ScienceDirect

Contents lists available at sciencedirect.com
Journal homepage: www.elsevier.com/locate/jval

Comparative-Effectiveness Research/HTA

Applying Risk-Based Follow-Up Strategies on the Dutch Breast Cancer Population: Consequences for Care and Costs

Teresa Draeger, MD, PhD, Vinzenz Voelkel, MD, PhD, Catharina G.M. Groothuis-Oudshoorn, PhD, Miha Lavric, PhD, Jeroen Veltman, MD, PhD, Anneriet Dassen, MD, PhD, Liesbeth J. Boersma, MD, PhD, Annemieke Witteveen, PhD, Gabe S. Sonke, MD, PhD, Hendrik Koffijberg, PhD, Sabine Siesling, PhD*

ABSTRACT

Objectives: An important aim of follow-up after primary breast cancer treatment is early detection of locoregional recurrences (LRR). This study compares 2 personalized follow-up scheme simulations based on LRR risk predictions provided by a time-dependent prognostic model for breast cancer LRR and quantifies their possible follow-up efficiency.

Methods: Surgically treated early patients with breast cancer between 2003 and 2008 were selected from the Netherlands Cancer Registry. The INFLUENCE nomogram was used to estimate the 5-year annual LRR. Applying 2 thresholds, they were defined according to Youden's J-statistic and a predefined follow-up sensitivity of 95%, respectively. These patient's risk estimations served as the basis for scheduling follow-up visits; 2 personalized follow-up schemes were simulated. The number of potentially saved follow-up visits and corresponding cost savings for each follow-up scheme were compared with the current Dutch breast cancer guideline recommendation and the observed utilization of follow-up on a training and testing cohort.

Results: Using LRR risk-predictions for 30 379 Dutch patients with breast cancer from 2003 to 2006 (training cohort), 2 thresholds were calculated. The threshold according to Youden's approach yielded a follow-up sensitivity of 62.5% and a potential saving of 62.1% of follow-up visits and €24.8 million in 5 years. When the threshold corresponding to 95% follow-up sensitivity was used, 17% of follow-up visits and €7 million were saved compared with the guidelines. Similar results were obtained by applying these thresholds to the testing cohort of 11 462 patients from 2007 to 2008. Compared with the observed utilization of follow-up, the potential cost-savings decline moderately.

Conclusions: Personalized follow-up schemes based on the INFLUENCE nomogram's individual risk estimations for breast cancer LRR could decrease the number of follow-up visits if one accepts a limited risk of delayed LRR detection.

Keywords: cancer registry, cost-effectiveness, follow-up, health services research, locoregional recurrence, mamma carcinoma, personalized care, prediction model.

VALUE HEALTH. 2020; 23(9):1149–1156

Introduction

Worldwide, breast cancer is the most frequent malignancy found among women.[1] In The Netherlands more than 14 000 women are diagnosed with invasive breast cancer every year.[2] Because of early detection and improved treatment strategies, the survival rates of patients with breast cancer have increased considerably in the last few decades.[3,4] On average, 85% of women diagnosed with breast cancer are alive after 5 years.[4] After treatment with curative intent, every patient receives regular follow-up care aimed at early detection of possible local and regional recurrences to reduce the risk of subsequent distant metastases and improve survival.[5,6] A local recurrence is defined as the reappearance of cancer in the ipsilateral breast or chest wall and a regional recurrence specifies tumor involvement in the regional lymph nodes.[7] According to the current Dutch breast cancer guidelines, follow-up for all women consists of an annual mammography and physical examination in the first 5 years after diagnosis unless a bilateral mastectomy was performed.[8] It has been demonstrated that the risk for local and regional recurrence is associated with various patient and tumor features (eg, breast cancer subtype, age, nodal status) and may change over time.[9-13] Thus personalized follow-up schemes based on individual risk profiles can be more efficient than the current uniform follow-up. Unnecessary follow-up visits leading to further invasive investigations and anxiety in the patients or their relatives can be

* Address correspondence to: Sabine Siesling, PhD; Netherlands Comprehensive Cancer Organisation (IKNL); PO Box 19079, 3501 DB, Utrecht, The Netherlands.
Email: s.siesling@iknl.nl

avoided. As a consequence, the patients' compliance with future follow-up visits might increase and the overall financial burden on the healthcare system could simultaneously be lowered.[14] In 2015, Witteveen et al[15] developed the INFLUENCE nomogram (available on www.evidencio.com),[16] a Time-Dependent Prognostic Nomogram for the Estimation of Annual Risk of Locoregional Recurrence in Early Breast Cancer Patients. For different patient, tumor, and treatment characteristics (such as age, tumor size, nodal involvement, grade, estrogen receptor [ER]- and progesterone receptor [PR]-status, multifocality, radiotherapy, chemotherapy, and endocrine therapy), the INFLUENCE nomogram estimates the individual overall 5-year risk, and the conditional annual risks, of developing a local or regional recurrence within 5 years after diagnosis. The predictions of INFLUENCE can be used to discern high-risk from low-risk patients based on a specific threshold. Using a large representative cohort of patients with breast cancer from The Netherlands, the present study proposes and simulates 2 INFLUENCE-based follow-up strategies and aims to quantify their impact on the healthcare system by calculating their potential to save follow-up visits and costs relative to the sensitivity of the follow-up program on a population level.

Methods

Patients

Patients were selected from the Netherlands Cancer Registry (NCR), a nationwide population-based cancer registry, which has collected data for all newly diagnosed malignancies in The Netherlands since 1989. The information stored there includes demographics, tumor characteristics, and treatment characteristics and is gathered from the patient files by specially trained data-managers.

To be included for further analyses, patients had to fulfill all of the inclusion criteria dictated by the INFLUENCE nomogram[15]: women with primary invasive breast cancer (International Classification for Disease-Oncology-10 C50), diagnosed between 2003 and 2008, with cM0 at time of diagnosis, resection with curative intent (no macroscopic residue), and no neo-adjuvant treatment. Furthermore, only patients with a locoregional recurrence (LRR) as the first recurrence event and a follow-up time of at least 5 years were included.

The study population was divided into 2 subsets: a training cohort comprising patients diagnosed between 2003 and 2006 (this cohort contained patients who were originally used to develop the INFLUENCE nomogram) to determine 2 thresholds for individualized INFLUENCE-based follow-up schemes, and a testing cohort comprising patients diagnosed in 2007 and 2008 to externally evaluate the schemes' performance.

Statistical Analysis

To obtain the cohort's interval-specific LRR rates, the life-table method based on 5 annual observation periods was used. The individual recurrence risks per year after surgery were estimated for every patient using the algorithm of the INFLUENCE nomogram. To provide an overview of these risk predictions, the mean, as well as the median, annual predicted risks with the corresponding interquartile ranges (IQR) based on all individual risks of the training cohort were calculated.

Using the risk predictions in clinical practice requires decision thresholds. Assuming that an LRR should be detected as quickly as possible, a follow-up visit at the end of the annual follow-up periods in which the corresponding predicted risk of a patient exceeds a certain threshold should be performed. Thus personalized follow-up schemes ranging between 0 and 5 follow-up visits in 5 follow-up years were generated. Because this article focuses on the design of risk-based follow-up schemes and not on the development of better examination techniques, it was further assumed that at every follow-up visit a hypothetical examination with perfect detection sensitivity and specificity for LRRs would be performed. Moreover, the time-window for successful LRR detection was set to a maximum of 1 year. Again, this simplifying definition follows the assumption that an LRR should be detected as quickly as possible. This means that if a patient developed an LRR in the second follow-up year, and the next follow-up was supposed to take place at the end of the same year, this recurrence was counted as detected; however, if the next follow-up visit was only scheduled for the third year according to the personalized scheme, the recurrence would have been detected 1 year too late and, thus, was counted as missed. Additionally, potential self-detections by the patients were not part of the simulation study's setup. Following these assumptions, the theoretical number and associated costs of follow-up visits according to the current Dutch guideline recommendation were calculated. These figures serve as a reference for the proposed personalized schemes.

The receiver operating curve of the training cohort, incorporating all individual annual risk estimations combined, was used to determine 2 thresholds for the annual risk predictions, corresponding to Youden's J-statistic (Y)[17] and a predefined follow-up sensitivity of 95% (S).

For Youden's J-statistic, the follow-up sensitivity $q(c)$ and specificity $s(c)$ for a given threshold value c were the probabilities of correctly identifying a person's recurrence status, with R^+/R^- being the true recurrence status of the training cohort[18]:

$$q(c) = Probability\ (prediction\ = +|R^+)$$

$$s(c) = Probability\ (prediction\ = -|R^-)$$

The optimal threshold according to this approach corresponded to the joint-optimum for follow-up sensitivity and specificity:

$$max\ (J) = max\ \{q(c) + s(c) - 1\}$$

$$= max\ \{q(c) - (1 - s(c))\}$$

In terms of the predefined follow-up sensitivity of 95%, if follow-up efficiency was increased, it would inevitably lead to a decreased sensitivity below 100%. Nevertheless, previous surveys revealed that in The Netherlands only about 90% of the patients make use of the recommended annual follow-up visits in the first 5 years after treatment.[19] It can be assumed that the current underutilization takes place at random, leading to a follow-up sensitivity of 90% if the simplified assumption of perfect diagnostic procedures is applied. Therefore, it was decided to set the follow-up sensitivity threshold to 95%, which is better than the actual observed sensitivity but still leaves room to improve efficiency.

In a second step, the newly developed strategies were simulated based on the testing cohort to evaluate their external validity. Following the previously introduced assumptions, the observed sensitivity and the number of recommended follow-up visits were calculated for the personalized follow-up schemes based on both suggested thresholds. The associated costs were estimated by quantifying resource use (diagnostic procedures, clinical follow-up visits). For the cost calculations, official Dutch hospital price lists were used.[20,21] Acknowledging that one of the main aims of follow-up after primary breast cancer is the early

detection of LRR and that the INFLUENCE nomogram only offers personalized predictions for LRRs and not for second primaries, only diagnostic procedures for the affected side were taken into account. Following this simplified assumption, all patients were assigned a clinical follow-up visit according to their individual follow-up scheme, but only patients who underwent breast-conserving surgery were eligible for a follow-up mammogram.

Still following the simplified assumptions concerning the diagnostic procedures, the follow-up sensitivity, the number, and the costs of the personalized follow-up visits were compared with the uniform annual follow-up strategy recommended by the current breast cancer guideline of The Netherlands.[8] Another comparison taking into account the observed follow-up utilization of 90% was performed to draw a more realistic picture of the saving potential.

For the analysis, IBM SPSS 25 (IBM Corp., SPSS for Windows, Armonk, NY), R version 3.5.1 (R Foundation for Statistical Computing, Vienna, Austria; http://www.R-project.org/), and the R packages "Hmisc" (www.CRAN.R-project.org/package=Hmisc) and "pROC"[22] were used.

Results

Study Population

In The Netherlands, 51 128 patients who were diagnosed with primary invasive breast cancer between 2003 and 2008 fulfilled the inclusion criteria for the INFLUENCE nomogram. Of all the patients meeting the basic inclusion criteria, 9287 patients (18.2%) had to be excluded because of missing data on at least 1 of the following items: TNM-stage, grading, ER status, PR status, or lymph node status. The definitive study population comprised 41 841 individual patients (Fig. 1). Comparing the 5-year LRR rates of all included and excluded patients, no substantial differences could be seen (5-year LRR rates included vs excluded patients were 3.4% vs 4.3%).

The study population was divided into a training cohort consisting of 30 379 patients diagnosed between 2003 and 2006 and a testing cohort with 11 462 patients diagnosed in 2007 and 2008. The training and testing cohort were highly similar when it came to patient, tumor, and treatment characteristics (Table 1).

In the training cohort, 1104 patients (3.6%) developed an LRR as first event within 5 years after diagnosis compared with 338 (2.9%) in the testing cohort. The mean annual LRR risk predicted by the INFLUENCE nomogram varied. Within the training cohort, it reached its maximum of 1.4% in the second year and its minimum in the fifth year (0.39%). The median annual predicted risk showed a comparable trend over time (year 1: 0.26%; IQR: 0.11% to 0.55%; year 2: 1.05%; IQR: 0.7% to 1.8%; year 5: 0.33%; IQR: 0.23% to 0.48%; Fig. 2). In the testing cohort, similar results were observed.

Threshold Determination and Screening Performance in the Training Cohort

Applying Youden's approach to the training cohort's receiver operating characteristic curve yielded a threshold of 0.6% for the individual annual risk. The following example illustrates how such a personalized follow-up scheme would look on an individual level for a random patient with the following characteristics: age 50 to 59, tumor size <2 cm, 1 to 3 nodes involved, grade 1, hormone status ER and PR negative, no multifocality, adjuvant radiotherapy, no adjuvant chemotherapy, and no adjuvant endocrine therapy. For this patient, the INFLUENCE nomogram estimates the following annual risks—year 1: 0.26%, year 2: 1.23%, year 3: 0.74%, year 4: 0.2%, and year 5: 0.18% (each annual risk estimation follows the condition that no LRR occurred in the previous year). Applying the annual risk threshold of 0.6% to this patient's risk estimations would result in personalized follow-up recommendations in the second and the third year (Fig. 3).

If only women with an annual risk exceeding Youden's threshold value of 0.6% were to receive further diagnostics at the end of the corresponding year, 690 of 1104 LRRs would be detected by follow-up, which is equivalent to a follow-up

Figure 1. Flowchart of patient selection.

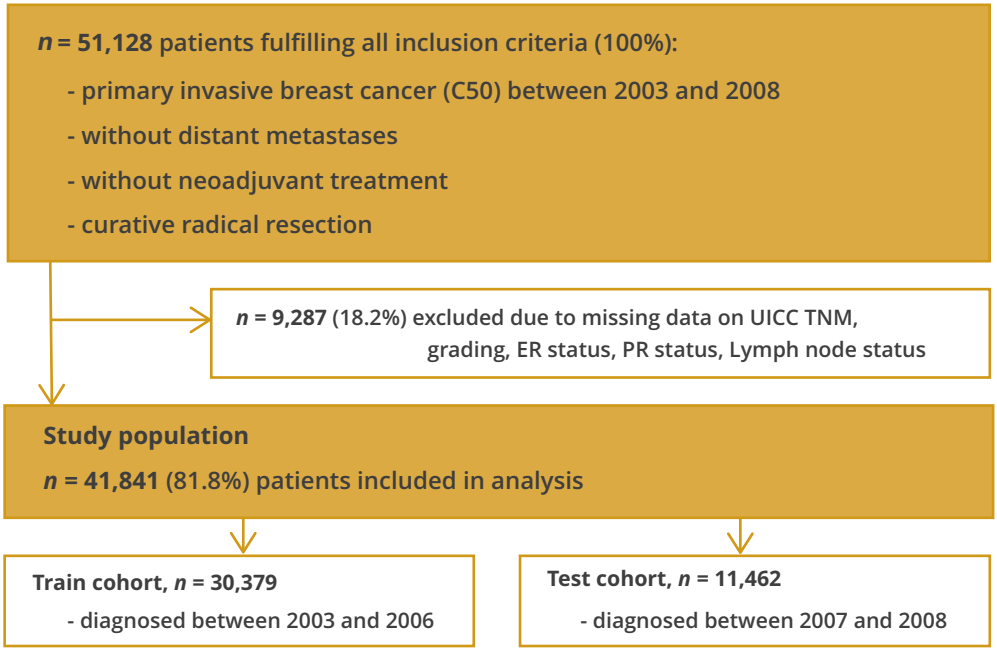


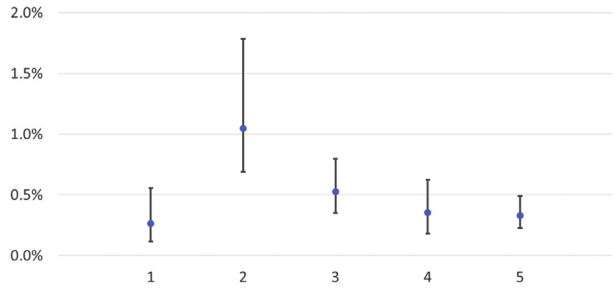
Table 1. Patient, tumor, and treatment characteristics.

Patient and tumor characteristics	Training cohort (2003-2006) *n* = 30 379		Testing cohort (2007-2008) *n* = 11 462	
	n	%*	*n*	%*
Age category at diagnosis (years)				
<50	7879	25.9%	2691	23.5%
50-59	8495	28.0%	3063	26.7%
60-69	6941	22.8%	2942	25.7%
≥70	7064	23.3%	2766	24.1%
Histologic type				
Ductal	23 815	78.4%	8927	77.9%
Lobular	3017	9.9%	1192	10.4%
Mixed	1777	5.8%	783	6.8%
Other	1770	5.8%	560	4.9%
Grading				
1	6684	22.0%	2881	25.1%
2	13 849	45.6%	5124	44.7%
3	9846	32.4%	3457	30.2%
UICC stage				
1	13 294	43.8%	5473	47.7%
2	13 483	44.4%	4842	42.2%
3	3602	11.9%	1147	10.0%
Tumor size (mm)				
<20	18 467	60.8%	7339	64.0%
20-50	11 062	36.4%	3878	33.8%
>50	850	2.8%	245	2.1%
Multifocal				
No	26 125	86.0%	9694	84.6%
Yes	4254	14.0%	1768	15.4%
Lymph node status				
Negative	18 690	61.5%	7449	65.0%
1-3 positive	8328	27.4%	2945	25.7%
>3 positive	3361	11.1%	1068	9.3%
ER status				
Negative	5461	18.1%	1930	16.9%
Positive	24 630	80.9%	9501	83.1%
Unknown	288		31	
PR status				
Negative	9604	33.2%	3505	31.6%
Positive	19 284	66.8%	7592	68.4%
Unknown	1491		365	
HR status				
negative	5038	16.6%	1776	15.5%
positive	25 341	83.4%	9686	84.5%

Treatment Characteristics	Training cohort (2003-2006) *n* = 30 379		Testing cohort (2007-2008) *n* = 11 462	
	n	%*	*n*	%*
Type of surgery				
Breast-conserving	17 009	56.0%	6699	58.4%
Mastectomy	13 370	44.0%	4763	41.6%
Chemotherapy				
No	19 405	63.9%	7209	62.9%
Yes	10 974	36.1%	4253	37.1%
Radiotherapy				
No	10 518	34.6%	3811	33.2%
Yes	19 861	65.4%	7651	66.8%
Endocrine therapy				
No	16 856	55.5%	6108	53.3%
Yes	13 523	44.5%	5354	46.7%

ER indicates estrogen receptor; HR, hormone receptor; n, number; PR, progesterone receptor; UICC, Union for International Cancer Control.
*Percentages do not consider patients with unknown variable values.

Figure 2. Median annual LRR risk over all patients with corresponding IQR of the training cohort predicted by the INFLUENCE nomogram.



IQR indicates interquartile range; LRR, locoregional recurrences.

sensitivity of 62.5%. In total, only 57 608 follow-up visits would be necessary. In the second year, 24 870 (43.2%) visits would take place. In the third year, 12 565 (21.8%) would take place. In the other years, less than 30% of the patients would be recommended to have a follow-up visit. On the individual level, great variation was seen regarding the recommended assignment of follow-up visits. There were 4475 patients (14.7%) without a recommendation for any follow-up visit within 5 years; the mean estimated cumulative 5-year LRR risk of this group was 1.3%. On the other hand, there were 2069 women (6.8%) who were assigned 5 follow-up visits; they had a mean estimated cumulative 5-year LRR risk prediction of 9.1%. Compared with the Dutch breast cancer guideline recommendations, the overall number of follow-up visits would have been reduced by 94 287 (62.1%). Following the previously mentioned assumptions, this is equivalent to a potential cost saving of approximately €24.8 million (64.4%) over the whole study population for all 5 years (note that the average cost of an individual clinical follow-up visit is €200.70 and the average cost of a mammogram is €94.58€).[20,21] On average, €163 per patient could be saved each year. Compared with the currently observed follow-up utilization, 57.9% of follow-up visits and 60.4% of costs could be saved if one would be willing to accept an absolute sensitivity reduction of 27.5%.

To achieve a predefined follow-up sensitivity of 95% would require an annual risk threshold of 0.19% based on the training cohort's risk estimations. If a personalized follow-up scheme were to rely on this value, 1049 of 1104 LRRs would be detected during the recommended 126 016 follow-up visits. This would result in a saving of 25 879 (17.0%) follow-up visits and €7 million (18.2%) over 5 years compared with the current guideline's recommendation (Table 2). In the first year €3.2 million could be saved, whereas in the second year nearly every woman would be eligible for a follow-up visit, which would yield almost no cost savings. The potential cost saving would be €46 per patient each year on average. Compared with the currently observed follow-up

Figure 3. Example patient. Patient and tumor characteristics (left). Annual risks predicted by the INFLUENCE nomogram and simulation of an individual follow-up scheme based on Youden's approach.

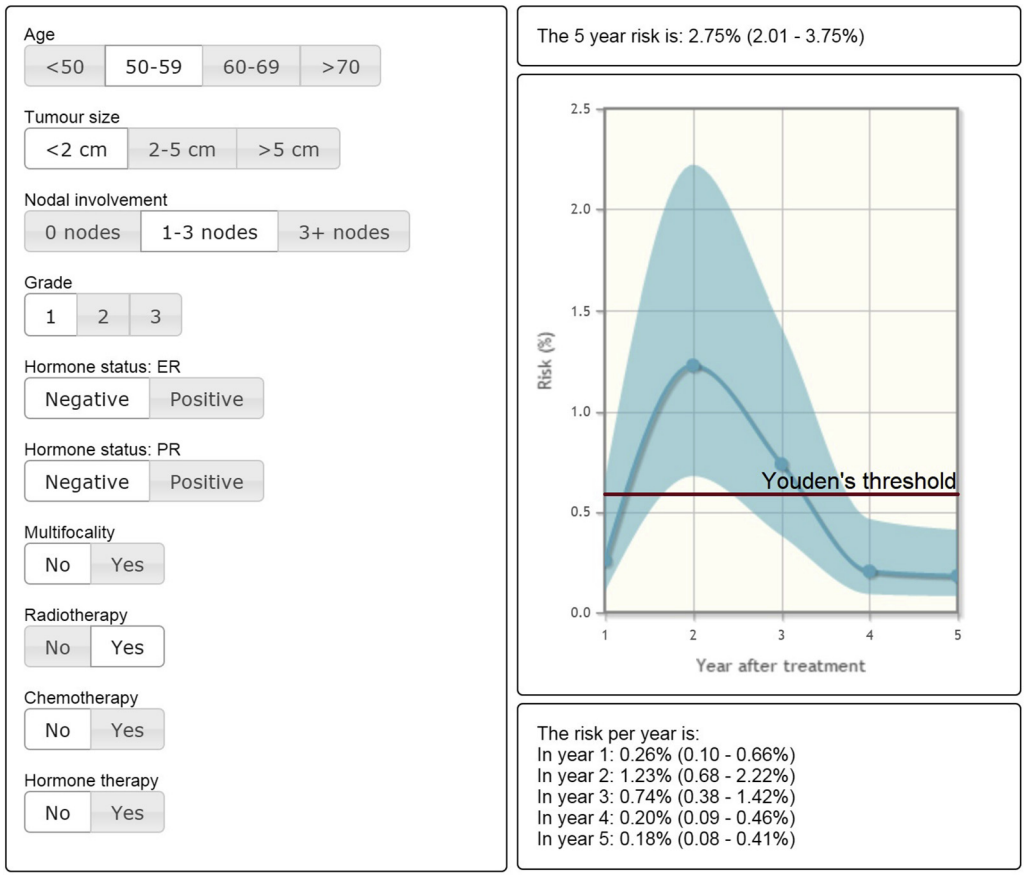


Table 2. Cost-effectiveness per year and overall cost-effectiveness.

Follow-up patterns according to INFLUENCE per year		
Time	**Number of follow-ups**	**Estimated costs**
Training cohort (2003-2006) *n* = 30 379		
95% sensitivity approach		
Year 1:	18 897	€4.5 million
Year 2:	30 365	€7.7 million
Year 3:	29 017	€7.3 million
Year 4:	22 516	€5.6 million
Year 5:	25 221	€6.4 million
Youden's approach		
Year 1:	6969	€1.5 million
Year 2:	24 870	€6.1 million
Year 3:	12 565	€2.9 million
Year 4:	8821	€2.2 million
Year 5:	4383	€1.1 million
Testing cohort (2007-2008) *n* = 11 462		
95% sensitivity approach		
Year 1:	6357	€1.4 million
Year 2:	11 446	€2.9 million
Year 3:	10 734	€2.7 million
Year 4:	7899	€2.0 million
Year 5:	9138	€2.3 million
Youden's approach		
Year 1:	2111	€0.5 million
Year 2:	8907	€2.2 million
Year 3:	4127	€0.9 million
Year 4:	2735	€0.7 million
Year 5:	1367	€0.3 million

Overall comparison of different follow-up approaches		
Number of follow-ups	**Estimated costs**	**Sensitivity**
Training cohort (2003-2006) *n* = 30 379		
Guideline recommendation		
151 895 (100%)	€38.5 million (100%)	100%
Observed utilization*		
136 706 (90.0%)	€34.7 million (90.0%)	90.0%
INFLUENCE - 95% sensitivity approach		
126 016 (83.0%)	€31.5 million (81.8%)	95.0%
INFLUENCE - Youden's approach		
57 608 (37.9%)	€13.7 million (35.6%)	62.5%
Testing cohort (2007-2008) *n* = 11 462		
Guideline recommendation		
57 310 (100%)	€14.7 million (100%)	100%
Observed utilization*		
51 579 (90.0%)	€13.2 million (90.0%)	90.0%
INFLUENCE - 95% sensitivity approach		
45 574 (79.5%)	€11.5 million (78.2%)	93.5%
INFLUENCE - Youden's approach		
19 247 (33.6%)	€4.6 million (31.3%)	57.1%

*Estimated values.

utilization, 7.8% of follow-up visits and 9.1% of costs could be saved and the absolute follow-up sensitivity would also increase by 5%.

Validation and Screening Performance in the Testing Cohort

In a second step, the thresholds calculated on the training cohort were applied to the testing cohort with patients diagnosed in 2007 and 2008. Using Youden's threshold of 0.6% to develop a follow-up scheme for the testing cohort resulted in a follow-up sensitivity of 57.1% (193 of 338 LRRs would be detected by follow-up). At the same time, 19 247 follow-up visits would have to be performed compared with 57 310 as recommended by the guidelines. Consequently, 38 063 (66.4%) follow-up visits and €10.1 million (68.7%) could be saved. In the first and the fifth year, over €2.4 million could be saved. In the second year, which again features follow-up recommendations for most of the women, the potential cost savings would be quite limited (€0.7 million; see Table 2). On average, €176 could be saved per patient each year. Compared with the currently observed follow-up utilization, 62.7% of follow-up visits and 65.2% of costs could be saved if one would be willing to accept an absolute sensitivity reduction of 32.9%.

Applying the 0.19% threshold corresponding to a 95% follow-up sensitivity (in the training cohort) would yield an external follow-up sensitivity for the testing cohort of 93.5%, equivalent to 316 out of 338 LRRs detected by follow-up. With this threshold, 45 574 follow-up visits would have to be performed, resulting in a saving of 11 736 follow-up visits (20.5%) and a potential overall cost saving of €3.2 million (21.8%), compared with the current guidelines. Concerning annual cost reduction, the highest amount could be saved in the first year (€1.5 million), whereas in the second year again almost no follow-up visits would be saved (see Table 2). The potential cost savings per patient would be €59 per year on average. Compared with the currently observed follow-up utilization, 11.6% of follow-up visits and 13.1% of costs could be saved and the absolute follow-up sensitivity would also increase by 3.5%.

Discussion

The present study simulates different personalized follow-up schemes for breast cancer LRR based on predictions of the INFLUENCE nomogram and therefore provides an evaluation of its potential effect in daily clinical practice. Depending on the thresholds applied to the individual risk predictions to generate personalized schemes, considerable reductions of clinical visits and associated costs could be achieved. Taking into account the current underutilization of the uniform annual follow-up program, the INFLUENCE nomogram even offers the chance to save money and increase the follow-up sensitivity at the same time. Moreover, the thresholds for an annual follow-up visit suggested in this article of 0.19% and 0.6% LRR risk per year are very conservative because the risk of a healthy woman in the general population developing breast cancer is 0.36%.[23]

The large sample size in the training and testing cohort are a considerable strength of this study. All of the simulations were based on a highly representative nationwide cohort obtained from the NCR. The results of this simulation study show that the predictions of INFLUENCE together with predefined thresholds can be used to effectively personalize breast cancer follow-up by discerning high-risk from low-risk patients. The fact that the performance of the newly developed strategies is very similar for the training cohort, which contains the patients on which the INFLUENCE nomogram was originally built, and the independent testing cohort underlines the generalizability of our findings.

To be able to provide the readers of this study with an estimation concerning potential cost savings, a simplified cost analysis was shown. The assumption that every follow-up imaging for patients with breast-conserving therapy was a mammogram is a simplification of daily practice, because there are patients or situations in which other procedures like sonography or the more expensive magnetic resonance imaging (MRI) might seem more appropriate or would be additionally used after a suspect

mammogram. In patients who had undergone mastectomy or in the case of an inconclusive mammogram, different or additional diagnostics like ultrasound or MRI could be applied, which were not part of the cost analysis. We also did not take into account the patient's travel costs and nonproductive time, which might have further augmented the cost savings. Moreover, additional costs caused by false positive findings if the current "full screening" recommendation was applied were not taken into account. On the other hand, additional costs that may arise as a consequence of a later detection of a LRR, and may counterbalance the benefit of the savings, were not considered either. Nevertheless, a recent survey by Lu et al showed that even in the current situation where every patient is eligible for an annual follow-up visit, underutilization leading to a suboptimal detection rate has taken place.[24] If a sufficiently high threshold, such as in our 95% sensitivity condition example, is chosen, INFLUENCE-based follow-up improves the allocation of clinic visits, leading to cost reductions and a higher sensitivity at the same time.

In this context, it is also important to look at the sensitivity of the diagnostic tools. In our simulation model, we assumed that a follow-up diagnostic procedure would detect all existing LRRs; however, there is no imaging method with a diagnostic sensitivity of 100%. The actual diagnostic sensitivity of a mammogram in women with a history of breast cancer is only 65.4%.[25] Nevertheless, this does not bias the relative comparison of the suggested individualized follow-up schemes and the current guideline, because both policies were equally assumed to rely on "perfect" mammograms as standard imaging procedure.

Another simplification was the definition of "detections" and "missing cases." The detection rate (follow-up sensitivity) was defined as the percentage of LRRs, which are detected at the end of the 365-day period in which they occurred. This is because the NCR data only includes the actual diagnosis dates. They do not leave room for any interpretation in terms of "how many days earlier it would have been possible to detect a recurrence" or "when would it have been too late to detect the recurrence." It might also be reasonable to change the model architecture of INFLUENCE toward making continuous rather than point-estimator predictions and, thus, enhance flexibility.

The external validity of the INFLUENCE nomogram's predictions has recently been demonstrated using a large German data set.[26] It was exclusively built to estimate the risk for LRRs, and consequently no predictions for the occurrence of second primary tumors or distant metastases can be done. Because the aim of the present study was to propose an application of the INFLUENCE nomogram in daily clinical practice, the cost analysis only involved the corresponding diagnostic procedures for the affected side. Nevertheless, breast cancer follow-up in clinical practice also aims to detect second primary tumors at an early stage. Of course, screening for second primary tumors uses the same diagnostic procedures, which was not part of the present study's simplified setup. Nevertheless, even women with a low risk for LRR are still at risk to develop a second primary tumor. Currently, the INFLUENCE nomogram cannot quantify this risk. Until this is possible, it might be reasonable to act cautiously and schedule biannual follow-up visits for patients with a low LRR-risk below the chosen follow-up threshold. Thus they would not be disadvantaged compared with woman from the general population, whho undergo a biannual screening program. Kaas et al[27] found that a biannual screening for second primary tumors is noninferior to annual screenings.

Furthermore, the INFLUENCE nomogram was not designed for patients with neo-adjuvant treatment and considers only a restricted number of variables, which might not cover all of the characteristics of a patient that might influence recurrences. Additional information (eg, family history, known genetic features, or comorbidity state), professional experience, or patient preferences can lead to a follow-up schedule not based on the INFLUENCE-risk predictions. Additionally, it must be stated that the algorithm of the INFLUENCE nomogram is based on patients from 2003 to 2006. Change of treatment modalities (ie, the Trastuzumab antibody-therapy for human epidermal growth factor receptor 2 (HER2) positive patients, which was introduced in 2005) could alter the original prediction model. Although HER2 status did not add statistical power to the model when it was originally built, a re-evaluation of this variable with more up-to-date data might be considered. The model performance and its clinical importance may be enhanced by adding more flexibility to the model architecture and by incorporating additional risk factors for LRR and second primary tumors as an outcome.

The aim of the personalized follow-up schemes was to lower the burden for the individual patient and for the entire healthcare system. Notwithstanding, a personalized follow-up approach should not only aim at minimizing the costs. A patient's security and quality of life are equally important. If a patient's risk is verifiably low, it can definitely be reasonable to omit overly frequent follow-up visits, even when a limited risk of delayed LRR-detection has to be accepted. Unnecessary investigations leading to false positive results and overtreatment can negatively affect a patient's quality of life.[14] Therefore, the INFLUENCE nomogram could be a valuable aid for clinical decision making. The term "aid" is important in this context. Like any guideline or prediction model, an INFLUENCE-based personalized follow-up scheme can only serve as a recommendation and support for shared decision making. After all, the underlying algorithm cannot take every eventuality into account, because it was generated using the observational data of a selected patient group for a specific outcome.

Conclusion

Personalized follow-up schemes based on the INFLUENCE nomogram's individual risk estimations for locoregional breast cancer recurrences may actively support health professionals in daily decision making. This study demonstrated that screening efficiency can be improved when focusing on high-risk patients, if one is willing to accept a limited risk of delayed LRR detection. In the future, the model performance may be enhanced by adding more flexibility to the model architecture and by incorporating additional risk factors for LRR, in addition to second primary tumors, as an outcome.

Article and Author Information

Accepted for Publication: May 26, 2020

Published Online: August 14, 2020

doi: https://doi.org/10.1016/j.jval.2020.05.012

Authors' Affiliations: Tumor Center Regensburg/ University of Regensburg, Institute for Quality Control and Health Services Research, Regensburg, Germany (Draeger, Voelkel); Department of Health Technology and Services Research, Technical Medical Centre, University of Twente, Enschede, The Netherlands (Draeger, Voelkel, Groothuis-Oudshoorn, Lavric, Koffijberg, Siesling); Department of Radiology, Ziekenhuisgroep Twente, Almelo/Hengelo, The Netherlands (Veltman); Department of Surgery, Medisch Spectrum Twente, Enschede, The Netherlands (Dassen); Department of Radiation Oncology (Maastro), GROW School for Oncology and Developmental Biology, Maastricht University Medical Centre, Maastricht, The Netherlands (Boersma); Personalized eHealth Technology, University of Twente, Enschede, The Netherlands (Witteveen); Department Medical Oncology, Netherlands Cancer Institute, Amsterdam, The Netherlands (Sonke); Department of Research and Development,

Netherlands Comprehensive Cancer Organisation (IKNL), Utrecht, The Netherlands (Siesling).

Author Contributions: *Concept and design:* Draeger, Groothuis-Oudshoorn, Dassen, Sonke, Koffijberg, Siesling
Acquisition of data: Draeger, Siesling
Analysis and interpretation of data: Draeger, Völkel, Groothuis-Oudshoorn, Lavric, Boersma, Witteveen, Siesling
Drafting of the manuscript: Draeger, Völkel, Groothuis-Oudshoorn, Veltman, Dassen, Witteveen, Sonke, Koffijberg, Siesling
Critical revision of the paper for important intellectual content: Draeger, Völkel, Groothuis-Oudshoorn, Lavric, Dassen, Boersma, Witteveen, Sonke, Koffijberg, Siesling
Statistical analysis: Draeger, Völkel, Lavric, Veltman
Provision of study materials or patients: Siesling
Obtaining funding: Völkel
Administrative, technical, or logistic support: Siesling
Supervision: Lavric, Veltman, Boersma, Siesling

Conflict of Interest Disclosures: Dr Sonke reported receiving institutional research support from AstraZeneca, Merck, Novartis, and Roche outside the submitted work. No other disclosures were reported.

Funding/Support: This work was supported by the Pioneers in Health Care Innovation Fund 2017. Vinzenz Voelkel received funding by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – project number 417891978.

Role of the Funder/Sponsor: The funder 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.

Acknowledgment: The authors would like to thank the registration teams of the Netherlands Comprehensive Cancer Organisation for the collection of data for the Netherlands Cancer Registry and the scientific staff of the Netherlands Cancer Registry.

REFERENCES

1. Stewart BW, Wild CP, eds. *World Cancer Report 2014*. Lyon, France: International Agency for Research on Cancer/World Health Organization; 2014.
2. IKNL - integraal kankercentrum Nederland Cijfers over kaner: Incidentie borst. https://www.cijfersoverkanker.nl/selecties/incidentie_borst/img5c597bed2fc4f. Accessed March 25, 2019.
3. Holleczek B, Brenner H. Trends of population-based breast cancer survival in Germany and the US: decreasing discrepancies, but persistent survival gap of elderly patients in Germany. *BMC Cancer*. 2012;12:317.
4. IKNL - integraal kankercentrum Nederland Cijfers over kanker: Overleving borst. https://www.cijfersoverkanker.nl/selecties/Overleving_borst/img5c597ecc567a0. Accessed March 25, 2019.
5. Lu WL, Jansen L, Post WJ, et al. Impact on survival of early detection of isolated breast recurrences after the primary treatment for breast cancer: a meta-analysis. *Breast Cancer Res Treat*. 2009;114(3):403–412.
6. van der Sangen MJC, Poortmans PMP, Scheepers SWM, et al. Prognosis following local recurrence after breast conserving treatment in young women with early breast cancer. *Eur J Surg Oncol*. 2013;39(8):892–898.
7. Moossdorff M, van Roozendaal LM, Strobbe LJA, et al. Maastricht Delphi consensus on event definitions for classification of recurrence in breast cancer research. *J Natl Cancer Inst*. 2014;106(12):dju288.
8. NABON. Breast Cancer, Dutch Guideline, Version 2.0. http://www.oncoline.nl/mammacarcinoom. Accessed December 20, 2020.
9. Wallgren A, Bonetti M, Gelber RD, et al. Risk factors for locoregional recurrence among breast cancer patients: results from International Breast Cancer Study Group Trials I through VII. *J Clin Oncol*. 2003;21(7):1205–1213.
10. Voogd AC, van Oost FJ, Rutgers EJT, et al. Long-term prognosis of patients with local recurrence after conservative surgery and radiotherapy for early breast cancer. *Eur J Cancer*. 2005;41(17):2637–2644.
11. Jagsi R, Raad RA, Goldberg S, et al. Locoregional recurrence rates and prognostic factors for failure in node-negative patients treated with mastectomy: implications for postmastectomy radiation. *Int J Radiat Oncol Biol Phys*. 2005;62(4):1035–1039.
12. Lin P-H, Yeh M-H, Liu L-C, et al. Clinical and pathologic risk factors of tumor recurrence in patients with node-negative early breast cancer after mastectomy. *J Surg Oncol*. 2013;108(6):352–357.
13. van Maaren MC, Munck L de, Strobbe LJ, et al. 10-year recurrence rates for breast cancer subtypes in the Netherlands: a large population-based study. *Int J Cancer*. 2019;144(2):263–272.
14. Puglisi F, Fontanella C, Numico G, et al. Follow-up of patients with early breast cancer: is it time to rewrite the story? *Crit Rev Oncol Hematol*. 2014;91(2):130–141.
15. Witteveen A, Vliegen IMH, Sonke GS, et al. Personalisation of breast cancer follow-up: a time-dependent prognostic nomogram for the estimation of annual risk of locoregional recurrence in early breast cancer patients. *Breast Cancer Res Treat*. 2015;152(3):627–636.
16. Witteveen A. Influence: five-year locoregional recurrence risk in breast cancer patients. https://www.evidencio.com/models/show/562. Accessed March 4, 2019.
17. Youden WJ. Index for rating diagnostic tests. *Cancer*. 1950;3(1):32–35.
18. Perkins NJ, Schisterman EF. The inconsistency of "optimal" cutpoints obtained using two criteria based on the receiver operating characteristic curve. *Am J Epidemiol*. 2006;163(7):670–675.
19. Lu W, Jansen L, Schaapveld M, et al. Underuse of long-term routine hospital follow-up care in patients with a history of breast cancer? *BMC Cancer*. 2011;11:279.
20. Erasmus Medical Center Passantenprijslijst 2018: Versie1-RZ18B. https://www.erasmusmc.nl/nl-nl/passantenprijslijst. Accessed March 4, 2019.
21. Nederlandse Zorgautoriteit Tariefbeschikking 2018: Eerstelijnsdiagnostiek. https://www.nza.nl/. Accessed April 10, 2019.
22. Robin X, Turck N, Hainard A, et al. pROC: an open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinformatics*. 2011;12:77.
23. van der Waal D, Verbeek ALM, den Heeten GJ, et al. Breast cancer diagnosis and death in the Netherlands: a changing burden. *Eur J Public Health*. 2015;25(2):320–324.
24. Lu W, Bock GH de, Schaapveld M, et al. The value of routine physical examination in the follow up of women with a history of early breast cancer. *Eur J Cancer*. 2011;47(5):676–682.
25. Houssami N, Abraham LA, Miglioretti DL, et al. Accuracy and outcomes of screening mammography in women with a personal history of early-stage breast cancer. *JAMA*. 2011;305(8):790–799.
26. Voelkel V, Draeger T, Groothuis-Oudshoorn CGM, et al. Predicting the risk of locoregional recurrence after early breast cancer: an external validation of the Dutch INFLUENCE-nomogram with clinical cancer registry data from Germany. *J Cancer Res Clin Oncol*. 2019;145(7):1823–1833.
27. Kaas R, Hart AA, Besnard AP, et al. Impact of mammographic interval on stage and survival after the diagnosis of contralateral breast cancer. *Br J Surg*. 2001;88(1):123–127.