



The impact of tumor detection method on genomic and clinical risk and chemotherapy recommendation in early hormone receptor positive breast cancer



Yael Bar ^a, Kfir Bar ^{b, c}, Itay Itzhak ^c, Chen Shitrit Niselbaum ^d, Nachum Dershowitz ^c, Eliya Shachar ^a, Ahuva Weiss-Meilik ^d, Orit Golan ^e, Ido Wolf ^a, Tehillah Menes ^{f, **, 1}, Amir Sonnenblick ^{a, *, 1}

^a Oncology Division, Tel Aviv Sourasky Medical Center and Sackler Faculty of Medicine, Tel Aviv University, Tel Aviv, Israel

^b School of Computer Science, The College of Management, Rishon LeZion, Israel

^c School of Computer Science, Tel Aviv University, Tel Aviv, Israel

^d Data Science Center, Tel-Aviv Sourasky Medical Center, Tel-Aviv, Israel

^e Radiology Department, Tel Aviv Sourasky Medical Center and Sackler Faculty of Medicine, Tel Aviv University, Tel Aviv, Israel

^f Department of Surgery, Chaim Sheba Medical Center, Tel-Hashomer and Sackler Faculty of Medicine, Tel Aviv University, Tel Aviv, Israel

ARTICLE INFO

Article history:

Received 12 April 2021

Received in revised form

16 August 2021

Accepted 3 September 2021

Available online 4 September 2021

Keywords:

Breast cancer

Tumor detection method

OncotypeDX

Genomic risk

Clinical risk

Adjuvant chemotherapy

ABSTRACT

Background: Symptomatic breast cancers share aggressive clinico-pathological characteristics compared to screen-detected breast cancers. We assessed the association between the method of cancer detection and genomic and clinical risk, and its effect on adjuvant chemotherapy recommendations.

Patients and methods: Patients with early hormone receptor positive (HR+) HER2neu-negative (HER2-) breast cancer, and known OncotypeDX Breast Recurrence Score test were included. A natural language processing (NLP) algorithm was used to identify the method of cancer detection. The clinical and genomic risks of symptomatic and screen-detected tumors were compared.

Results: The NLP algorithm identified the method of detection of 401 patients, with 216 (54%) diagnosed by routine screening, and the remainder secondary to symptoms. The distribution of OncotypeDX recurrence score (RS) varied between the groups. In the symptomatic group there were lower proportions of low RS (13% vs 23%) and higher proportions of high RS (24% vs. 13%) compared to the screen-detected group. Symptomatic tumors were significantly more likely to have a high clinical risk (59% vs 40%). Based on genomic and clinical risk and current guidelines, we found that women aged 50 and under, with a symptomatic cancer, had an increased probability of receiving adjuvant chemotherapy recommendation compared to women with screen-detected cancers (60% vs. 37%).

Conclusions: We demonstrated an association between the method of cancer detection and both genomic and clinical risk. Symptomatic breast cancer, especially in young women, remains a poor prognostic factor that should be taken into account when evaluating patient prognosis and determining adjuvant treatment plans.

© 2021 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

* Corresponding author. Oncology Division, Tel Aviv Sourasky Medical Center, 6 Weizman Street, Tel Aviv, 64239, Israel.

** Corresponding author. Department of Surgery, Chaim Sheba Medical Center, Tel-Hashomer, 52621, Israel.

E-mail addresses: tehillah.menes@sheba.health.gov.il (T. Menes), amirson@tlvmc.gov.il (A. Sonnenblick).

¹ Contributed equally.

1. Introduction

Breast cancer is detected by screening mammography, self-detection and, rarely, by clinical breast exam. At the time of diagnosis, 90% of cases are found to be early breast cancer, in which the disease is confined to the breast and regional lymph nodes; these patients are treated with curative intent [1].

Following the implementation of screening mammography, most breast cancers in the United States are diagnosed by screening

mammography, while approximately one-third are diagnosed as a symptomatic tumor [2–4]. Detection methods vary with age. Young women (<50 years) are more likely to present with palpable tumors, whereas identification of cancer by screening mammography increases with age [2,3].

Several studies have demonstrated that symptomatic breast cancers share more aggressive clinico-pathological characteristics including larger size, higher proportion of node positive disease, higher grade and hormone receptor negative subtypes, compared to screen-detected breast cancers [5–12]. Accordingly, patients with screen-detected breast cancers have improved survival rates compared to those with symptomatic cancers, specifically among those diagnosed with Luminal A subtype [5,6,9,11].

The decision to administer adjuvant chemotherapy in addition to the standard preventive endocrine therapy for hormone receptor positive (HR+) HER2neu-negative (HER2-) breast cancer patients has evolved and currently is often guided by the combination of clinical parameters and genomic tests [13,14]. The OncotypeDX Breast Recurrence Score test was found to be both prognostic for disease recurrence [15,16] and predictive for adjuvant chemotherapy benefit in node negative patients [17–20] and in postmenopausal women with 1–3 positive nodes [21–24]. OncotypeDX did not demonstrate predictivity for chemotherapy benefit in premenopausal node positive patients [24].

In the current study using data on recurrence scores of consecutive women with early HR+ HER2- breast cancer, we assessed the association between the method of cancer detection and the genomic and clinical risk. We implemented a natural language processing (NLP) algorithm to extract the method of tumor detection from the electronic medical record (EMR) and evaluated the contribution of the cancer detection method to adjuvant chemotherapy recommendations, based on the most contemporary treatment guidelines.

2. Methods

2.1. Patients and data retrieval

All patients with known OncotypeDX recurrence score (RS), diagnosed with HR positive HER2 negative early breast cancer between 2004 and 2020 at the Tel Aviv Sourasky Medical Center (TASMC), were included. Patients with HER2 positive disease (HER2 +3 or HER2 +2 with positive HER2 FISH) were excluded. OncotypeDX test was completed to guide chemotherapy recommendation for women with tumors larger than 1 cm. Patients with node negative disease comprised the majority of the cases. Node positive patients for whom the institutional tumor board believed chemotherapy could potentially be omitted based on the test results were included as well. The test was conducted in the majority upon surgical specimens (90%) and only a minority on biopsy specimens. Women with clear indications for neoadjuvant therapy were not referred for OncotypeDX testing. We retrospectively retrieved pathological characteristics including age, tumor size, grade, Ki67, progesterone receptor (PR) status, HER2 level and nodal status. Luminal subtype was defined with pathological based surrogate definitions (ESMO criteria); Luminal A-like subtype was defined if PR>20% and Ki67 ≤ 14%. Luminal B-like subtype was defined if PR<20% and/or Ki67 > 14% [25]. The study was approved by the local Institutional Ethics Committee, Number TLV18-0426.

We analyzed all free-text patient visit summaries from the breast-oncology unit and the breast-cancer surgery unit. We developed a rule-based NLP information-extraction algorithm to identify the initial method of tumor detection, analyzing free-text medical reports, written in Hebrew. Our algorithm was designed to search for terms indicating the method of tumor detection. A

symptomatic tumor was defined if prior to the diagnostic process there was a palpable tumor, breast pain, or changes in the morphological appearance of the breast noted by the patient or her physician; a screen-detected tumor was defined if the tumor was detected during a routine screening mammography or ultrasound exam. A full description of the algorithm and the validation process are provided in [Appendix A](#).

2.2. Assessment of the genomic and clinical risk

Genomic risk: Genomic risk was defined according to the TAILORx study [19], OncotypeDX RS ≤ 10 was considered low risk, 10 < RS ≤ 25 was considered intermediate risk and RS ≥ 26 was considered high risk.

Clinical risk: The clinical risk assessment was based on the Adjuvant! Online algorithm (version 8) integrating tumor size, grade and nodal status [26,27]. Since Adjuvant! is no longer available online, we used a binary clinical-risk categorization (low vs. high) model based on the algorithm, as applied in the MINDACT trial (See appendix table S13 in Ref. [28]). A low clinical risk was defined as greater than 92% probability of breast cancer-specific survival at 10 years in women with HR+ HER2- tumors who received endocrine therapy alone [28]. For N0/N1mic patients clinical risk was defined as low if one of the following conditions was present: grade I and tumor size ≤3 cm, or grade II and tumor size ≤2 cm, or grade III and tumor size ≤1 cm. For N1 patients clinical risk was defined as low only if it was Grade I and the tumor size was ≤2 cm. Otherwise, the clinical risk was defined as high.

2.3. Assessment of the probability for adjuvant chemotherapy recommendation

The probability for adjuvant chemotherapy recommendation was calculated based on the genomic and clinical risk using the model suggested by the phase III TAILORx study [19], the subsequent analysis by Sparano et al. [29] and the recently published RxPonder results [24].

In the TAILORx trial, node-negative patients with RS ≤ 25 did not benefit, while node-negative patients with high genomic risk (RS ≥ 26) did benefit from adjuvant chemotherapy [18–20]. However, an exploratory analysis revealed that younger women under the age of 50 may benefit from adjuvant chemotherapy, even with lower genomic risk (RS 16–25) [19]. A secondary analysis of the TAILORx trial demonstrated that in node-negative younger women the benefit from adjuvant chemotherapy is defined by both the genomic risk and the clinical risk [29]. According to this model, adjuvant chemotherapy should be considered in node-negative women of all ages with a high RS (≥26). Additionally, chemotherapy should be considered in younger women (Age ≤50) with a RS 16 or higher, based on their clinical risk; In high clinical risk tumors, chemotherapy should be considered with a RS ≥ 16 and in low clinical risk with a RS ≥ 21.

The RxPonder trial demonstrated that postmenopausal node-positive (N1) patients with RS ≤ 25 did not benefit from adjuvant chemotherapy. However, premenopausal node-positive patients (N1) with RS ≤ 25 did benefit from adjuvant chemotherapy regardless of OncotypeDX RS [24]. According to those results, node-positive postmenopausal patients should be recommended for adjuvant chemotherapy only with a high RS (≥26). Chemotherapy should be advised to all node-positive premenopausal patients regardless of RS. Patients with more than 3 positive lymph nodes (N2–N3) were excluded from this analysis as there is no evidence that chemotherapy can be omitted in this population.

2.4. Statistical analysis

Demographic and clinico-pathological data of patients in the two groups (symptomatic and screen-detected) were compared, using the chi-square test for categorical variables and the *t*-test for continuous variables. The genomic risk and the clinical risk as well as the probability of adjuvant chemotherapy was compared between the screen-detected and the symptomatic groups using the chi-square test. All *p* values were two-sided and *p* < .05 was considered significant. Statistical analysis was performed with IBM SPSS statistics for Windows, version 25 (IBM Corp., Armonk, N.Y., USA).

3. Results

3.1. Patient characteristics

The cohort included 962 consecutive patients with known OncotypeDX scores and available EMRs, who were diagnosed between 2004 and 2020. The NLP algorithm successfully extracted the method of cancer detection in 401 patients. For the remainder of the cohort access to initial diagnostic data was not available as it was stored in a different software system. Most of the women (N = 216; 53.9%) were detected by routine screening, and 185 (46.1%) patients by self-examination or symptoms (Fig. 1). Patient characteristics are summarized in Table 1. Symptomatic women were younger (mean age 53 vs. 61, *p* < .001), had larger tumors (46% were ≥ 2 cm compared to 18% of screen-detected tumors; *p* < .0001) and high Ki67 (24% were Ki67 > 14% compared to 15% of screen-detected tumors; *p* < .004). A quarter (N = 104) of the cohort presented with node positive disease, with no significant differences in the proportions of node positive patients between the two groups. A higher proportion of patients for whom the biopsy specimen was used for genomic testing had symptomatic tumors (16% compared to 6% of screen-detected tumors; *p* < .001). Approximately half of the cohort had known Ki67 and PR status which enabled us to define the luminal subtype based on ESMO criteria. No significant differences were found in the proportions of Luminal A-like and Luminal B-like tumors (50% were Luminal B-like in the screen-detected group compared to 59% in the symptomatic group; *p* = .2).

3.2. Impact of initial method of breast cancer detection on the genomic risk of recurrence

The distribution of OncotypeDX RS was significantly different

between the two groups (Fig. 2A, *p* = .003). The proportion of patients with low RS (0-10) was higher in the screen-detected group compared to the symptomatic group (23% vs. 13%). Conversely, the proportion of patients with high RS (≥26) was higher in the symptomatic group (24% vs. 13% in the screen-detected group).

The association between the method of cancer detection and the RS was even more pronounced in women 50 years or younger (Fig. 2B; *p* = .02). A smaller proportion of women with low or intermediate RS and a higher proportion of women with high RS were identified in the symptomatic group (11% vs. 21%, 56% vs. 72% and 33% vs. 7%, respectively). In women older than 50 there was a higher proportion of low RS in the screen detected group, but the proportions of high RS were similar in both groups (Fig. 2C; *p* = .28).

3.3. Impact of initial method of breast cancer detection on the clinical risk of recurrence

The clinical risk as assessed using tumor size, grade and nodal status was higher in the symptomatic group; Fifty-nine percent of patients who presented with symptomatic cancer had high clinical risk of recurrence compared to only 40% with screen-detected tumors (Fig. 3A, *p* = .0001). Similar trends were seen when assessing separately the different age groups (Fig. 3B, Age ≤50; *p* = .04 and Fig. 3C, Age >50; *p* = .002).

3.4. Probability of adjuvant chemotherapy recommendation based on the initial method of tumor detection

When applying the model based on the TAILORx analysis [29] and the RxPonder results [24] to our data (Table 2), in women over 50, the probability for adjuvant chemotherapy recommendation is comparable in the symptomatic and screen-detected groups (15% vs. 13%, respectively; *p* = .7). In women who were 50 years or younger, the probability of adjuvant chemotherapy recommendation was significantly higher in women presenting with symptomatic tumors compared to women with screen-detected tumors (60% vs. 37%, respectively. *p* = .03).

4. Discussion

We examined the association between the method of cancer detection and the genomic and clinical risk of disease recurrence in women with early HR+ HER2- breast cancer. Women with symptomatic tumors had both higher clinical risk and higher genomic risk for disease recurrence, compared to patients whose tumors were detected by routine screening. These findings are consistent

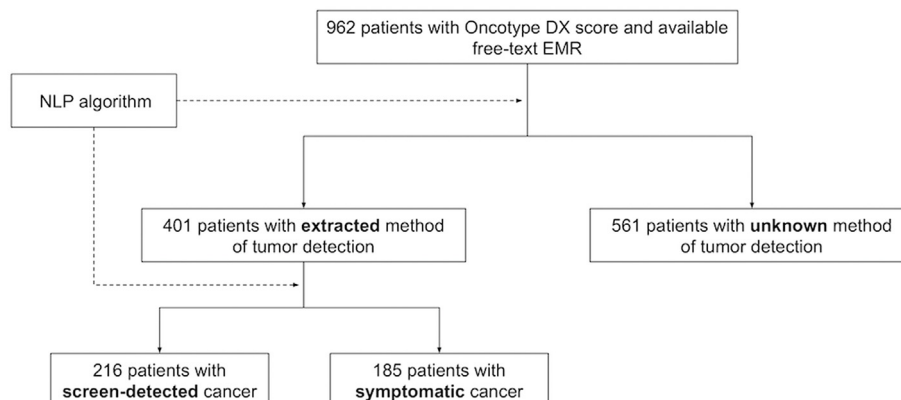


Fig. 1. Study flow diagram and implantation of NLP algorithm. Abbreviations: NLP, Natural language processing; EMR, Electronic medical records.

Table 1
Patient and tumor characteristics by method of detection.

Characteristics	Screen-detected (n = 216)	n (%)	symptomatic (n = 185)	p value
Age (mean)	61		53	<.00001
≤50	29 (13.4)		91 (49.1)	<.00001
>50	187 (86.5)		94 (50.8)	
Tumor size				
≤2 cm	178 (82.4)		99 (53.5)	<.00001
>2 cm	38 (17.5)		86 (46.4)	
Histologic grade				
G1: Well differentiated	7 (3.2)		9 (4.8)	.14
G2: Moderately differentiated	163 (75.4)		122 (65.9)	
G3: Poorly differentiated/Undiff	43 (19.9)		49 (26.4)	
Missing data	3 (1.4)		5 (2.7)	
Lymph Node Status				
N0	161 (74.6)		136 (73.5)	.78
N1mic	11 (5.1)		14 (7.6)	
N1	30 (13.9)		24 (13)	
N2–N3	14 (6.5)		11 (6)	
Ki67				
Ki67 ≤ 14%	79 (36.6)		42 (22.7)	<.004
Ki67 > 14%	33 (15.3)		45 (24.3)	
Missing data	104 (48.1)		98 (53)	
Progesterone Receptor (PR)				
Positive	151 (70)		144 (77.9)	.1
Negative	62 (28.7)		39 (21.1)	
Missing data	3 (1.4)		2 (1.1)	
Tissue for Genomic testing				
Surgical specimen	203 (94)		156 (84.3)	<.001
Biopsy	13 (6)		29 (15.7)	

with previous reports which have demonstrated that symptomatic tumors portend more aggressive clinico-pathological characteristics than screen-detected tumors [5–12]. Accordingly, the method of detection was found to be prognostic for disease survival [5,6] and therefore was incorporated in the PREDICT online prognostication tool together with clinical and tumor characteristics [30,31].

In accordance with previous literature [2,3,8,10,11], our study demonstrates that patients with symptomatic breast cancer tend to be significantly younger (49% of women with symptomatic cancer were ≤50 years compared to 13% of screen-detected patients). This observation is attributed, at least in part, to the widespread adoption of screening mammography in women over 50. The Israeli breast screening program invites average-risk women aged 50–74 to undergo screening mammography every two years. The screening compliance in this group is 75% [32]. Younger age at diagnosis is associated with more aggressive tumor behavior [33] and may explain the greater prevalence of high genomic and clinical risk tumors in symptomatic patients. However, in a subset analysis of women under the age of 50, the association between the method of detection on the genomic and clinical risk remained significant. Therefore, the higher prevalence of genomic and clinical risk tumors in symptomatic patients cannot be explained by age alone.

There are few reports examining the association between method of detection and genomic risk. Esserman et al. [34] compared the 70-gene signature MammaPrint in two groups of women in the Netherlands. The first group included women diagnosed between 1984 and 1992, before the era of widespread use of screening mammography. The second group included women diagnosed between 2004 and 2006 when image-based screening reached 75–80% of the population. Similarly, Drukker et al. [35] analyzed 1165 patients in the MINDACT trial and compared the 70-gene signature of cancers detected by image-based screening to interval cancers. In accordance with our findings, both reports suggest that screen-detected cancers are more likely to be of low genomic risk. Additionally, in the predominantly-screened group

almost a third had an ultra-low genomic risk [34] leading the authors to suggest that these tumors may account for clinical over-diagnosis. Conversely, in our cohort, in the group of women presenting with symptomatic breast cancer 13% were found to have a low RS (≤10), suggesting that even ultra-low risk tumors can become symptomatic.

Adjuvant treatment recommendations have evolved tremendously over the past two decades. Presently the decision to add adjuvant chemotherapy to HR + early breast cancer patients is determined by clinical factors in addition to genomic features derived from molecular tests. We examined the contribution of the method of cancer detection to adjuvant chemotherapy recommendation based on the model suggested by the TAILORx results [19], the subsequent analysis by Sparano et al. [29] and the related results of the RxPonder [24]. When applying this model to our results, we observed that in women age 50 and under, symptomatic cancer significantly increases the likelihood of adjuvant chemotherapy recommendation compared to screen-detected tumors (60% vs. 37%, respectively). These results demonstrate the relevance of the tumor detection-method, underlining the fact that symptomatic tumor at presentation is prognostic, especially in young women.

The TAILORx trial demonstrated the benefit of adjuvant chemotherapy in young women with a RS > 16. However only a minority (13%) of the premenopausal women who participated in the study received ovarian suppression in addition to endocrine therapy [19]. In light of the clear benefit observed with ovarian suppression in addition to endocrine therapy in the SOFT and TEXT studies [36], it is unclear if the benefit of chemotherapy in preventing disease recurrence among premenopausal women, can be attributed at least in part to chemotherapy induced ovarian failure. Our work does not address this question and it is plausible that some patients who were recommended for chemotherapy, would have similarly benefited from ovarian suppression alone.

Approximately 25% of the women in the study had involved lymph nodes, the majority with 1–3 positive nodes (N1). Women

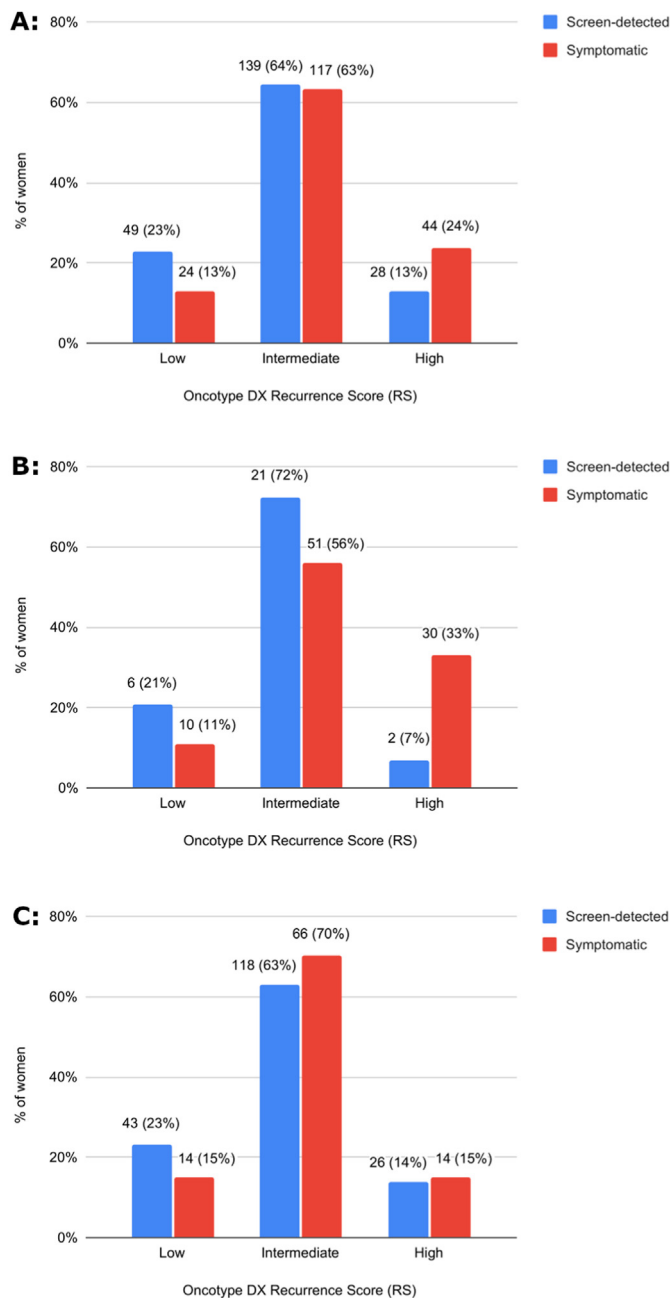


Fig. 2. Genomic risk distribution by the initial method of cancer detection. **(A) Among all women** (N = 401; $p = .003$) **(B) Age ≤ 50** (N = 120; $p = .02$) **(C) Age > 50** (N = 281; $p = .28$). Variables are shown as n (%). The p values are based on chi-squared tests.

with lymph node involvement were evenly distributed between the two study groups. A number of studies have suggested that OncotypeDX is prognostic also in women with positive lymph nodes [21–23]. The recently published results of the prospective RxPonder trial demonstrated that chemotherapy can be spared in postmenopausal women with 1–3 involved nodes and a RS ≤ 25 [24]. Most contemporary clinical guidelines have integrated OncotypeDX in the treatment algorithm of node positive (N1) HR+ HER2- early breast cancer patients [37]. Accordingly, we included node-positive (N1) patients in our analysis and assessed the probability of chemotherapy recommendation among these patients based on RS and menopausal status, in line with the RxPONDER results.

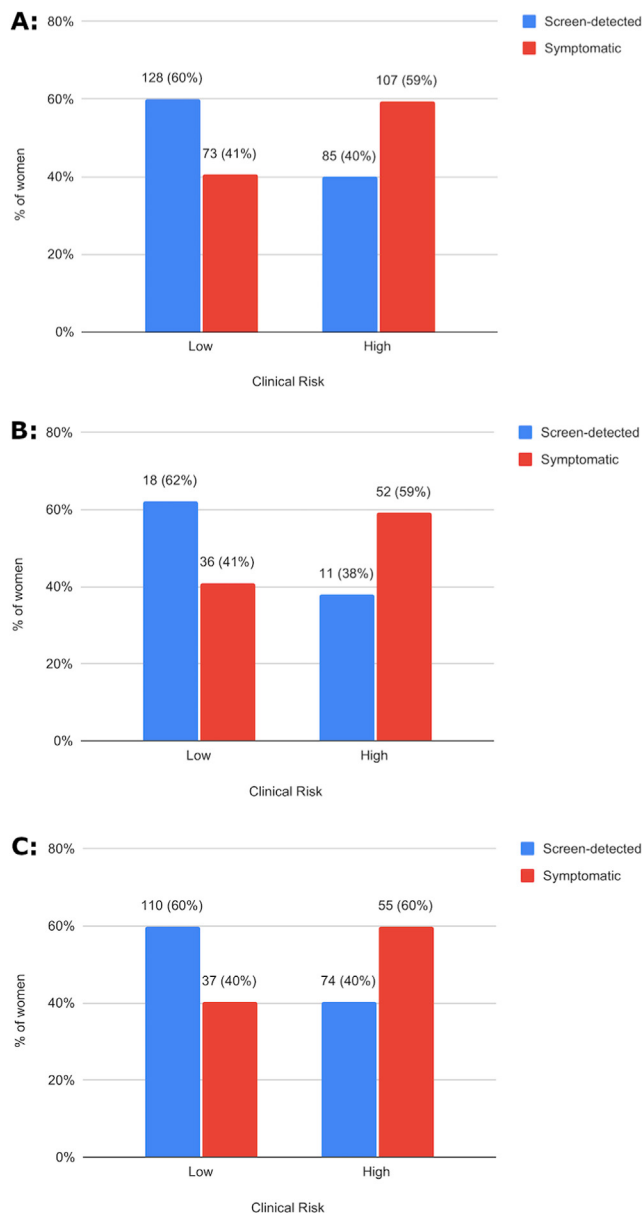


Fig. 3. Clinical risk distribution by the initial method of cancer detection. **(A) Among all women** (N = 393; $p = .0001$) **(B) Age ≤ 50** (N = 117; $p = .04$) **(C) Age > 50** (N = 276; $p = .002$). Variables are shown as n (%). The p values are based on chi-squared tests.

In this work we used NLP algorithms to extract the method of tumor detection from free-text visit summaries. There is a growing body of literature in which computational approaches are applied for processing unstructured records to retrieve information and improve diagnosis performance [38,39], support treatment decisions [40], and improve cancer research by providing a better interface to existing knowledge platforms [41]. We believe that in our work, we have demonstrated the potential of integrating a computational approach for extracting information from oncological EMRs that outline the disease course of a patient, formatted in a completely unstructured way.

Our work has several limitations. The main limitation is the selection of women for genomic studies. Over time the recommendations for genomic tests have changed. Moreover, one can assume that in the screen-detected group many low-risk women were not selected for a genomic test, whereas in the symptomatic

Table 2
Probability of adjuvant chemotherapy recommendation by the initial method of cancer detection. Colored boxes represent patients who are candidates for adjuvant chemotherapy. Variables are shown as n (%). Patients with more than 3 positive nodes (N2–N3) were excluded from this analysis. The p values are based on chi-squared tests. Abbreviations: RS, OncotypeDX Recurrence Score.

Age ≤ 50 (n=113)					
	Symptomatic (n=86)		Screen-detected (n=27)		p value
	Low Clinical Risk	High Clinical Risk	Low Clinical Risk	High Clinical Risk	
RS<16 (N0)	13 (15.1)	15 (17.4)	12 (44.4)	2 (7.4)	
RS<16 (N1)	0 (0)	4 (4.6)	0 (0)	1 (3.7)	
16<= RS<=20	6 (7)	8 (9.3)	3 (11.1)	3 (11.1)	
20<RS<=25	8 (9.3)	6 (7)	2 (7.4)	2 (7.4)	
RS>25	9 (10.4)	17 (19.7)	1 (3.7)	1 (3.7)	
Chemotherapy Recommendation	52 (60)		10 (37)		<i>p</i> =.03
Age > 50 (n=255)					
	Symptomatic (n=83)		Screen-detected (n=172)		p value
	Low Clinical Risk	High Clinical Risk	Low Clinical Risk	High Clinical Risk	
RS<=25	34 (41)	37 (44.5)	102 (59.3)	48 (27.9)	
RS>25	3 (3.6)	9 (14.4)	8 (4.6)	14 (8.1)	
Chemotherapy Recommendation	12 (15)		22 (13)		<i>p</i> =.7

group many high risk women were recommended for neoadjuvant or adjuvant chemotherapy without undergoing a genomic test. However, such a differential selection bias would be expected to weaken the association we found between method of detection and RS. Additionally, the method of breast cancer detection was defined partially by self-reported data which can be influenced by recall bias and lead to misclassification. The accuracy of the NLP algorithm was 91%, allowing for misclassification of several cases as well. The relatively small final sample size limited our ability to perform multivariable analysis and examine the independent contribution of the method of detection to the recurrence score. Finally, our study does not include long term follow-up which limits our ability to examine the independent contribution of the method of diagnosis to long-term survival.

5. Conclusion

We demonstrated a strong association between the method of cancer detection and the genomic and clinical risk of recurrence in

HR+ early breast cancer patients. Based on our data and current guidelines, most young women presenting with symptomatic breast cancer will be recommended for adjuvant chemotherapy.

Declaration of competing interest

A.Sonnenblick reports personal fees from Eli Lilly, Pfizer, Teva, Novartis, Medison and Roche; grants from Novartis and Roche, all outside the submitted work. I.Wolf reports research grants from MSD, BMS, Roche and Novartis, all outside the submitted work. The rest of the authors declare that they have no conflict of interest.

Acknowledgments

We thank Oncotest for providing OncotypeDX Breast Recurrence Score test results.

Appendix A

Automatic extraction of the method of tumor detection using a rule-based NLP algorithm

The computational process was executed in three steps:

1. Visit summary chronology: the algorithm organized visit summaries in chronological order for each patient
2. Expression detection: We compiled a list of terms and phrases capturing events of tumor-detection methods. The list included variations of each expression in an attempt to reflect and include different writing styles, synonyms, paraphrases, common misspellings and inflections. Overall, we applied 35 expressions indicating the symptomatic method, and 19 expressions designating the screen-detection method. Hebrew equivalent expressions of “palpable tumor” and “felt pain in the (right/left) breast” are examples of the phrases identifying the symptomatic method, while Hebrew terms stating “routine mammography” and “screening mammography” are examples of phrases identifying screen-detection method.

The search for these expressions was applied to the earliest visit summaries for each patient, in a chronological order. Once the appropriate expression was identified and validated (see Validation section), the algorithm returned the relevant detection method. If none of the expressions were identified in any of the patient’s visit summaries, the algorithm halted and returned ‘unknown’, reflecting its inability to identify the method of tumor detection of the patient. The summaries were written in Hebrew, a highly inflected language; Hebrew words are derived from a root and a pattern, combined with prefixes and suffixes, which may interfere with the traditional way of searching text. Therefore, for each expression we considered all the relevant inflections possible in the text.

3. Validation: To eliminate false positives, some of the expressions required extra validation steps. For example, some symptomatic expressions needed validation to confirm that they were not mentioned in negation (e.g., “did not palpate a tumor”). Therefore, we created a few simple negation detection rules. Additional confirmation involved validation that mention of a routine screening mammography did in fact result in detection of a tumor.

In addition, our algorithm was validated by a breast medical oncologist. Collectively, our human evaluation set contains 101 cases (10.4% of total cases): 29 cases of the screen-detection method, 21 cases of the symptomatic method, and 51 cases of unknown detection method. The overall accuracy was 91%. The sensitivity and specificity for screen-detection were 96.5% and 95.8%, respectively, and the sensitivity and specificity for symptomatic cancers were 85.7% and 95%, respectively.

Funding

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

Data sharing

The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request

Ethical approval

The study was approved by the local Institutional Ethics

Committee and was performed in line with the principles of the Declaration of Helsinki.

Consent to participate

As data were aggregative and anonymous no informed consent was required by the institutional Committee.

References

- [1] Waks AG, Winer EP. Breast cancer treatment: a review. *J Am Med Assoc* 2019;321:288. <https://doi.org/10.1001/jama.2018.19323>.
- [2] Roth MY, Elmore JG, Yi-Frazier JP, Reisch LM, Oster NV, Miglioretto DL. Self-detection remains a key method of breast cancer detection for U.S. Women. *J Wom Health* 2011;20:1135–9. <https://doi.org/10.1089/jwh.2010.2493>.
- [3] Caughran J, Braun TM, Breslin TM, Smith DR, Kreinbrink JL, Parish GK, Davis AT, Bacon-Baguley TA, Silver SM, Henry NL. The Effect of the 2009 USPSTF breast cancer screening recommendations on breast cancer in Michigan: a longitudinal study. *Breast J* 2018;24:730–7. <https://doi.org/10.1111/tbj.13034>.
- [4] Bleyer A, Welch HG. Effect of three decades of screening mammography on breast-cancer incidence. *N Engl J Med* 2012;367:1998–2005. <https://doi.org/10.1056/NEJMoa1206809>.
- [5] Shen Y, Yang Y, Inoue LYT, Munsell MF, Miller AB, Berry DA. Role of detection method in predicting breast cancer survival: analysis of randomized screening trials. *JNCI: J Nat Canc Inst* 2005;97:1195–203. <https://doi.org/10.1093/jnci/dji239>.
- [6] Dawson SJ, Duffy SW, Blows FM, Driver KE, Provenzano E, LeQuesne J, Greenberg DC, Pharoah P, Caldas C, Wishart GC. Molecular characteristics of screen-detected vs symptomatic breast cancers and their impact on survival. *Br J Canc* 2009;101:1338–44. <https://doi.org/10.1038/sj.bjc.6605317>.
- [7] Mook S, Van 't Veer LJ, Rutgers EJ, Ravdin PM, van de Velde AO, van Leeuwen FE, Visser O, Schmidt MK. Independent prognostic value of screen detection in invasive breast cancer. *JNCI: J Nat Canc Inst* 2011;103:585–97. <https://doi.org/10.1093/jnci/djr043>.
- [8] Kim J, Lee S, Bae S, Choi M-Y, Lee J, Jung SP, Kim S, Choe J-H, Kim J-H, Kim JS, Lee JE, Nam SJ, Yang J-H. Comparison between screen-detected and symptomatic breast cancers according to molecular subtypes. *Breast Canc Res Treat* 2012;131:527–40. <https://doi.org/10.1007/s10549-011-1836-0>.
- [9] Redondo M, Funez R, Medina-Cano F, Rodrigo I, Acebal M, Tellez T, Roldan MJ, Hortas ML, Bellinvia A, Pereda T, Domingo L, Morales-Suarez Varela M, Sala M, Rueda A. Detection methods predict differences in biology and survival in breast cancer patients. *BMC Canc* 2012;12:604. <https://doi.org/10.1186/1471-2407-12-604>.
- [10] Hayse B, Hooley RJ, Killelea BK, Horowitz NR, Chagpar AB, Lannin DR. Breast cancer biology varies by method of detection and may contribute to overdiagnosis. *Surgery* 2016;160:454–62. <https://doi.org/10.1016/j.surg.2016.03.031>.
- [11] Crispo A, Barba M, D'Aiuto G, De Laurentis M, Grimaldi M, Rinaldo M, Caolo G, D'Aiuto M, Capasso I, Esposito E, Amore A, Di Bonito M, Botti G, Montella M. Molecular profiles of screen detected vs. symptomatic breast cancer and their impact on survival: results from a clinical series. *BMC Canc* 2013;13:15. <https://doi.org/10.1186/1471-2407-13-15>.
- [12] Molino A, Pavarana M, Micciolo R, Nortilli R, Pedersini R, Manno P, Bozzi P, Bonetti F, Piubello Q, Cetto LG. Comparative study of clinical, pathological and biological characteristics of symptomatic versus asymptomatic breast cancers. *Ann Oncol* 2000;11:581–6. <https://doi.org/10.1023/A:1008320317114>.
- [13] Harris LN, Ismaila N, McShane LM, Andre F, Collyar DE, Gonzalez-Angulo AM, Hammond EH, Kuderer NM, Liu MC, Mennel RG, Van Poznak C, Bast RC, Hayes DF. Use of biomarkers to guide decisions on adjuvant systemic therapy for women with early-stage invasive breast cancer: American society of clinical oncology clinical practice guideline. *J Clin Oncol* 2016;34:1134–50. <https://doi.org/10.1200/JCO.2015.65.2289>.
- [14] Krop I, Ismaila N, Andre F, Bast RC, Barlow W, Collyar DE, Hammond ME, Kuderer NM, Liu MC, Mennel RG, Van Poznak C, Wolff AC, Stearns V. Use of biomarkers to guide decisions on adjuvant systemic therapy for women with early-stage invasive breast cancer: American society of clinical oncology clinical practice guideline focused update. *J Clin Oncol* 2017;35:2838–47. <https://doi.org/10.1200/JCO.2017.74.0472>.
- [15] Paik S, Shak S, Tang G, Kim C, Baker J, Cronin M, Baehner FL, Walker MG, Watson D, Park T, Hiller W, Fisher ER, Wickerham DL, Bryant J, Wolmark N. A multigene assay to predict recurrence of tamoxifen-treated, node-negative breast cancer. *N Engl J Med* 2004;351:2817–26. <https://doi.org/10.1056/NEJMoa041588>.
- [16] Wolf I, Ben-Baruch N, Shapira-Frommer R, Rizel S, Goldberg H, Yaal-Hahoshen N, Klein B, Geffen DB, Kaufman B. Association between standard clinical and pathologic characteristics and the 21-gene recurrence score in breast cancer patients: a population-based study. *Cancer* 2008;112:731–6. <https://doi.org/10.1002/cncr.23225>.
- [17] Paik S, Tang G, Shak S, Kim C, Baker J, Kim W, Cronin M, Baehner FL, Watson D, Bryant J, Costantino JP, Geyer CE, Wickerham DL, Wolmark N. Gene expression and benefit of chemotherapy in women with node-negative, estrogen

- receptor-positive breast cancer. *J Clin Orthod* 2006;24:3726–34. <https://doi.org/10.1200/JCO.2005.04.7985>.
- [18] Sparano JA, Gray RJ, Makower DF, Pritchard KI, Albain KS, Hayes DF, Geyer CE, Dees EC, Perez EA, Olson JA, Zujewski J, Lively T, Badve SS, Saphner TJ, Wagner LJ, Whelan TJ, Ellis MJ, Paik S, Wood WC, Ravdin P, Keane MM, Gomez Moreno HL, Reddy PS, Goggins TF, Mayer IA, Brufsky AM, Toppmeyer DL, Kaklamani VG, Atkins JN, Berenberg JL, Sledge GW. Prospective validation of a 21-gene expression assay in breast cancer. *N Engl J Med* 2015;373:2005–14. <https://doi.org/10.1056/NEJMoa1510764>.
- [19] Sparano JA, Gray RJ, Makower DF, Pritchard KI, Albain KS, Hayes DF, Geyer CE, Dees EC, Goetz MP, Olson JA, Lively T, Badve SS, Saphner TJ, Wagner LJ, Whelan TJ, Ellis MJ, Paik S, Wood WC, Ravdin PM, Keane MM, Gomez Moreno HL, Reddy PS, Goggins TF, Mayer IA, Brufsky AM, Toppmeyer DL, Kaklamani VG, Berenberg JL, Abrams J, Sledge GW. Adjuvant chemotherapy guided by a 21-gene expression assay in breast cancer. *N Engl J Med* 2018;379:111–21. <https://doi.org/10.1056/NEJMoa1804710>.
- [20] Sparano JA, Gray RJ, Makower DF, Albain KS, Saphner TJ, Badve SS, Wagner LJ, Kaklamani VG, Keane MM, Gomez HL, Reddy PS, Goggins TF, Mayer IA, Toppmeyer DL, Brufsky AM, Goetz MP, Berenberg JL, Mahalcioiu C, Desbiens C, Hayes DF, Dees EC, Geyer CE, Olson JA, Wood WC, Lively T, Paik S, Ellis MJ, Abrams J, Sledge GW. Clinical outcomes in early breast cancer with a high 21-gene recurrence score of 26 to 100 assigned to adjuvant chemotherapy plus endocrine therapy: a secondary analysis of the TAILORx randomized clinical trial. *JAMA Oncol* 2020;6:367. <https://doi.org/10.1001/jamaoncol.2019.4794>.
- [21] Gluz O, Nitz UA, Christgen M, Kates RE, Shak S, Clemens M, Kraemer S, Aktas B, Kuemmel S, Reimer T, Kusche M, Heyl V, Lorenz-Salehi F, Just M, Hofmann D, Degenhardt T, Liedtke C, Svedman C, Wuerstlein R, Kreipe HH, Harbeck N. West German study group phase III PlanB trial: first prospective outcome data for the 21-gene recurrence score assay and concordance of prognostic markers by central and local pathology assessment. *J Clin Orthod* 2016;34:2341–9. <https://doi.org/10.1200/JCO.2015.63.5383>.
- [22] Albain KS, Barlow WE, Shak S, Hortobagyi GN, Livingston RB, Yeh I-T, Ravdin P, Bugarini R, Baehner FL, Davidson NE, Sledge GW, Winer EP, Hudis C, Ingle JN, Perez EA, Pritchard KI, Shepherd L, Gralow JR, Yoshizawa C, Allred DC, Osborne CK, Hayes DF. Prognostic and predictive value of the 21-gene recurrence score assay in postmenopausal women with node-positive, oestrogen-receptor-positive breast cancer on chemotherapy: a retrospective analysis of a randomised trial. *Lancet Oncol* 2010;11:55–65. [https://doi.org/10.1016/S1470-2045\(09\)70314-6](https://doi.org/10.1016/S1470-2045(09)70314-6).
- [23] Stemmer SM, Steiner M, Rizel S, Geffen DB, Nisenbaum B, Peretz T, Soussan-Gutman L, Bareket-Samish A, Isaacs K, Rosengarten O, Fried G, McCullough D, Svedman C, Shak S, Liebermann N, Ben-Baruch N. Clinical outcomes in ER+HER2- node-positive breast cancer patients who were treated according to the Recurrence Score results: evidence from a large prospectively designed registry. *npj Breast Canc* 2017;3:32. <https://doi.org/10.1038/s41523-017-0033-7>.
- [24] Kalinsky K, Barlow WE, Meric-Bernstam F, Gralow JR, Albain KS, Hayes D, Lin N, Perez EA, Goldstein LJ, Chia S, Dhesy-Thind S, Rastogi P, Alba E, Delalogo S, Martin M, Gil MG, Arce-Salinas C, Brain E, Park IH, Pierga J-Y, Lluch AH, Vasquez MR, Borrego MR, Jung KH, Ferrero J-M, Schott A, Shak S, Sharma P, Lew DL, Miao J, Tripathy D, Hortobagyi G, Pusztai L. Abstract GS3-00: first results from a phase III randomized clinical trial of standard adjuvant endocrine therapy (ET) +/- chemotherapy (CT) in patients (pts) with 1-3 positive nodes, hormone receptor-positive (HR+) and HER2-negative (HER2-) breast cancer (BC) with recurrence score (RS) < 25: SWOG S1007 (RxPonder). *Canc Res* 2021;81. <https://doi.org/10.1158/1538-7445.SABCS20-GS3-00>. GS3-00.
- [25] Goldhirsch A, Winer EP, Coates AS, Gelber RD, Piccart-Gebhart M, Thürlimann B, Senn H-J, Albain KS, André F, Bergh J, Bonnefoi H, Bretel-Morales D, Burstein H, Cardoso F, Castiglione-Gertsch M, Coates AS, Colleoni M, Costa A, Curigliano G, Davidson NE, Di Leo A, Ejlersen B, Forbes JF, Gelber RD, Gnant M, Goldhirsch A, Goodwin P, Goss PE, Harris JR, Hayes DF, Hudis CA, Ingle JN, Jassem J, Jiang Z, Karlsson P, Loibl S, Morrow M, Namer M, Kent Osborne C, Partridge AH, Penault-Llorca F, Perou CM, Piccart-Gebhart MJ, Pritchard KI, Rutgers EJ, Sedlmayer F, Semiglazov V, Shao Z-M, Smith I, Thürlimann B, Toi M, Tutt A, Untch M, Viale G, Watanabe T, Wilcken N, Winer EP, Wood WC. Personalizing the treatment of women with early breast cancer: highlights of the st gallen international expert consensus on the primary therapy of early breast cancer 2013. *Ann Oncol* 2013;24:2206–23. <https://doi.org/10.1093/annonc/mdt303>.
- [26] Olivetto IA, Bajdik CD, Ravdin PM, Speers CH, Coldman AJ, Norris BD, Davis GJ, Chia SK, Gelmon KA. Population-based validation of the prognostic model ADJUVANT! For early breast cancer. *J Clin Orthod* 2005;23:2716–25. <https://doi.org/10.1200/JCO.2005.06.178>.
- [27] Mook S, Schmidt MK, Rutgers EJ, van de Velde AO, Visser O, Rutgers SM, Armstrong N, van't Veer LJ, Ravdin PM. Calibration and discriminatory accuracy of prognosis calculation for breast cancer with the online Adjuvant! program: a hospital-based retrospective cohort study. *Lancet Oncol* 2009;10:1070–6. [https://doi.org/10.1016/S1470-2045\(09\)70254-2](https://doi.org/10.1016/S1470-2045(09)70254-2).
- [28] Cardoso F, van't Veer LJ, Bogaerts J, Slaets L, Viale G, Delalogo S, Pierga J-Y, Brain E, Causeret S, DeLorenzi M, Glas AM, Golfopoulos V, Goulioti T, Knox S, Matos E, Meulemans B, Neijenhuis PA, Nitz U, Passalacqua R, Ravdin P, Rubio IT, Saghatian M, Smilde TJ, Sotiriou C, Stork L, Straehle C, Thomas G, Thompson AM, van der Hoeven JM, Vuylsteke P, Bernards R, Tryfonidis K, Rutgers E, Piccart M. 70-Gene signature as an aid to treatment decisions in early-stage breast cancer. *N Engl J Med* 2016;375:717–29. <https://doi.org/10.1056/NEJMoa1602253>.
- [29] Sparano JA, Gray RJ, Ravdin PM, Makower DF, Pritchard KI, Albain KS, Hayes DF, Geyer CE, Dees EC, Goetz MP, Olson JA, Lively T, Badve SS, Saphner TJ, Wagner LJ, Whelan TJ, Ellis MJ, Paik S, Wood WC, Keane MM, Gomez Moreno HL, Reddy PS, Goggins TF, Mayer IA, Brufsky AM, Toppmeyer DL, Kaklamani VG, Berenberg JL, Abrams J, Sledge GW. Clinical and genomic risk to guide the use of adjuvant therapy for breast cancer. *N Engl J Med* 2019;380:2395–405. <https://doi.org/10.1056/NEJMoa1904819>.
- [30] Wishart GC, Azzato EM, Greenberg DC, Rashbass J, Kearins O, Lawrence G, Caldas C, Pharoah PD. PREDICT: a new UK prognostic model that predicts survival following surgery for invasive breast cancer. *Breast Cancer Res* 2010;12:R1. <https://doi.org/10.1186/bcr2464>.
- [31] Predict Breast. <https://breast.predict.nhs.uk/tool>. Accessed 9 Nov 2020.
- [32] Hayek S, Enav T, Shohat T, Keinan-Boker L. Factors associated with breast cancer screening in a country with national health insurance: did we succeed in reducing healthcare disparities? *J Wom Health* 2017;26:159–68. <https://doi.org/10.1089/jwh.2016.5835>.
- [33] Pronzato P, Mustacchi G, De Matteis A, Di Costanzo F, Rulli E, Floriani I, Cazzaniga ME. Biological characteristics and medical treatment of breast cancer in young women—a featured population: results from the NORA study. *Int J Breast Canc* 2011;2011:1–6. <https://doi.org/10.4061/2011/534256>.
- [34] Esserman LJ, Shieh Y, Rutgers EJT, Knauer M, Retèl VP, Mook S, Glas AM, Moore DH, Linn S, van Leeuwen FE, van't Veer LJ. Impact of mammographic screening on the detection of good and poor prognosis breast cancers. *Breast Canc Res Treat* 2011;130:725–34. <https://doi.org/10.1007/s10549-011-1748-z>.
- [35] Drukker CA, Schmidt MK, Rutgers EJT, Cardoso F, Kerlikowske K, Esserman LJ, van Leeuwen FE, Pijnappel RM, Slaets L, Bogaerts J, van't Veer LJ. Mammographic screening detects low-risk tumor biology breast cancers. *Breast Canc Res Treat* 2014;144:103–11. <https://doi.org/10.1007/s10549-013-2830-5>.
- [36] Francis PA, Pagani O, Fleming GF, Walley BA, Colleoni M, Láng I, Gómez HL, Tondini C, Ciruelos E, Burstein HJ, Bonnefoi HR, Bellet M, Martino S, Geyer CE, Goetz MP, Stearns V, Pinotti G, Pujlii F, Spazzapan S, Climent MA, Pavesi L, Ruhstaller T, Davidson NE, Coleman R, Debled M, Buchholz S, Ingle JN, Winer EP, Maibach R, Rabaglio-Poretto M, Ruepp B, Di Leo A, Coates AS, Gelber RD, Goldhirsch A, Regan MM. Tailoring adjuvant endocrine therapy for premenopausal breast cancer. *N Engl J Med* 2018;379:122–37. <https://doi.org/10.1056/NEJMoa1803164>.
- [37] National Comprehensive Cancer Network. NCCN: NCCN clinical practice guidelines in oncology—breast cancer. 2021, Version 5.2021. https://www.nccn.org/professionals/physician_gls/pdf/breast.pdf.
- [38] Kehl KL, Xu W, Lepisto E, Elmarakeby H, Hassett MJ, Van Allen EM, Johnson BE, Schrag D. natural language processing to ascertain cancer outcomes from medical oncologist notes. *JCO Clin Canc Info* 2020;680–90. <https://doi.org/10.1200/CCI.20.00020>.
- [39] Senders JT, Karhade AV, Cote DJ, Mehrtash A, Lamba N, DiRisio A, Muskens IS, Gormley WB, Smith TR, Broekman MLD, Arnaout O. natural language processing for automated quantification of brain metastases reported in free-text radiology reports. *JCO Clin Canc Info* 2019;1–9. <https://doi.org/10.1200/CCI.18.00138>.
- [40] Xu F, Sepúlveda M-J, Jiang Z, Wang H, Li J, Liu Z, Yin Y, Roebuck MC, Shortliffe EH, Yan M, Song Y, Geng C, Tang J, Purcell Jackson G, Preininger AM, Rhee K. Effect of an artificial intelligence clinical decision support system on treatment decisions for complex breast cancer. *JCO Clin Canc Info* 2020;824–38. <https://doi.org/10.1200/CCI.20.00018>.
- [41] Hanauer DA, Barnholtz-Sloan JS, Beno MF, Del Fiol G, Durbin EB, Gologorskaya O, Harris D, Harnett B, Kawamoto K, May B, Meeks E, Pfaff E, Weiss J, Zheng K. Electronic medical record search engine (EMERSE): an information retrieval tool for supporting cancer research. *JCO Clin Canc Info* 2020;454–63. <https://doi.org/10.1200/CCI.19.00134>.