledge, this is the first NTCP model for late AO measured with BCCT.core after BCT. AO results from a combination of breast asymmetry, skin colour differences and scar visibility [8]. Breast asymmetry is primarily dependent on the amount of tissue excised, but radiotherapy-associated fibrosis also has an impact [30,31]. Four large studies have analysed the sub-outcome of breast fibrosis after BCT. Mukesh et al. have pooled patient data from two trials including breast RT with or without a boost and estimated NTCP parameters for breast fibrosis. A small value for the volume parameter ' n ' was obtained (0.01), suggesting that the highest dose part of the DVH is the most important parameter to influence breast fibrosis. In the analysis, however, only limited dose distribution data were available and a calculation model was used to estimate the boost/breast volume ratio and dose distribution [32]. To extend the application of NTCP models for fibrosis after whole-breast irradiation to partial breast irradiation, Avanzo et al. have compared

calculated and observed toxicity rates for patients treated with partial breast irradiation. Radiobiological parameters were determined by fitting fibrosis incidences in published randomized studies and they also obtained a low n -value of 0.06 [33]. These results are conflicting with the data reported by Alexander et al. [34]. In their study, clinical results on breast fibrosis were also collected from available literature and prescribed doses rather than dose distributions were used. The recently developed model by the group of Groningen estimated the risk for late physician-rated \geq grade 2 fibrosis in the boost area following 3D conformal RT with simultaneous integrated boost. 13.4% of the 546 analysed patients developed fibrosis \geq grade 2 in the boost area and the model included age, V55 of the CTV breast and D_{max} [15]. In contrary to the above cited studies, in the present cohort the dose-volume metrics for the whole breast were taken into account, thus including the skin. Pastore et al. have estimated LKB parameters for severe acute radio-induced skin toxicity from dose-surface histograms and concluded a high n -value of 0.4, consistent with a surface effect for skin [35].

AUC values for the two NTCP models in the present study were 0.76 and 0.74 (for the regression analysis and the LKB model, respectively) when training the models on the whole dataset. The AUC corrected after bootstrap sampling was slightly worse for the LKB model (0.74 with 95% CI 0.61; 0.83) compared to the logistic regression model (0.75 with 95% CI 0.64; 0.85). This is probably due to the hotspots taken into account in the EUD calculation. Even with an n -parameter of 1, hotspots affect the EUD/mean dose. The hotspot information in DVHs from patients treated with a brachytherapeutic boost could thus not be eliminated from the LKB model, while it could be eliminated in the regression model by only incorporating V55. As D_{max} is not withheld in both NTCP models, we might suggest that high hot spot doses to a small volume as obtained in brachytherapy boost doses do not seem to impair AO. Based on the logistic regression model, we suggest to implement a new constraint in treatment planning. In case of no seroma or ALND, a late unfavourable AO is expected in less than 10% of breast cancer patients when the volume of breast tissue receiving 55 Gy (EQD2_{3,6}) is limited to 8%.

A true limitation of the study is that there was no baseline AO available so there may have been patients with an unfavourable AO before the start of RT. We tried to overcome this issue by taking those clinicopathological variables correlated with unfavourable AO into account, as reported by Hennigs et al. [10]. Added to that, a prospective ongoing study at our institution revealed a low rate of baseline unfavourable AO after breast-conserving surgery: 3 patients out of 175 (1.7%). Bias could theoretically also have been introduced due to losses in follow-up, however a strict follow-up regimen is followed at our institution, so we believe this would not alter the results. More prospective studies however are warranted and an external and independent validation of the obtained model is important before widespread implementation. Also, there is evidence that patients evaluate AO more positively than objective measures and professionals [7]. Furthermore, the relationship between the

objective AO and quality of life is unclear. Ideally, patient reported outcome measures should be incorporated in the decision-making when weighing up different treatment options. The deformable registration of the brachytherapy CT scan with the external beam planning CT scan for whole breast irradiation might lead to inaccuracies. This deformation of the breast however is inherent to the brachytherapeutic technique with the use of interstitial needles. Furthermore, it is possible that due to the higher radiobiological effectiveness (RBE around 1.3) of brachytherapy the calculated brachytherapy doses were underestimated [36]. On the contrary, the LQ model tends to overestimate the effectiveness of cell killing in high dose rate brachytherapy [37]. The present model should therefore also not be used for extreme hypofractionation schedules with fraction sizes increasing 3.25 Gy without further testing [37]. Bentzen et al. have also reported the RBE for electrons was 0.88 relative to photons at 4.1 mm depth. As the RBE at different depths is unknown, no attempts were made to correct for this [38]. Another important remark is that a lot of dose-volume metrics between V30 and V65 were significantly correlated with unfavourable AO in univariable analysis. Because V55 was the most significant and a clinically relevant parameter, it was included in the multivariable analysis. Furthermore, as described by Marks et al. the correlations between dosimetric parameters are technique dependent [39]. With intensity-modulated radiotherapy for example, dose gradients are typically steep and high-dose volumes small. Thus the dose-volume characteristics and their correlations would be quite different from those in the present cohort. The presented results are based on 3D conformal RT with wedges or field-in-field RT and a sequential external beam or brachytherapeutic boost and they can thus not be extrapolated to other treatment techniques without further testing [39].

In conclusion, we have developed an NTCP model for late AO after BCT. Our data suggest that a seroma and an axillary lymphadenectomy are the most important clinical factors, that the effect of the maximum dose is small ($n = 1.00$) and that the V55 is the most important dose-volume metric to limit unfavourable AO after breast-conserving therapy. The multivariable logistic regression model with three single parameters (seroma, ALND and V55) is considered practically useful. Independent validation is necessary.

Disclosure statement

No potential conflict of interest was reported by the authors.

ORCID

Isabelle Kindts  <http://orcid.org/0000-0002-6570-141X>

References

- [1] Ferlay J, Soerjomataram I, Ervik M, et al. GLOBOCAN 2012 v1.1, Cancer Incidence and [Online] Lyon, France: International Agency for Research on Cancer; 2014. [cited 2017 Oct 9]. Available from: <http://globocan.iarc.fr>
- [2] Fisher B, Anderson S, Bryant J, et al. Twenty-year follow-up of a randomized trial comparing total mastectomy, lumpectomy, and lumpectomy plus irradiation for the treatment of invasive breast cancer. *N Engl J Med.* 2002;347:1233–1241.
- [3] Litière S, Werutsky G, Fentiman IS, et al. Breast conserving therapy versus mastectomy for stage I-II breast cancer: 20 year follow-up of the EORTC 10801 phase 3 randomised trial. *Lancet Oncol.* 2012;13:412–419.
- [4] Darby S, McGale P, Correa C, et al. Effect of radiotherapy after breast-conserving surgery on 10-year recurrence and 15-year breast cancer death: meta-analysis of individual patient data for 10,801 women in 17 randomised trials. *Lancet.* 2011;378:1707–1716.
- [5] Taylor C, Correa C, Duane FK, et al. Estimating the risks of breast cancer radiotherapy: evidence from modern radiation doses to the lungs and heart and from previous randomized trials. *J Clin Oncol.* 2017;35:1641–1649.
- [6] Waljee J, Hu E, Ubel P, et al. Effect of esthetic outcome after breast-conserving surgery on psychosocial functioning and quality of life. *J Clin Oncol.* 2008;26:3331–3337.
- [7] Cardoso M, Cardoso J, Vrieling C, et al. Recommendations for the aesthetic evaluation of breast cancer conservative treatment. *Breast Cancer Res Treat.* 2012;135:629–637.
- [8] Cardoso J, Cardoso M. Towards an intelligent medical system for the aesthetic evaluation of breast cancer conservative treatment. *Artif Intell Med.* 2007;40:115–126.
- [9] Harris J, Levene M, Svensson G, et al. Analysis of cosmetic results following primary radiation therapy for stages I and II carcinoma of the breast. *Int J Radiat Oncol Biol Phys.* 1979;5:257–261.
- [10] Hennigs A, Hartmann B, Rauch G, et al. Long-term objective esthetic outcome after breast-conserving therapy. *Breast Cancer Res Treat.* 2015;153:345–351.
- [11] Collins G, Reitsma J, Altman D, et al. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. *Br J Cancer.* 2015;112:251–259.
- [12] Verhoeven K, Kindts I, Laenen A, et al. A comparison of three different radiotherapy boost techniques after breast conserving therapy for breast cancer. *Breast.* 2015;24:391–396.
- [13] Taylor ME, Perez CA, Halverson KJ, et al. Factors influencing cosmetic results after conservation therapy for breast cancer. *Int J Radiat Oncol Biol Phys.* 1995;31:753–764.
- [14] Poortmans P, Bartelink H, Horiot J, et al. The influence of the boost technique on local control in breast conserving treatment in the EORTC 'boost versus no boost' randomised trial. *Radiother Oncol.* 2004;72:25–33.
- [15] Hammer C, Maduro J, Bantema-Joppe E, et al. Radiation-induced fibrosis in the boost area after three-dimensional conformal radiotherapy with a simultaneous integrated boost technique for early-stage breast cancer: a multivariable prediction model. *Radiother Oncol.* 2017;122:45–49.
- [16] Cardoso M, Cardoso J, Santos A, et al. Factors determining esthetic outcome after breast cancer conservative treatment. *Breast J.* 2007;13:140–146.
- [17] Offersen BV, Boersma LJ, Kirkhove C, et al. ESTRO consensus guidelines on target volume delineation for elective radiation therapy of early stage breast cancer. *Radiother Oncol.* 2015;114:3–10.
- [18] Bentzen S, Agrawal R, Aird E, et al. The UK standardisation of Breast Radiotherapy (START) Trial A of radiotherapy hypofractionation for treatment of early breast cancer: a randomised trial. *Lancet Oncol.* 2008;9:331–341.
- [19] Niemierko A. Reporting and analyzing dose distributions: a concept of equivalent uniform dose. *Med Phys.* 1997;24:103–110.
- [20] American Association of Physicists in Medicine. Report of AAPM task group 166 of the therapy physics committee. The use and QA of biologically related models for treatment planning. March 2012.

- [21] Tucker S, Liu H, Liao Z, et al. Analysis of radiation pneumonitis risk using a generalized Lyman model. *Int J Radiat Oncol Biol Phys.* 2008;72:568–574.
- [22] Kutcher G, Burman C. Calculation of complication probability factors for non-uniform normal tissue irradiation: the effective volume method. *Int J Radiat Oncol Biol Phys.* 1989;16:1623–1630.
- [23] Lyman J. Complication probability as assessed from dose-volume histograms. *Radiat Res Suppl.* 1985;8:S13–S19.
- [24] Hanley J, McNeil B. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology.* 1982;143:29–36.
- [25] Hosmer D, Lemeshow S. *Applied Logistic regression.* New York: John Wiley & Sons Inc. 1989.
- [26] Van Limbergen E, Rijnders A, van der Schueren E, et al. Cosmetic evaluation of breast conserving treatment for mammary cancer. 2. A quantitative analysis of the influence of radiation dose, fractionation schedules and surgical treatment techniques on cosmetic results. *Radiother Oncol.* 1989;16:253–267.
- [27] Senkus E, Kyriakides S, Ohno S, on behalf of the ESMO Guidelines Committee, et al. Primary breast cancer: ESMO clinical practice guidelines for diagnosis, treatment and follow-up. *Ann Oncol.* 2015;26(Suppl5):v8–30.
- [28] Van der Laan H, Dolsma W, Maduro J, et al. Three-dimensional conformal simultaneously integrated boost technique for breast-conserving radiotherapy. *Int J Radiat Oncol Biol Phys.* 2007;68:1018–1023.
- [29] Hurkmans C, Meijer G, van Vliet-Vroegindewey C, et al. High-dose simultaneously integrated breast boost using intensity-modulated radiotherapy and inverse optimization. *Int J Radiat Oncol Biol Phys.* 2006;66:923–930.
- [30] Vrieling C, Collette L, Fourquet A, et al. The influence of patient, tumor and treatment factors on the cosmetic results after breast-conserving therapy in the EORTC 'boost vs. no boost' trial. *EORTC Radiotherapy and Breast Cancer Cooperative Groups. Radiother Oncol.* 2000;55:219–232.
- [31] Van Limbergen E, van der Schueren E, Van Tongelen K. Cosmetic evaluation of breast conserving treatment for mammary cancer. 1. Proposal of a quantitative scoring system. *Radiother Oncol.* 1989;16:159–167.
- [32] Mukesh M, Harris E, Collette S, et al. Normal tissue complication probability (NTCP) parameters for breast fibrosis: pooled results from two randomised trials. *Radiother Oncol.* 2013;108:293–298.
- [33] Avanzo M, Stancanella J, Trovò M, et al. Complication probability model for subcutaneous fibrosis based on published data of partial and whole breast irradiation. *Phys Med.* 2012;28:296–306.
- [34] Alexander MA, Brooks WA, Blake SW. Normal tissue complication probability modelling of tissue fibrosis following breast radiotherapy. *Phys Med Biol.* 2007;52:1831–1843.
- [35] Pastore F, Conson M, D'Avino V, et al. Dose-surface analysis for prediction of severe acute radio-induced skin toxicity in breast cancer patients. *Acta Oncol.* 2016;55:466–473.
- [36] Wu C, Kliauga P, Zaider M, Amols H. Microdosimetric evaluation of relative biological effectiveness for ^{103}Pd , ^{125}I , ^{241}Am , and ^{192}Ir brachytherapy sources. *Int J Radiat Oncol Biol Phys.* 1996;36:689–697.
- [37] Wang J, Huang Z, Lo S, et al. A generalized linear-quadratic model for radiosurgery, stereotactic body radiation therapy, and high-dose rate brachytherapy. *Sci Transl Med.* 2010;2:39ra48.
- [38] Bentzen S, Overgaard M. Relationship between early and late normal-tissue injury after postmastectomy radiotherapy. *Radiother Oncol.* 1991;20:159–165.
- [39] Marks L, Bentzen S, Deasy J, et al. Radiation dose-volume effects in the lung. *Int J Radiat Oncol Biol Phys.* 2010;76(3 Suppl): S70–S76.