Validation of an AI-based solution for breast cancer risk stratification using routine digital histopathology images

Abhinav Sharma[1], Sandy Kang Lövgren[5], Kajsa Ledesma Eriksson[5], Yinxi Wang[5], Stephanie Robertson[2], Johan Hartman[2,3,4], Mattias Rantalainen[1,4]

[1.] Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden
[2.] Department of Oncology-Pathology, Karolinska Institutet, Stockholm, Sweden
[3.] Department of Clinical Pathology and Cancer Diagnostics, Karolinska University Hospital, Stockholm, Sweden
[4.] MedTechLabs, BioClinicum, Karolinska University Hospital, Solna, Sweden
[5.] Stratipath AB, Solna, Sweden

Abstract

Background

Stratipath Breast is a CE-IVD marked artificial intelligence-based solution for prognostic risk stratification of breast cancer patients into high- and low-risk groups, using haematoxylin and eosin (H&E)-stained histopathology whole slide images (WSIs). In this validation study, we assess the prognostic performance of Stratipath Breast in independent breast cancer cases.

Methods

This is a retrospective multi-site study including patients (N=2475) diagnosed with primary breast cancer at two hospitals in Sweden. The Stratipath Breast tool was applied to stratify patients based on digitised WSIs of the diagnostic H&E-stained tissue sections from surgically resected tumours. The prognostic performance was evaluated using time-to-event analysis by multivariable Cox Proportional Hazards analysis with progression-free survival (PFS) as the primary endpoint.

Results

In the clinically relevant oestrogen receptor (ER)+/human epidermal growth factor receptor 2 (HER2)- patient subgroup, the estimated Hazard Ratio (HR) associated with PFS between low- and high-risk groups was 2.61 (95% CI: 1.50 - 4.50, $p < 0.001$) after adjusting for established risk factors. In the ER+/HER2- Nottingham histological grade (NHG) 2 (intermediate risk) subgroup, the HR was 1.99 (95% CI: 1.10 - 3.60, $p = 0.023$) between low- and high-risk groups.

Conclusion

The results indicate an independent prognostic value of Stratipath Breast in both the general breast cancer population, in the clinically relevant ER+/HER2- subgroup and in the NHG2/ER+/HER2- subgroup. Improved risk stratification of intermediate-risk ER+/HER2- breast cancers provides information relevant for treatment decisions of adjuvant chemotherapy and has the potential to reduce both under- and over-treatment. Image-based risk stratification provides the added benefit of short lead times and substantially lower cost

compared to molecular diagnostics and therefore has the potential to reach broader patient groups.

Introduction

In breast cancer, Nottingham histological grade (NHG) is a well-established prognostic factor that provides information for clinical decision-making (1,2). Patients with oestrogen receptor (ER)-positive/human epidermal growth factor receptor 2 (HER2)-negative tumours with high-risk clinical factors such as high histological grade (NHG3) are considered for adjuvant chemotherapy whereas patients with low-grade (NHG1) tumours are recommended endocrine therapy only. However, more than 50% of patients belong to the intermediate risk category (NHG2), with limited clinical value (3). Consequently, many patients may be overtreated with considerable side effects.

Several multigene expression-based methods have been developed to predict the risk of recurrence for ER+ early-stage breast cancer patients (4–6). The most commonly used commercial solutions for clinical use include MammaPrint (Agendia Inc., Amsterdam, The Netherlands) (7), Oncotype DX (Exact Sciences Corp., Madison, WI, USA) (8), EndoPredict/EPclin (Myriad Genetics Inc., Salt Lake City, UT, USA) (9) and Prosigna (Veracyte Inc., South San Francisco, CA, USA) (10). Oncotype DX recurrence score (RS) was one of the first multigene assay-based tests to predict the distant recurrence in ER+ node-negative tamoxifen-treated breast cancer patients by stratifying patients into low-, intermediate- and high-risk groups (8). RS was validated to predict distant recurrence (11) and was further compared with the PAM50-based Prosigna Risk of Recurrence (ROR) score in the same study population (12). In addition, RS was prospectively validated and shown to be predictive of chemotherapy benefit (13). Prosigna ROR score has been validated to predict the distant and late-distant (5-10 years) recurrence in postmenopausal hormone receptor-positive early breast cancer patients (4,10,14). Furthermore, models based on molecular signatures, like the Genomic grade index (GGI) (15), have been proposed to improve the stratification of intermediate-grade patients (16). However, in clinical practice, the use of molecular multigene assays implies long lead times and remains expensive (17).

Enabled by the emergence of digital and computational pathology, deep learning-based whole slide image (WSI) classification models have recently been demonstrated to enable improved prognostic stratification of patients compared to routine pathology, and to enable faster and more precise information for clinical decision-making in several cancer types (17–21). Stratipath Breast (Stratipath AB, Stockholm, Sweden) is the first CE-IVD marked artificial intelligence (AI)-based image analysis tool for primary breast cancer risk stratification that is commercially available for clinical use. It provides risk stratification of breast cancer patients into low- and high-risk groups based on the assessment of routine haematoxylin and eosin (H&E)-stained tumour tissue slides from surgical resection specimens. Risk stratification of intermediate-risk ER+/HER2- patients is of particular interest in the clinical setting, as improved prognostic information can provide guidance relating to treatment decisions for adjuvant chemotherapy, and thus reduce over- and under-treatment of patients. Stratipath

Breast only requires an H&E-stained WSI from the resected tumour as input and uses a deep learning-based model to risk-stratify patients into low- and high-risk groups.

It is of clinical importance that the prognostic performance and generalisability of AI-based tools for decision support are validated in independent study materials to ensure that the solution provides significant prognostic value and that it generalises well to other sites. In this study, we validated the prognostic performance of Stratipath Breast in two fully independent breast cancer cohorts from two sites in Sweden.

Methods

Study Materials

In this study, we included breast cancer patients from two Swedish studies: the CHIME breast KS Solna cohort and the SCAN-B (Lund) study. CHIME breast KS Solna is a population representative retrospective cohort including primary breast cancer patients (N=2922) from the Karolinska University Hospital Solna, Stockholm, diagnosed between 2009-2018. The SCAN-B (Lund) study includes a subset of the patients (N=1262) from the prospective SCAN-B study (22), diagnosed between 2010-2019 in Lund, Sweden. Archived H&E-stained WSIs were retrieved and scanned at 40x magnification using Hamamatsu NanoZoomer XR or S360 whole slide scanner (Hamamatsu Photonics K.K., Shizuoka, Japan) and Aperio GT 450 DX Digital Pathology slide scanner (Leica Biosystems, Wetzlar, Germany). One H&E-stained tumour WSI for each patient was used for analyses. Patients diagnosed with invasive breast cancer without receiving neoadjuvant treatment were included in this study. Only WSIs from resection specimens were included. 160 WSIs from each study were used for the Stratipath Breast system setup procedure, and these were excluded from further analyses. All exclusion criteria are described in the CONSORT diagram in Figure 1. A total of 2475 patients were included for further analyses and the baseline characteristics of the cohorts are available in Supplementary Table 1.

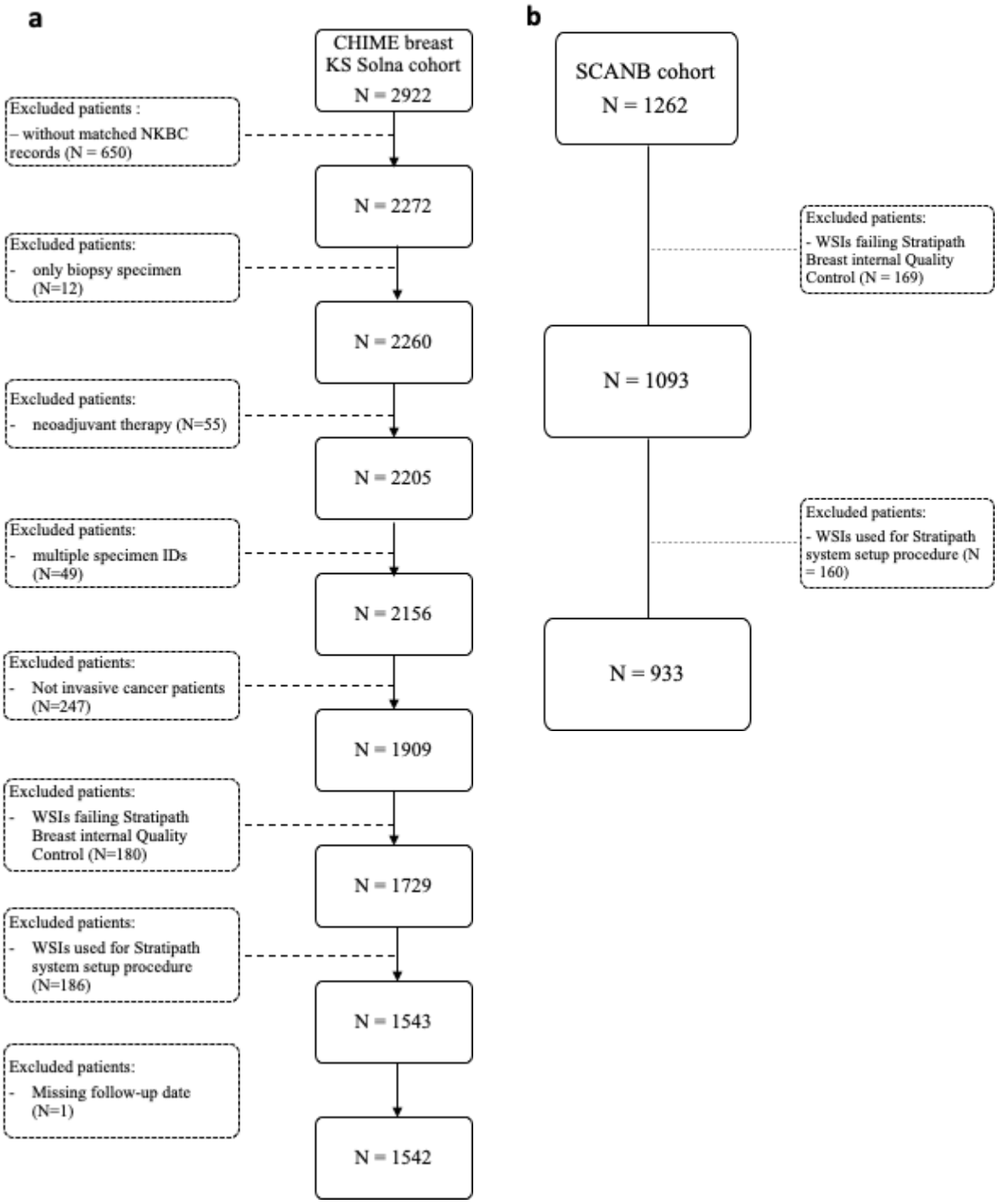

Figure 1. CONSORT diagram for the cohorts from two different sites in Sweden. **a.** CHIME breast KS Solna cohort with included patient whole slide images from Stockholm, Sweden. **b.** SCAN-B cohort with included patient whole slide images from Lund, Sweden. NKBC=National quality registry for breast cancer. WSI=whole slide image.

Statistical analyses

We evaluated the prognostic performance of the patient risk groups (high risk and low risk) assigned by the Stratipath Breast analysis. Progression-free survival (PFS) was the survival

endpoint, defined as the time to local recurrence, distant metastasis or detection of contralateral tumours. The patients were followed from the initial date of diagnosis to the date of reported progression, or the last follow-up date (whichever came first). Kaplan-Meier (KM) curves were used to visualise the survival probability over time, for the risk groups assigned by Stratipath Breast, as well as for clinical NHG. The association between the different patient groups and PFS was assessed by estimating the Hazard ratio (HR) with 95% confidence interval (CI) using the Cox Proportional Hazard (PH) model. The multivariable Cox PH model was fitted by adjusting for established clinical covariates including: age at the time of diagnosis, tumour size, lymph node status, ER status, and HER2 status. Tumour size was converted to a binary covariate with tumour size <=20mm and >20mm. Lymph node status was defined as positive (metastasis to one or more lymph nodes) or negative (no lymph node metastasis). ER status was defined in accordance with local clinical guidelines, as the presence of >= 10% ER-positive cells in the immunohistochemistry (IHC) staining. HER2 status was defined according to national guidelines using IHC staining and for equivocal cases, gene amplification was confirmed by fluorescent or silver *in situ* hybridisation assays. Clinical information was retrieved from the National Quality Registry for Breast Cancer (NKBC). Patients with missing clinical information were excluded from the analyses. All time-to-event analysis was performed using the *survival* package in R (23). P-value < 0.05 was considered as statistically significant in all the analyses. All statistical analyses were performed in R version 4.2.0.

Results

Prognostic patient stratification by tumour grade and Stratipath Breast

First, we assessed the prognostic stratification based on clinically assigned NHG, and the risk groups assigned by Stratipath Breast in all patients. We observed a significant difference in the PFS rate over time between the NHG groups (log-rank p-value < 0.0001) and between the Stratipath Breast low- and high-risk groups (log-rank p-value < 0.0001) (Figure 2a and 2b). The association between NHG3 and PFS with NHG1 as the reference was not found to be statistically significant (p-value > 0.05) in the multivariable Cox PH analysis (Figure 2c). The adjusted HR for PFS between the Stratipath Breast low- and high-risk groups was estimated to 2.42 (95% CI: 1.44-4.10, p-value < 0.001) (Figure 2d).

Next, we performed subgroup analysis in the ER+/HER2- patient subgroup, by clinical NHG and Stratipath Breast. We observed a significant HR for PFS between the NHG groups (log-rank p-value = 0.0038), as well as between the Stratipath Breast low- and high-risk groups (log-rank p-value < 0.0001) (Figure 2e and 2f). However, we did not observe a statistically significant association between the NHG3 and PFS in the multivariable cox PH model (Figure 2g). We observed an adjusted HR for PFS of 2.61 (95% CI: 1.50-4.50, p-value < 0.001) for Stratipath low- vs high-risk groups (Figure 2h).

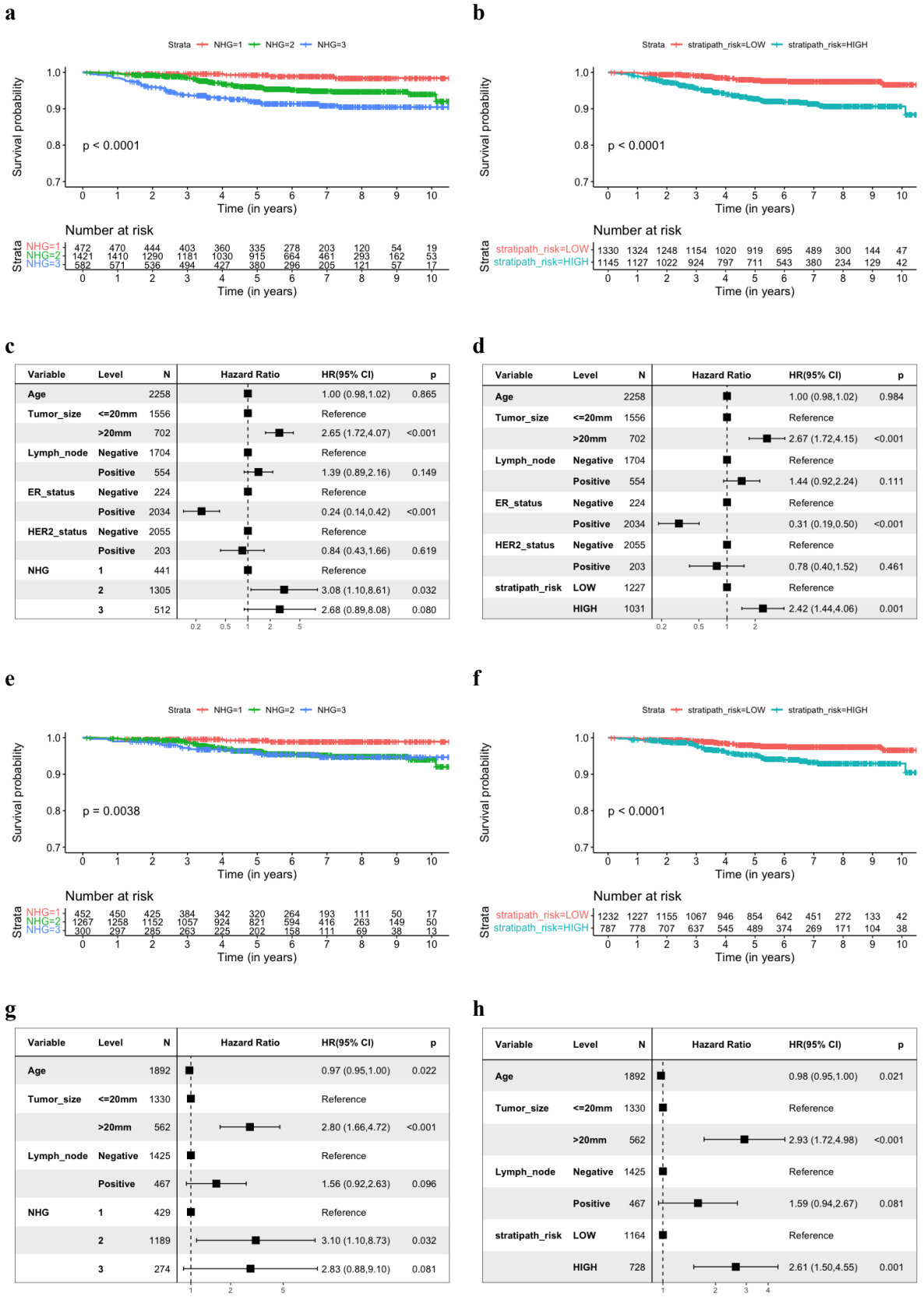

Figure 2. Prognostic stratification of patients by clinical NHG and Stratipath Breast. **a** Kaplan Meier (KM) curves for progression-free survival (PFS) among all patients stratified

by clinical NHG. **b** KM curves for PFS among all patients stratified by Stratipath Breast risk groups. **c** Multivariable Cox Proportional Hazard (PH) regression estimating the association between NHG and PFS. **d** Multivariable Cox PH regression estimating the association between Stratipath Breast risk groups and PFS. **e** KM curves for PFS among ER+/HER2- patients stratified by clinical NHG. **f** KM curves for PFS among ER+/HER2- patients stratified by Stratipath Breast. **g** Multivariable Cox PH regression estimating the association between NHG and PFS in the ER+/HER2- patient subgroup. **h** Multivariable Cox PH regression estimating the association between Stratipath Breast risk groups and PFS in the ER+/HER2- subgroup. NHG=Nottingham histological grade, ER=oestrogen receptor, HER2=human epidermal growth factor receptor 2.

Prognostic performance in the subgroup of intermediate-risk patients

Next, we evaluated the prognostic performance of the Stratipath Breast in the clinically relevant subgroup of intermediate-risk breast cancers. We investigated the NHG2 subgroup (Figure 3a), as well as the subgroup of NHG2/ER+/HER2- patients (Figure 3b) and observed significant difference in the PFS rate over time between the Stratipath Breast low- and high-risk groups in the KM curves (log-rank p-value < 0.0001) (Figure 3a and 3b). In multivariable Cox PH analysis, we observed an HR for PFS of 2.10 in the NHG2 subgroup (95% CI: 1.19-3.7, p-value = 0.01), and an HR of 1.99 (95% CI: 1.10-3.60, p-value = 0.023) in the NHG2/ER+/HER2- subgroup (Figure 3c and 3d).

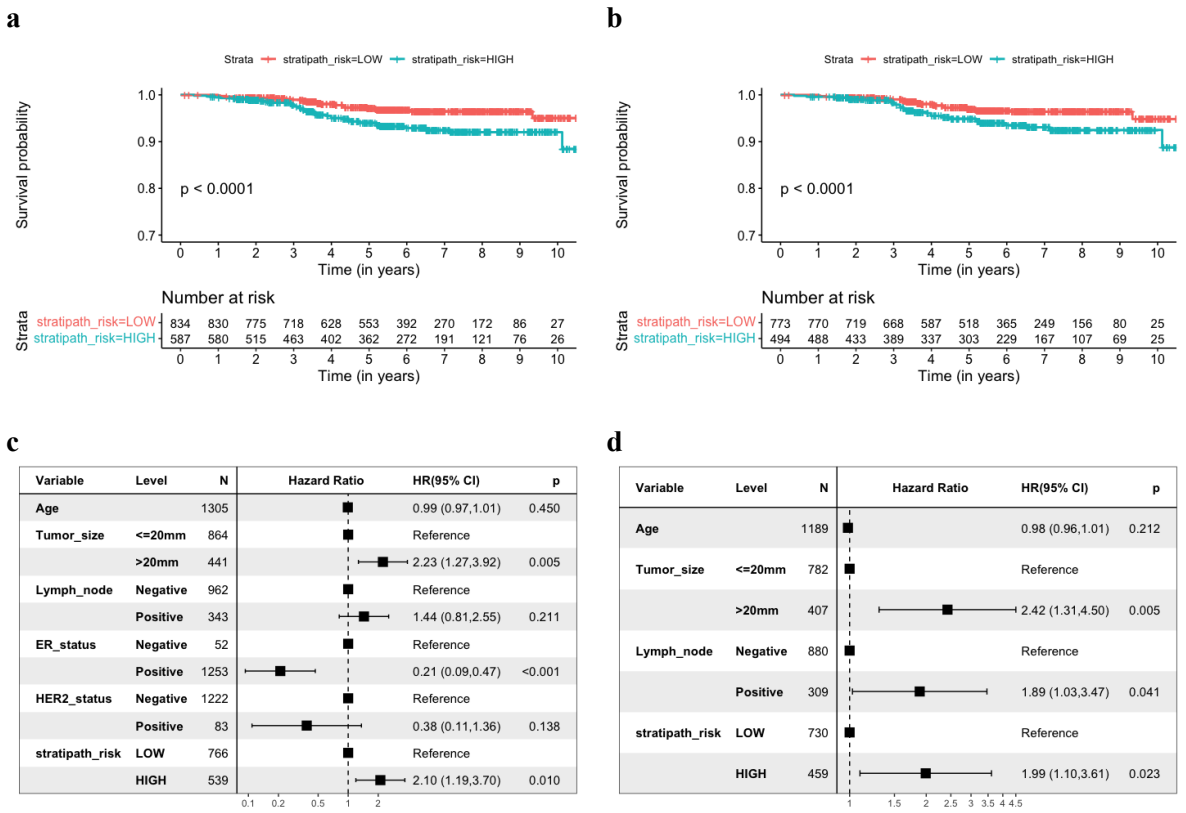

Figure 3. Prognostic stratification of patients with intermediate grade (NHG2) using

Stratipath Breast with progression-free survival (PFS) as the outcome. **a** Kaplan Meier (KM) curves for PFS among NHG2 patients stratified by Stratipath Breast risk groups. **b** KM curves for PFS among NHG2/ER+/HER2- patients stratified by Stratipath Breast risk groups. **c** Multivariate Cox Proportional Hazard (PH) regression estimating the association between Stratipath Breast risk groups and PFS in NHG2 patients. **d** Multivariate Cox PH regression estimating the association between Stratipath Breast risk groups and PFS in the NHG2/ER+/HER2- patient subgroup. NHG=Nottingham histological grade, ER=oestrogen receptor, HER2=human epidermal growth factor receptor 2.

Discussion

In this study, we validated the prognostic performance, using PFS as the endpoint, of the deep learning-based CE-IVD marked Stratipath Breast solution for risk stratification of invasive breast tumours from H&E-stained histopathology WSIs. We assessed the prognostic performance across all breast cancer cases from two different hospitals, in the clinically relevant ER+/HER2- subgroup, and in the intermediate-risk subgroup.

Histological grading has a high inter-observer variability in clinical routine (3,24–26) and more than 50% of the patients in the ER+/HER2- patient subgroup are assigned as the intermediate grade (NHG2) (3). However, the NHG2 assignment is not informative in clinical decision-making and is not directly related to the molecular characteristic of the tumours (27). Therefore, there is a need for assays or methods that can provide a refined risk stratification of the intermediate-grade patients to provide decision support when considering patients for adjuvant chemotherapy. Moreover, as a consequence of the low reproducibility in histological grading, it would be clinically valuable to enable more precise and reliable risk stratification of all breast cancer patients.

In terms of clinical value, the ER+/HER2- subgroup is not only the largest but also the group for which risk stratification is most valuable for therapy decision-making. In the other subgroups of breast cancer, namely HER2+ and triple negative, most patients are treated with adjuvant chemotherapy and/or targeted drugs. In the present study, we observed an adjusted HR for PFS of 1.99 (95% CI: 1.10-3.61) for the re-stratification of intermediate-grade patients (NHG2) in the ER+/HER2- subgroup. Wang et al. reported an adjusted HR for recurrence-free survival of 1.91 in the re-stratification of NHG2 patients into low- and high-risk groups among ER+/HER2- patients, using an academic model (research use only) in their independent external validation cohort (17).

Several gene expression-based methods have been developed for clinical use to predict the risk of distant recurrence, especially for the selection of patients benefiting from adjuvant chemotherapy among ER+/HER2- node-negative early breast cancer patients. Popular methods include PAM50-based Prosigna ROR and Oncotype DX RS that stratify ER+/HER2- patients into low-, intermediate- and high-risk groups. Oncotype DX RS has been validated in retrospective analysis of prospective clinical trials (28,29) and prospectively

in a prospective clinical trial (13,30,31). However, some of the retrospective comparative studies assessing the prognostic performance of the various multigene assays have reported better predictive performance of the overall and late distant recurrence using PAM50-based ROR score, Breast Cancer Index (14) and EndoPredict (EPclin) (32) over Oncotype DX RS in ER+/HER2- node-negative breast cancer patients (5,12). Most importantly, there is poor correlation between the different gene expression profiling tests (33). Another limitation of the molecular profiling methods is the use of bulk tumour tissue analysis, in which intra-tumour heterogeneity is not accounted for (34). Furthermore, the effect of intra-tumour heterogeneity on the prognostic misclassification of the individual patients using multigene assays has been shown (35).

Three-risk group stratification from Oncotype DX RS showed an adjusted HR of 2.5 for distant recurrence (95% CI: 1.30-4.50) between intermediate- vs low-risk groups and HR of 5.2 (95% CI: 2.7-10.1) for high- vs low-risk groups in the node-negative postmenopausal hormonal receptor-positive early breast cancer patients from the TransATAC study population (11). PAM50-based Prosigna ROR derived risk-groups reported adjusted HRs of 2.15 for distant recurrence (95% CI: 1.21-3.81, p=0.009) between intermediate- vs low-risk groups and of 4.26 (95% CI: 2.44-7.43, p<0.0001) for high- vs low-risk groups in the postmenopausal hormonal receptor-positive early breast cancer patients derived from the ABCSG-8 trial (4). Further, for the postmenopausal hormonal receptor-positive early breast cancer patients from the combined cohort derived from the combined TransATAC and ABCSG-8 trial, they observed HRs of 3.75 (95% CI: 2.19-6.41) for intermediate- vs low-risk and 5.49 (95% CI: 2.92-10.35) for high- vs low-risk groups in the HER2- node-negative patient subgroup with distant recurrence between 5 to 10 years as the survival endpoint (10). Notably, the patients that were retrospectively derived from the clinical trials in the mentioned validation studies, were randomised to the treatment arm (tamoxifen or anastrozole) and did not receive adjuvant chemotherapy. This could lead to a lower survival rate, especially among high-risk patients and might result in larger effect sizes for high- vs low- and intermediate- vs low-risk groups.

The study has several limitations. Since, the discussed multigene assays, except for EPclin, stratify patients into three-tier risk groups (low, intermediate and high) and have been validated in different cohorts, it is difficult to directly compare the prognostic performance with the binary stratification method (low and high) by Stratipath Breast. However, it remains an important question for the future to be answered, preferably with comparisons in the same study population. Another limitation of the study is the relatively short median follow-up time (6.25 years) in the present study in comparison with prognostic validation studies of multigene assays in retrospective analysis of the clinical trials.

Furthermore, molecular profiling-based academic models have shown significant prognostic value in re-stratification of intermediate-grade patients into low-and high-risk groups (15,16). Sotiriou et al. developed the genomic grade index (GGI) based on the identified 97-gene signatures and observed an estimated HR of 3.61 associated with the risk of recurrence (15). Wang et al. reported an estimated HR of 2.43 associated with recurrence-free survival based

on the 37 gene RNA-seq profiling for the re-stratification of the intermediate-grade patients (16). Using AI-based image analysis, Stratipath Breast provides prognostic re-stratification based on routinely available digitised WSIs of H&E-stained breast tumours without any further tissue preparation. Image-based risk stratification also provides the added benefit of short lead times and substantially lower cost compared to molecular diagnostics and therefore has the potential to reach broader patient populations.

In conclusion, this study assesses the prognostic performance of Stratipath Breast in independent data from two Swedish hospitals, using PFS as the endpoint. This image analysis tool provides a fast and cost-effective solution to provide precise prognostic patient stratification that can be used as a decision support tool in clinical decision-making, with the potential to contribute to the reduction of both over- and under-treatment of breast cancer patients.

List of abbreviations

H&E: Haematoxylin & Eosin
WSI: Whole Slide Image
ER: Oestrogen Receptor
HER2: Human Epidermal Growth Factor Receptor 2
HR: Hazard Ratio
PFS: Progression-Free Survival
NHG: Nottingham Histological Grade
RS: Risk Score
ATAC: Arimidex, Tamoxifen Alone or in Combination
ROR: Risk Of Recurrence
ABCSG-8: Austrian Breast and Colorectal Study Group-8
GGI: Genomic Grade Index
AI: Artificial Intelligence
KM: Kaplan-Meier
IHC: Immunohistochemistry
CI: Confidence Interval
PH: Proportional Hazard

Ethics approval and consent to participate

The study has approval from the Swedish Ethical Review Authority (2017/2106-31, with amendments 2018/1462-32 and 2019-02336).

Consent for publication

Not Applicable.

Availability of data and materials

Data cannot be made public due to local legislation since it contains sensitive patient data.

Author's contribution:

AS performed statistical analysis and computing, visualisation, drafted and edited the manuscript.
SKL, KLE and YW contributed to data analysis and editing of the draft manuscript.
SR contributed to editing of the draft manuscript. JH conceived and conceptualised the study, contributed to preparing and editing of the draft manuscript, and provided resources and supervision. MR conceived, conceptualised and directed the study, contributed to preparing and editing of the draft manuscript, and provided resources and supervision. All authors approved of the final manuscript.

Acknowledgements:

The authors acknowledge patients, clinicians, and hospital staff participating in the SCAN-B study, the staff at the central SCAN-B laboratory at the Division of Oncology, Lund University, the Swedish National Breast Cancer Quality Registry (NKBC), Regional Cancer Center South and the South Swedish Breast Cancer Group (SSBCG). We also acknowledge the help and support from Dr Johan Vallon-Christersson at Lund University (Sweden) with the preparation of clinicopathological information for the SCAN-B study. SCAN-B was funded by the Swedish Cancer Society, the Mrs. Berta Kamprad Foundation, the Lund-Lausanne L2-Bridge/Biltema Foundation, the Mats Paulsson Foundation, and Swedish governmental funding (ALF). Stratipath AB provided the Stratipath Breast analysis.

Funding statement:

This work was supported by funding from the Swedish Research Council, Swedish Cancer Society, Karolinska Institutet, ERA PerMed (ERAPERMED2019-224-ABCAP), VINNOVA and SweLIFE (SwAIPP project), MedTechLabs, Swedish e-science Research Centre (SeRC) - eCPC, Stockholm Region, Stockholm Cancer Society and Swedish Breast Cancer Association.

Competing interest:

JH has obtained speaker's honoraria or advisory board remunerations from Roche, Novartis, AstraZeneca, Pfizer, Eli Lilly, MSD and Gilead, and has received institutional research support from Cepheid, Novartis, Roche and AstraZeneca. MR and JH are co-founders and

shareholders of Stratipath AB. SR is an employee at Stratipath AB and holds employee stock options. All other authors have declared no conflicts of interest.

References

1. Rakha EA, Reis-Filho JS, Baehner F, Dabbs DJ, Decker T, Eusebi V, et al. Breast cancer prognostic classification in the molecular era: the role of histological grade. Breast Cancer Res. 2010 Jul 30;12(4):207.

2. Rakha EA, El-Sayed ME, Lee AHS, Elston CW, Grainge MJ, Hodi Z, et al. Prognostic significance of Nottingham histologic grade in invasive breast carcinoma. J Clin Oncol. 2008 Jul 1;26(19):3153–8.

3. Acs B, Fredriksson I, Rönnlund C, Hagerling C, Ehinger A, Kovács A, et al. Variability in Breast Cancer Biomarker Assessment and the Effect on Oncological Treatment Decisions: A Nationwide 5-Year Population-Based Study. Cancers [Internet]. 2021 Mar 9;13(5). Available from: http://dx.doi.org/10.3390/cancers13051166

4. Gnant M, Filipits M, Greil R, Stoeger H, Rudas M, Bago-Horvath Z, et al. Predicting distant recurrence in receptor-positive breast cancer patients with limited clinicopathological risk: using the PAM50 Risk of Recurrence score in 1478 postmenopausal patients of the ABCSG-8 trial treated with adjuvant endocrine therapy alone. Ann Oncol. 2014 Feb;25(2):339–45.

5. Sestak I, Buus R, Cuzick J, Dubsky P, Kronenwett R, Denkert C, et al. Comparison of the Performance of 6 Prognostic Signatures for Estrogen Receptor-Positive Breast Cancer: A Secondary Analysis of a Randomized Clinical Trial. JAMA Oncol. 2018 Apr 1;4(4):545–53.

6. Wallden B, Storhoff J, Nielsen T, Dowidar N, Schaper C, Ferree S, et al. Development and verification of the PAM50-based Prosigna breast cancer gene signature assay. BMC Med Genomics. 2015 Aug 22;8:54.

7. Bueno-de-Mesquita JM, van Harten WH, Retel VP, van ’t Veer LJ, van Dam FS, Karsenberg K, et al. Use of 70-gene signature to predict prognosis of patients with node-negative breast cancer: a prospective community-based feasibility study (RASTER). Lancet Oncol. 2007 Dec;8(12):1079–87.

8. Paik S, Shak S, Tang G, Kim C, Baker J, Cronin M, et al. A multigene assay to predict recurrence of tamoxifen-treated, node-negative breast cancer. N Engl J Med. 2004 Dec 30;351(27):2817–26.

9. Filipits M, Rudas M, Jakesz R, Dubsky P, Fitzal F, Singer CF, et al. A new molecular predictor of distant recurrence in ER-positive, HER2-negative breast cancer adds independent information to conventional clinical risk factors. Clin Cancer Res. 2011 Sep 15;17(18):6012–20.

10. Sestak I, Cuzick J, Dowsett M, Lopez-Knowles E, Filipits M, Dubsky P, et al. Prediction of late distant recurrence after 5 years of endocrine treatment: a combined analysis of

patients from the Austrian breast and colorectal cancer study group 8 and arimidex, tamoxifen alone or in combination randomized trials using the PAM50 risk of recurrence score. J Clin Oncol. 2015 Mar 10;33(8):916–22.

11. Dowsett M, Cuzick J, Wale C, Forbes J, Mallon EA, Salter J, et al. Prediction of risk of distant recurrence using the 21-gene recurrence score in node-negative and node-positive postmenopausal patients with breast cancer treated with anastrozole or tamoxifen: a TransATAC study. J Clin Oncol. 2010 Apr 10;28(11):1829–34.

12. Dowsett M, Sestak I, Lopez-Knowles E, Sidhu K, Dunbier AK, Cowens JW, et al. Comparison of PAM50 risk of recurrence score with oncotype DX and IHC4 for predicting risk of distant recurrence after endocrine therapy. J Clin Oncol. 2013 Aug 1;31(22):2783–90.

13. Sparano JA, Gray RJ, Makower DF, Pritchard KI, Albain KS, Hayes DF, et al. Prospective Validation of a 21-Gene Expression Assay in Breast Cancer. N Engl J Med. 2015 Nov 19;373(21):2005–14.

14. Sgroi DC, Sestak I, Cuzick J, Zhang Y, Schnabel CA, Schroeder B, et al. Prediction of late distant recurrence in patients with oestrogen-receptor-positive breast cancer: a prospective comparison of the breast-cancer index (BCI) assay, 21-gene recurrence score, and IHC4 in the TransATAC study population. Lancet Oncol. 2013 Oct;14(11):1067–76.

15. Sotiriou C, Wirapati P, Loi S, Harris A, Fox S, Smeds J, et al. Gene expression profiling in breast cancer: understanding the molecular basis of histologic grade to improve prognosis. J Natl Cancer Inst. 2006 Feb 15;98(4):262–72.

16. Wang M, Klevebring D, Lindberg J, Czene K, Grönberg H, Rantalainen M. Determining breast cancer histological grade from RNA-sequencing data. Breast Cancer Res. 2016 May 10;18(1):48.

17. Wang Y, Acs B, Robertson S, Liu B, Solorzano L, Wählby C, et al. Improved breast cancer histological grading using deep learning. Ann Oncol. 2022 Jan;33(1):89–98.

18. Ström P, Kartasalo K, Olsson H, Solorzano L, Delahunt B, Berney DM, et al. Artificial intelligence for diagnosis and grading of prostate cancer in biopsies: a population-based, diagnostic study. Lancet Oncol. 2020 Feb;21(2):222–32.

19. Campanella G, Hanna MG, Geneslaw L, Miraflor A, Werneck Krauss Silva V, Busam KJ, et al. Clinical-grade computational pathology using weakly supervised deep learning on whole slide images. Nat Med. 2019 Aug;25(8):1301–9.

20. Couture HD, Williams LA, Geradts J, Nyante SJ, Butler EN, Marron JS, et al. Image analysis with deep learning to predict breast cancer grade, ER status, histologic subtype, and intrinsic subtype. NPJ Breast Cancer. 2018 Sep 3;4:30.

21. Coudray N, Ocampo PS, Sakellaropoulos T, Narula N, Snuderl M, Fenyö D, et al. Classification and mutation prediction from non-small cell lung cancer histopathology images using deep learning. Nat Med. 2018 Oct;24(10):1559–67.

22. Vallon-Christersson J, Häkkinen J, Hegardt C, Saal LH, Larsson C, Ehinger A, et al.

Cross comparison and prognostic assessment of breast cancer multigene signatures in a large population-based contemporary clinical series. Sci Rep. 2019 Aug 21;9(1):12184.

23. Therneau TM. A Package for Survival Analysis in R [Internet]. 2023. Available from: https://CRAN.R-project.org/package=survival

24. Ginter PS, Idress R, D'Alfonso TM, Fineberg S, Jaffer S, Sattar AK, et al. Histologic grading of breast carcinoma: a multi-institution study of interobserver variation using virtual microscopy. Mod Pathol. 2021 Apr;34(4):701–9.

25. Zhang R, Chen HJ, Wei B, Zhang HY, Pang ZG, Zhu H, et al. Reproducibility of the Nottingham modification of the Scarff-Bloom-Richardson histological grading system and the complementary value of Ki-67 to this system. Chin Med J . 2010 Aug 5;123(15):1976–82.

26. van Dooijeweert C, van Diest PJ, Willems SM, Kuijpers CCHJ, van der Wall E, Overbeek LIH, et al. Significant inter- and intra-laboratory variation in grading of invasive breast cancer: A nationwide study of 33,043 patients in the Netherlands. Int J Cancer. 2020 Feb 1;146(3):769–80.

27. Ivshina AV, George J, Senko O, Mow B, Putti TC, Smeds J, et al. Genetic reclassification of histologic grade delineates new clinical subtypes of breast cancer. Cancer Res. 2006 Nov 1;66(21):10292–301.

28. Paik S, Tang G, Shak S, Kim C, Baker J, Kim W, et al. Gene expression and benefit of chemotherapy in women with node-negative, estrogen receptor-positive breast cancer. J Clin Oncol. 2006 Aug 10;24(23):3726–34.

29. Sparano JA, Paik S. Development of the 21-gene assay and its application in clinical practice and clinical trials. J Clin Oncol. 2008 Feb 10;26(5):721–8.

30. Sparano JA, Gray RJ, Makower DF, Pritchard KI, Albain KS, Hayes DF, et al. Adjuvant Chemotherapy Guided by a 21-Gene Expression Assay in Breast Cancer. N Engl J Med. 2018 Jul 12;379(2):111–21.

31. Sparano JA, Gray RJ, Ravdin PM, Makower DF, Pritchard KI, Albain KS, et al. Clinical and Genomic Risk to Guide the Use of Adjuvant Therapy for Breast Cancer. N Engl J Med. 2019 Jun 20;380(25):2395–405.

32. Buus R, Sestak I, Kronenwett R, Denkert C, Dubsky P, Krappmann K, et al. Comparison of EndoPredict and EPclin With Oncotype DX Recurrence Score for Prediction of Risk of Distant Recurrence After Endocrine Therapy. J Natl Cancer Inst [Internet]. 2016 Nov;108(11). Available from: http://dx.doi.org/10.1093/jnci/djw149

33. Bartlett JMS, Bayani J, Marshall A, Dunn JA, Campbell A, Cunningham C, et al. Comparing Breast Cancer Multiparameter Tests in the OPTIMA Prelim Trial: No Test Is More Equal Than the Others. J Natl Cancer Inst [Internet]. 2016 Sep;108(9). Available from: http://dx.doi.org/10.1093/jnci/djw050

34. Pece S, Disalvatore D, Tosoni D, Vecchi M, Confalonieri S, Bertalot G, et al. Identification and clinical validation of a multigene assay that interrogates the biology of cancer stem cells and predicts metastasis in breast cancer: A retrospective consecutive

study. EBioMedicine. 2019 Apr;42:352–62.

35. Gyanchandani R, Lin Y, Lin HM, Cooper K, Normolle DP, Brufsky A, et al. Intratumor Heterogeneity Affects Gene Expression Profile Test Prognostic Risk Stratification in Early Breast Cancer. Clin Cancer Res. 2016 Nov 1;22(21):5362–9.