



OPEN

## Integrating age, BMI, and serum N-glycans detected by MALDI mass spectrometry to classify suspicious mammogram findings as benign lesions or breast cancer

Calvin R. K. Blaschke<sup>1</sup>, Elizabeth G. Hill<sup>2</sup>, Anand S. Mehta<sup>1</sup>, Peggi M. Angel<sup>1</sup>, Christine Laronga<sup>3</sup> & Richard R. Drake<sup>1</sup>✉

While mammograms are the standard tool for breast cancer screening, there remains challenges for mammography to effectively distinguish benign lesions from breast cancers, leading to many unnecessary biopsy procedures. A blood-based biomarker could provide a minimally invasive supplemental assay to increase the specificity of breast cancer screening. Serum N-glycosylation alterations have associations with many cancers and several of the clinical characteristics of breast cancer. The current study utilized a high-throughput mass spectrometry workflow to identify serum N-glycans with differences in intensities between patients that had a benign lesion from patients with breast cancer. The overall N-glycan profiles of the two patient groups had no differences, but there were several individual N-glycans with significant differences in intensities between patients with benign lesions and ductal carcinoma in situ (DCIS). Many N-glycans had strong associations with age and/or body mass index, but there were several of these associations that differed between the patients with benign lesions and breast cancer. Accordingly, the samples were stratified by the patient's age and body mass index, and N-glycans with significant differences between these subsets were identified. For women aged 50–74 with a body mass index of 18.5–24.9, a model including the intensities of two N-glycans, 1850.666 m/z and 2163.743 m/z, age, and BMI were able to clearly distinguish the breast cancer patients from the patients with benign lesions with an AUROC of 0.899 and an optimal cutoff with 82% sensitivity and 84% specificity. This study indicates that serum N-glycan profiling is a promising approach for providing clarity for breast cancer screening, especially within the subset of healthy weight women in the age group recommended for mammograms.

Breast cancer has the highest incidence rate of all cancers, and is the second leading cause of cancer deaths in US women<sup>1</sup>. The stage of the breast cancer at diagnosis has a large impact on the prognosis of the patient, highlighting the need for early detection strategies<sup>1</sup>. Accordingly, the implementation of mammography as a screening tool has decreased breast cancer mortality<sup>2</sup>. While mammograms have a high cancer detection rate, issues of overdiagnosis and false-positives continue to affect their clinical utility<sup>3</sup>. Due to the similar presentations of benign lesions and breast cancer in a mammogram, abnormal interpretations are followed up with additional imaging or a biopsy. A study of over 1.6 million mammograms found that 1.2% of mammograms were classified as suspicious or highly suggestive of malignancy, but after the recommended biopsy, only 71.6% of those patients had breast cancer<sup>4</sup>. With over 39 million mammograms performed in the US every year, this results in over 345,000 unnecessary biopsies yearly in the US alone<sup>5</sup>. To improve breast cancer screening, a supplemental assay is needed to aid in the discrimination of benign lesions and breast cancer.

While mammography is an essential tool in breast cancer screening, issues of low specificity frequently result in unnecessary biopsies or follow-up imaging. These false positives have significant financial costs on a national

<sup>1</sup>Department of Cell & Molecular Pharmacology & Experimental Therapeutics, Medical University of South Carolina, Charleston 29425, USA. <sup>2</sup>Hollings Cancer Center, Medical University of South Carolina, Charleston 29425, USA. <sup>3</sup>Department of Breast Oncology, H. Lee Moffitt Cancer Center & Research Institute, Tampa 33612, USA. ✉email: draker@musc.edu

and individual patient level, as well as a psychological cost for the patient<sup>6–9</sup>. A minimally or non-invasive biomarker for breast cancer detection is needed to increase the specificity of breast cancer screening. Alternative imaging modalities have displayed improvements in cancer detection, especially for women with a higher breast density, but suffer from increased radiation exposure, high costs and low availability, or similar issues with false-positives for digital breast tomosynthesis (DBT), magnetic resonance imaging (MRI), and ultrasound (US), respectively<sup>10–16</sup>. Using blood-based biomarkers for screening benefits from cost-effective and minimally invasive collection. The currently used blood-based biomarkers for breast cancer, such as cancer antigen (CA) 15-3, CA 27.29, carcinoembryonic antigen (CEA), and soluble human epidermal growth factor receptor 2 (sHER2), are primarily used for prognosis, staging, and therapy monitoring, but have limited utility in detection<sup>17–21</sup>. These limitations prompt an investigation into other blood-based biomarkers that can detect differences in the systemic response to breast cancer and benign lesions.

Serum is a rich source of potential biomarkers because of the variety of molecules circulating in the body with altered abundance, activation, and/or composition due to the effects of a disease. Serum biomarkers can be readily adapted into clinical assays because non-invasive collection is cost-effective and large cohorts are available for thorough and accurate analysis of diagnostic potential. The composition of serum collected during blood draws reflects a dynamic biofluid comprised of thousands of proteins, lipids, metabolites, and nucleic acids. A large portion of serum consists of glycosylated proteins produced by the liver and immunoglobulins secreted by B-cells<sup>22</sup>. Glycosylation is a highly regulated metabolic process shown to have essential roles in protein folding, molecular trafficking, protein clearance, and many other processes<sup>23–26</sup>. Being closely aligned with function, altered glycoforms of serum proteins can be dynamic indicators of systemic responses to disease and have been interrogated for potential clinical applications for decades<sup>27–29</sup>.

The creation and optimization of high-throughput techniques for the profiling of serum N-glycans, i.e. a class of glycans linked to a protein via an asparagine residue, has enabled the interrogation of large sample cohorts for robust detection of N-glycan alterations associated with diseases and conditions<sup>30,31</sup>. These methods have been essential for identifying N-glycan alterations associated with clinical factors, such as age and BMI. Age has been characterized by an increase in non-galactosylated serum N-glycans, while BMI is associated with an increase in sialylation of biantennary serum N-glycans<sup>32,33</sup>. Recently, our group has adapted matrix-assisted laser desorption/ionization (MALDI) mass spectrometry techniques for analysis of serum, urine, prostatic fluids, cultured cells, and antibody-captured serum proteins<sup>34–39</sup>. Here, we utilized a high-throughput MALDI mass spectrometry technique<sup>34</sup> to identify significant differences in the serum N-glycan profiles of 199 women with benign lesions from 99 women with breast cancer. The relative intensity of individual N-glycans and N-glycan classes were compared and evaluated for the ability to discriminate between benign and cancerous conditions, with and without stratification by age and BMI.

## Methods

**Serum samples.** Donors were women who reported to the breast surgical clinic at the Moffitt Cancer Center & Research Institute for a breast biopsy after a Breast Imaging Reporting and Data System (BI-RADS) 4 imaging designation. The samples were collected after H. Lee Moffitt Cancer Center & Research Institute institutional review board approval and informed consent from the donors. All experiments were performed in accordance with this protocol following the approved, relevant guidelines and regulations. Serum samples were collected immediately prior to tissue biopsy and deidentified. Age, body mass index (BMI), ethnicity/race, and pathology results of benign lesions (fibroadenomas, fibrocystic, etc.) and breast cancers (invasive or ductal carcinoma in situ (DCIS)) were linked with each specimen. Patients with both an invasive breast cancer and DCIS were treated as invasive. Patients self-identified their ethnicity, and patients identifying as non-Hispanic were asked to self-identify race. The benign samples (n = 199) and breast cancer samples (n = 99) were stratified by age groups according to the current United States Preventative Services Task Force mammogram recommendations (younger than 40, 40–49, 50–74, older than 74) and the established BMI classifications (less than 18.5, 18.5–24.9, 25–29.9, more than 29.9)<sup>40</sup>. The clinical characteristics of the samples analyzed in this study are summarized in Table 1.

**Materials.** Trifluoroacetic acid (TFA), sodium bicarbonate, and  $\alpha$ -cyano-4-hydroxycinnamic acid (CHCA) were purchased from Sigma-Aldrich (St. Louis, MO). Ethanol was purchased from Decon Labs (King of Prussia, PA). Acetonitrile, phosphate buffered saline, HPLC grade water, citraconic anhydride, xylene, methanol, glacial acetic acid, and chloroform were purchased from Fisher Scientific (Hampton, NH). The amine-reactive slides (Nexterion® Slide H) were purchased from Applied Microarray (Tempe, AZ). The attachable well chambers (Pro-Plate Multi-Array Slide System, 64-well) were purchased from Grace Bio-Laboratories (Bend, OR). Peptide-N-glycosidase F (PNGase F Prime) was purchased from N-Zyme Scientifics (Doylestown, PA).

**Sample preparation.** Serum sample preparation was performed as previously described<sup>34</sup>. Serum samples were diluted in sodium bicarbonate and spotted onto an amine-reactive slide. Each slide had spot of PBS diluted in sodium bicarbonate added as a blank. Additionally, a standard healthy serum sample was added to each slide for normalization across slides. All samples and standards were spotted in technical triplicate. The slide was then placed into a humidity chamber for 1 h to bind the serum proteins to the slide. Well chambers were attached to the slide to isolate each sample, and the samples were washed with Carnoy's solution and water to remove lipids and salts, respectively. After drying the slides, PNGase F was sprayed across the slide with an automated sprayer (M5 TM-Sprayer, HTX Technologies, Chapel Hill, NC) and incubated in a humidity chamber at 37 °C for 2 h to enzymatically cleave the N-glycans from the captured serum glycoproteins. The HTX M5 sprayer was then used to spray CHCA matrix across the slide.

		Benign (n = 199)	Cancer (n = 99)
Ethnicity/Race	Caucasian	151	85
	Hispanic	38	10
	African American	8	1
	Asian	2	3
BMI	Mean $\pm$ SD	28.3 $\pm$ 5.9	28.7 $\pm$ 6.19
	< 18.5	1	2
	18.5–24.9	68	21
	25–29.9	60	42
	> 30	70	34
Age	Mean $\pm$ SD	56.7 $\pm$ 11.3	57.3 $\pm$ 11.9
	< 40	14	7
	40–49	35	17
	50–74	140	68
	> 74	10	7
Pathology	DCIS	–	30
	Invasive	–	69

**Table 1.** Clinical characteristics of samples by diagnosis. BMI, body mass index; SD, standard deviation; DCIS, ductal carcinoma in situ.

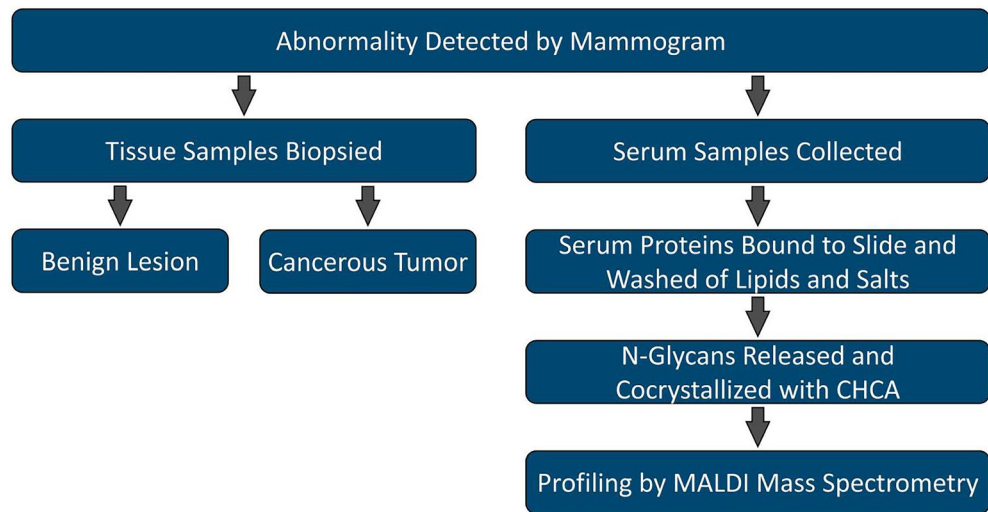
**Mass spectrometry data collection.** The slides were imaged with a timsTOF fleX MALDI-QTOF mass spectrometer (Bruker, Billerica, MA) with a SmartBeam 3D laser operating at 10,000 Hz and 20  $\mu$ m laser spot size. There were 300 laser shots collected per pixel, and a 150  $\mu$ m raster was used. A 700–4000 m/z range was scanned in positive ion mode.

**Data processing.** The mass spectra were imported into SCiLS Lab software (2022b Pro, Bruker, Billerica, MA). Total ion current was used to normalize the images. N-glycan peaks were selected manually using theoretical mass values. Putative structures were assigned to N-glycan peaks based on previously reported assignments<sup>41–43</sup>. N-glycans are reported as m/z values and using the Oxford nomenclature for the putative structures, where A represents the number of antennae present, F indicates the fucose, B indicates the presence of a bisecting N-acetylglucosamine, G represents galactoses, and S denotes sialic acids<sup>44</sup>. Multiply sodiated species of the same N-glycan structure were processed, analyzed, and reported individually. N-glycan compositions and mass error are reported in Supplementary Table S1. Maximum peak values were extracted and used for N-glycan absolute intensities. For each slide, the intensity of the blank sample was subtracted from the N-glycan intensity across the samples. N-glycans with less than 2 arbitrary units (au) of intensity after blank subtraction were converted to 1 au to avoid 0 values for the statistical analysis. If a N-glycan was less than 2 au for more than 20% of samples, it was discarded from further analysis. Next, the standard sample was used to create normalization factors for each N-glycan on each slide, where the intensity of the individual slide's standard was divided by the average intensity of the standards across all slides. Each slide's N-glycans were then multiplied by the corresponding normalization factor. The intensities were then converted to relative intensities by dividing individual N-glycan intensities by the summed intensities of all the N-glycans in that sample. N-glycan class intensities were calculated by summing the relative intensities of the N-glycans containing the structural or compositional traits of the N-glycan class.

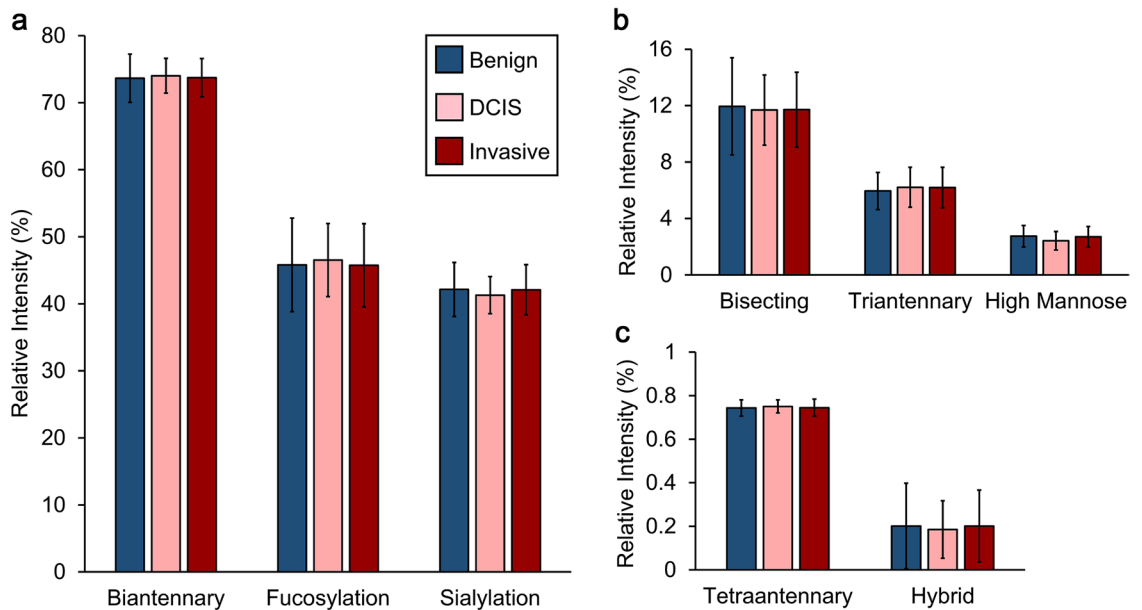
**Statistics.** Linear regression was used to model log-transformed N-glycan relative intensities as a function of disease status, controlling for age, ethnicity/race, and BMI. Logarithmic transformation was used to induce approximate normality and stabilize variance. Differences in log-relative intensities across disease status levels were evaluated using model-based contrasts. A 5% false-discovery rate was used to adjust for test multiplicity<sup>45,46</sup>. A similar modeling approach was used to evaluate the association between N-glycan intensities and age or BMI. Loess smoothing was used for graphical exploration of the functional relationship prior to modeling<sup>47</sup>. The performance of N-glycans to classify disease status was evaluated using the area under the receiver operating characteristic curve (AUROC) of a logistic regression model of the patient pathology as a function of the selected N-glycans, N-glycan classes, and clinical characteristics. AUROCs are reported with 95% confidence intervals. Optimal sensitivity and specificity were defined as the sensitivity and specificity of the threshold at the cutoff point of the curve closest to (0,1). All statistical analyses were performed using R version 4.1.0.

## Results

Serum N-glycan profiles of women with either a benign lesion or breast cancer detected after a mammogram were determined by MALDI mass spectrometry of enzymatically cleaved N-glycans from serum glycoproteins captured and washed on an amine-reactive slide as previously described (Fig. 1)<sup>34</sup>. After data processing and application of data quality criteria, 55 N-glycans, including 10 multiply sodiated N-glycans, were detected across the samples with an average coefficient of variation (CV) of 7.9% across the technical triplicates (Supplementary



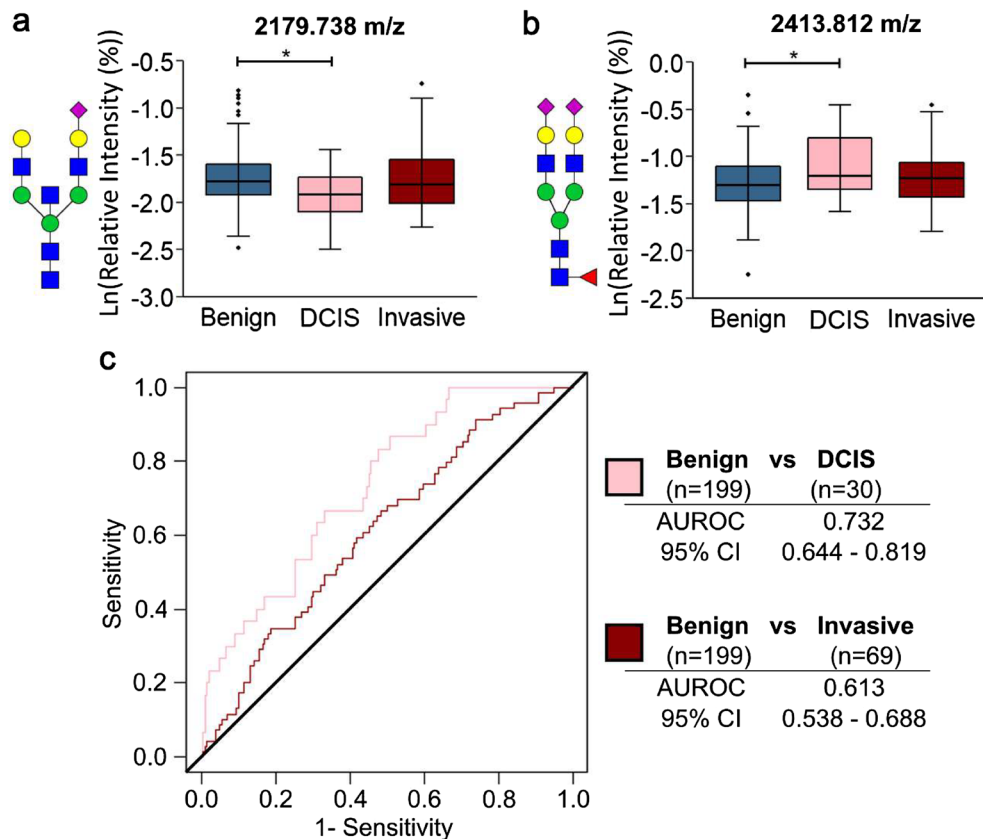
**Figure 1.** Sample collection, processing, and analysis overview. MALDI, matrix-assisted laser/desorption ionization; CHCA,  $\alpha$ -cyano-4-hydroxycinnamic acid.



**Figure 2.** N-glycan class intensities of serum samples with benign lesions, DCIS, and invasive breast cancer. (a) High, (b) medium, and (c) low intensity serum N-glycan classes of benign, DCIS, and invasive samples. Data was analyzed by linear regression modeling of log-transformed relative intensities controlled for age, ethnicity/race, and BMI to identify differences across disease status. Using a 5% false-discovery rate, no significant associations were found.

Table S1). The clinical descriptions of the patients included ethnicity/race, age, and BMI, and had no significant differences between the benign and cancerous samples (Table 1).

As the clinical characteristics and prognosis of DCIS and invasive breast cancer vary greatly, initial analysis was focused on detecting differences between benign, DCIS, and invasive breast cancer. No differences were detected between the N-glycan classes, with the majority of the N-glycan profile intensity being attributed to biantennary, fucosylated, and/or sialylated N-glycans (Fig. 2). As no differences in the overall N-glycan structural class profiles were detected, individual N-glycan intensities were investigated. There were two N-glycans (2179.738 m/z (A2BG2S1) and 2413.812 m/z (A2G2F1S2)) with significantly different intensities in the benign and DCIS serum samples (Fig. 3, Supplementary Table S2). The 2179.738 m/z N-glycan has a bisected and sialylated structure and had a lower intensity in the DCIS samples. The 2413.812 m/z N-glycan has a fucosylated and multiply sialylated biantennary structure, and had a higher intensity in the DCIS samples (Fig. 3a, b). Based on a logistic model using the intensities of these two N-glycans, age, ethnicity/race, and BMI, an AUROC of 0.732



**Figure 3.** Serum N-glycan differences between benign, DCIS, and invasive samples. Log-transformed relative intensity values of the (a) 2179.738 m/z (A2BG2S1) and (b) 2413.812 m/z (A2G2F1S2) N-glycans had significant differences in benign and DCIS samples using a 5% false discovery rate. (c) Using the log-transformed relative intensities of these N-glycans, age, BMI, and ethnicity/race, benign and DCIS samples could be distinguished with an AUROC of 0.732, and benign and invasive samples could be distinguished with an AUROC of 0.613. N-glycan compositions are represented by blue squares for N-acetylglucosamine, green circles for mannose, yellow circles for galactose, purple diamonds for sialic acid, and red triangles for fucose.

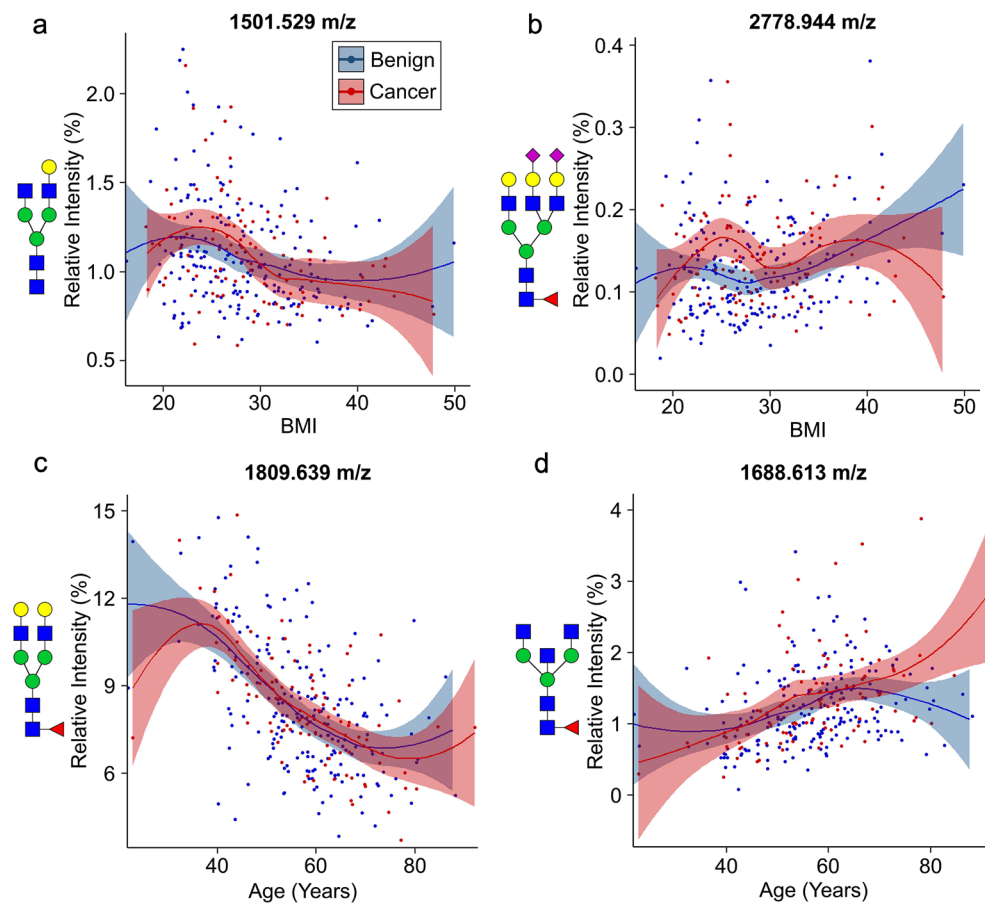
(0.644–0.819) was found for the discrimination of benign and DCIS samples (Fig. 3c). Using the same variables, an AUROC of 0.613 (0.538–0.688) was found for the discrimination of benign and invasive samples (Fig. 3c). For the evaluation of the entire set based on clinical designation, there were no N-glycan classes or individual N-glycans with differences between the benign and invasive or benign and all cancerous samples.

Additional analysis identified significant relationships between N-glycan intensities and patient age or BMI. There were 24 N-glycans associated with age, as well as the triantennary, tetraantennary, and fucosylation N-glycan classes (Supplementary Table S3). Additionally, there were 12 N-glycans associated with BMI, as well as the biantennary, triantennary, and tetraantennary N-glycan classes (Supplementary Table S4). While some of these associations were present in both benign and cancerous samples (Fig. 4a, c), others were only present in benign samples or displayed differences between the samples at certain ranges (Fig. 4b, d). These findings prompted an in-depth analysis of the samples stratified by age and BMI.

The samples were separated by age (less than 40, 40–49, 50–74, and older than 74) and/or by BMI (less than 18.5, 18.5–24.9, 25–30, and more than 30). There were no significant findings between the benign and cancerous samples in any individual or combination of these groups, except for the patients that were 50–74 years old and within 18.5–24.9 BMI. These samples had significant differences in intensity between benign (n = 45) and cancerous samples (n = 11) for two N-glycans with, 1850.666 m/z (A2BG1F1) and 2163.743 m/z (A2BG1F1S1), and the bisecting N-glycan class (Fig. 5a–c, Supplementary Table S5). Based on a logistic model using the intensities of 1850.666 m/z and 2163.743 m/z, age, and BMI an AUROC of 0.899 (0.801–0.997) found for the discrimination of benign and cancerous samples with an 82% sensitivity and 84% specificity (Fig. 5d).

## Discussion

In this study, the serum N-glycan profiles of patients with an abnormality detected by a mammogram were compared to evaluate their ability to distinguish benign lesions from breast cancer. The samples were processed using a recently developed high-throughput MALDI mass spectrometry workflow<sup>34</sup>. When analyzing all the samples, two N-glycans had significant differences in the benign and DCIS samples. Incorporating the intensities of these N-glycans with age, ethnicity/race, and BMI had moderate diagnostic potential for samples from patients with

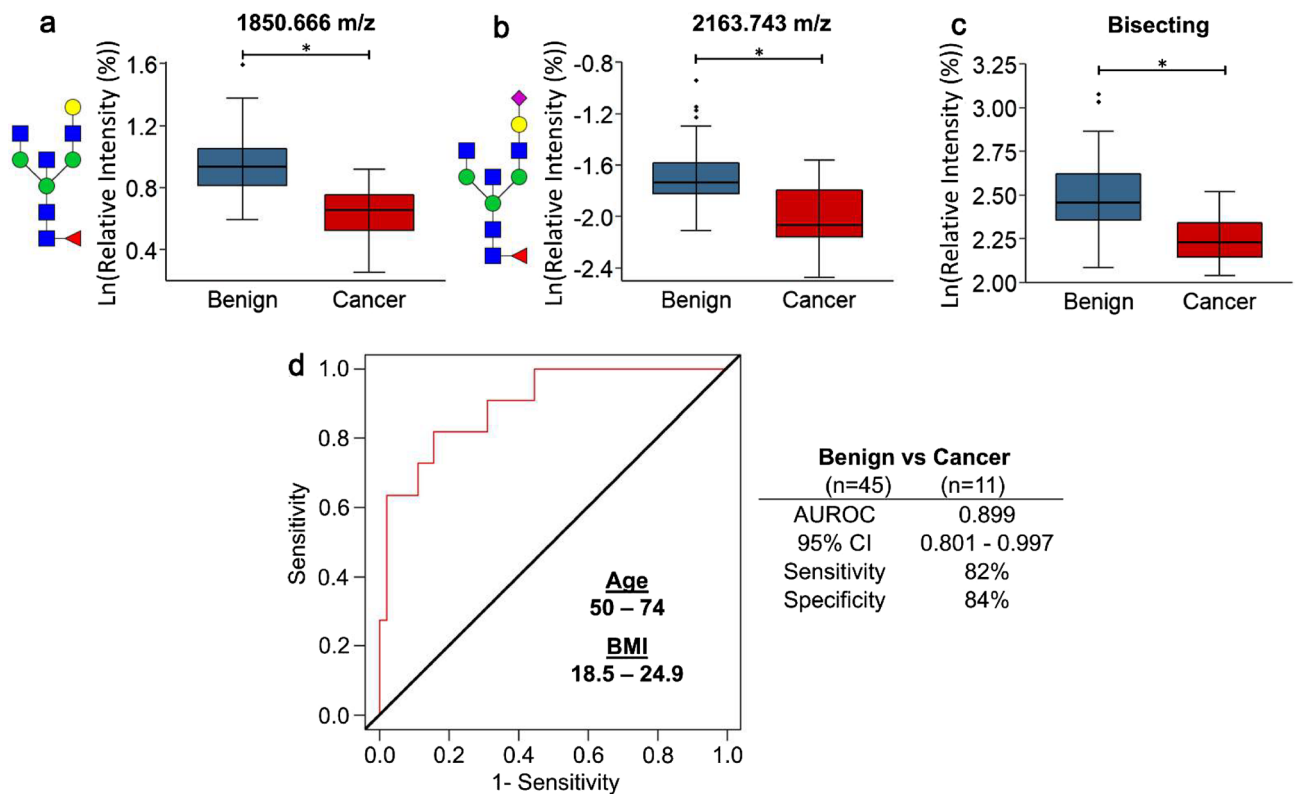


**Figure 4.** Correlation of serum N-glycan intensities with the patient age and BMI. The patient BMI and the intensities of both (a) 1501.529 m/z (A2G1) and (b) 2778.944 m/z (A3G3F1S2) N-glycans, but the 2778.944 m/z N-glycan and BMI association diverged for benign and cancer samples in the patients with a BMI from 20 to 30. The patient age and the intensities of both (c) 1809.639 m/z (A2G2F1) and (d) 1688.613 m/z (A2BF1) N-glycans were significantly associated, but the 1688.613 m/z N-glycan and age association diverged for benign and cancer samples in the older patients. N-glycan compositions are represented by blue squares for N-acetylglucosamine, green circles for mannose, yellow circles for galactose, purple diamonds for sialic acid, and red triangles for fucose.

DCIS but not from patients with invasive breast cancer. There were many N-glycan intensities that had altered associations with age or BMI, depending on the diagnosis of the patient. Stratifying the patients by age and BMI identified two N-glycans and two N-glycan classes with significant intensity differences in the samples from patients with benign lesions and breast cancer aged 50–74 with a BMI of 18.5–24.9, and when combined with age and BMI had high diagnostic ability for this patient subset.

Inflammatory stimuli due to the presence of a disease can cause an alteration in the glycosylation of plasma cells and hepatocytes, which produce the bulk of serum proteins<sup>48,49</sup>. Altered serum N-glycan profiles have been associated with a range of diseases and conditions, including Crohn's disease, gastric cancer, ovarian cancer, Alzheimer's disease, and hepatocellular carcinoma<sup>50–54</sup>. Serum analysis of breast cancer patients have identified N-glycan changes associated with circulating tumor cell counts, breast cancer subtypes, prognosis, breast cancer progression, and lymph node metastasis<sup>55–60</sup>. While several studies have investigated breast cancer detection by serum N-glycan analysis, they have either had healthy women donors as controls, used methods unable to identify individual N-glycans, or only analyzed certain types of N-glycans<sup>61–63</sup>. To properly address the issue of false-positives in mammograms, the samples from patients with breast cancer need to be compared to samples from patients with a lesion detected by a mammogram that requires a biopsy, because healthy women with no abnormalities detected by mammogram do not suffer from the issues of false-positives. This cohort can appropriately investigate the potential of serum N-glycan analysis for breast cancer detection because of the clinical relevance of the samples collected from women with benign lesions, the controlled surgical collection site, and the availability and incorporation of age, ethnicity/race, and BMI into the analysis.

Age and BMI have strong associations with breast cancer risk, as well as specific alterations in the serum N-glycan profile<sup>29,64–66</sup>. In this study, there were 24 N-glycans with significant associations with age, and 12 N-glycans with significant associations with BMI (Supplementary Tables S3 and S4). Age and BMI were controlled for in the initial analysis to account for these associations and isolate N-glycan intensity differences that were due to the presence of breast cancer. Further investigation found that only samples within certain ranges of age and BMI



**Figure 5.** Serum N-glycan differences between benign ( $n=45$ ) and cancer ( $n=11$ ) samples from patients aged 50–74 and with a BMI of 18.5–24.9. Log-transformed relative intensity values of the (a) 1850.666 m/z (A2BG1F1) and (b) 2163.743 m/z (A2BG1F1S1) N-glycans and (c) bisecting N-glycan class had significant differences in benign and cancer samples using a 5% false discovery rate. (d) Using the log-transformed relative intensities of the two individual N-glycans, age, and BMI, benign and cancer samples could be distinguished with an AUROC of 0.899, sensitivity of 82%, and specificity of 84%. N-glycan compositions are represented by blue squares for *N*-acetylglucosamine, green circles for mannose, yellow circles for galactose, purple diamonds for sialic acid, and red triangles for fucose.

had disease-associated serum N-glycan changes (Fig. 4b–d). To identify any disease-associated serum N-glycan intensities changes only present in a subset of the cohort, the samples were stratified according to age and BMI. To remain clinically relevant, the samples were grouped based on the USPSTF mammogram recommendation for age and the BMI classifications for underweight, healthy weight, overweight, and obese<sup>40</sup>.

The subset of samples from patients with a BMI of 18.5–24.9 and aged 50–74 could be readily distinguished as having a benign lesion or breast cancer using the intensities of two N-glycans, age, and BMI (Fig. 5d). While this is a subset of the population, it makes up a significant portion of the U.S. women that will have mammogram screening for breast cancer. For women with an average-risk of breast cancer, the USPSTF recommends only women aged 50–74 to have biennial mammography screening and make up the bulk of women that will have a mammogram<sup>40</sup>. Results from the 2017–2018 National Health and Nutrition Examination Survey (NHANES) indicated that approximately one third of U.S. adult women are within the 18.5–24.9 BMI range<sup>67</sup>. The healthy weight women may have had the best N-glycan biomarker performance due to the lack of systemic inflammation that is present in overweight and obese women<sup>68</sup>. The plasma cells and hepatocytes, from which the majority of serum glycoproteins originate, undergo significant cellular changes in presence of inflammatory stimuli. The serum N-glycan changes found in cancer patients may be masked, altered, or negated by the N-glycan changes due to the obesity-induced inflammatory stimuli.

There were several limitations to this study. Breast cancer is an incredibly heterogeneous disease, that includes multiple molecular and genomic subtypes with distinct clinical behavior and biological characteristics: luminal-A, luminal-B, HER2-positive, and basal-like. The risk factors for breast cancer, including age, body mass index (BMI), onset of menopause, parity, family-history, and personal history of benign breast disease differ in effect for each molecular subtype<sup>69</sup>. As subtype information was not available for this cohort, there may be N-glycan alterations only in certain subtypes that are not being detected. Additionally, there were fewer samples from Asian or African American patients. This makes it more difficult to account for variances due to ethnicity/race, especially after stratifying the cohort into smaller subsets. In the 11 breast cancer patients who were aged 50–74 and within the 18.5–24.9 BMI range, 4 had DCIS and 7 had invasive breast cancer. While both breast cancer groups are represented in these samples, future studies with more patients to represent this subset of women can solidify these findings.

Additional studies identifying the protein(s) that the distinctive N-glycans originated from could increase the performance of the N-glycans as breast cancer detection biomarkers. Recent glycoproteomic approaches have

made high-throughput analysis of complex biological fluids quicker and more sensitive<sup>70</sup>. Glycopeptide analysis of this cohort could identify the proteins that contain the cancer-associated N-glycans identified in this study. Alternatively, a technique using MALDI mass spectrometry to N-glycan profile antibody-captured serum proteins could be used to identify breast cancer-associated N-glycan alterations across an array of targets<sup>39,71</sup>. Isolating low-abundant proteins shed from breast cancer cells or released from dying cancer cells could identify N-glycan changes with a higher specificity across all age and BMI ranges. To further investigate the fucosylated bisecting N-glycans that had a significant association with breast cancer, the location of the fucose could be determined by applying Endoglycosidase F3 to the samples instead of PNGase F to detect only core fucosylated N-glycans<sup>34</sup>.

In summary, a MALDI mass spectrometry approach was utilized to detect serum N-glycans and N-glycan classes that could distinguish patients with an abnormal mammogram result as having a benign lesion or breast cancer. For a significant portion of the women screened by mammogram, fucosylated bisecting N-glycans could readily distinguish patient groups.

### Data availability

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

Received: 9 September 2022; Accepted: 28 November 2022

Published online: 02 December 2022

### References

1. Siegel, R. L., Miller, K. D., Fuchs, H. E. & Jemal, A. Cancer statistics, 2022. *CA Cancer J. Clin.* **72**, 7–33 (2022).
2. Pace, L. E. & Keating, N. L. A systematic assessment of benefits and risks to guide breast cancer screening decisions. *JAMA* **311**, 1327–1335 (2014).
3. Nelson, H. D. *et al.* Effectiveness of breast cancer screening: Systematic review and meta-analysis to update the 2009 U. S. Preventive services task force recommendation. *Ann. Intern. Med.* **164**, 244–255 (2016).
4. Lehman, C. D. *et al.* National performance benchmarks for modern screening digital mammography: Update from the Breast Cancer Surveillance Consortium. *Radiology* **283**, 49–58 (2017).
5. US Food and Drug Administration. MQSA national statistics. *MQSA Insights*. <https://www.fda.gov/radiation-emitting-products/mqsa-insights/mqsa-national-statistics> (2022).
6. Tosteson, A. N. A. *et al.* Consequences of false-positive screening mammograms. *JAMA Intern. Med.* **174**, 954–961 (2014).
7. Ong, M. S. & Mandl, K. D. National expenditure for false-positive mammograms and breast cancer overdiagnoses estimated at \$ 4 billion a year. *Health Aff.* **34**, 576–583 (2015).
8. Vlahiotis, A., Griffin, B., Stavros, A. T. & Margolis, J. Analysis of utilization patterns and associated costs of the breast imaging and diagnostic procedures after screening mammography. *Clin. Outcomes Res.* **10**, 157 (2018).
9. Chubak, J., Boudreau, D. M., Fishman, P. A. & Elmore, J. G. Cost of breast-related care in the year following false positive screening mammograms. *Med. Care* **48**, 815 (2010).
10. Kolb, T. M., Lichy, J. & Newhouse, J. H. Comparison of the performance of screening mammography, physical examination, and breast US and evaluation of factors that influence them: An analysis of 27,825 patient evaluations. *Radiology* **225**, 165–175 (2002).
11. Boyd, N. F. *et al.* Mammographic density: The risk and detection of breast cancer. *J. Biol. Chem.* **282**, 60–61 (2007).
12. Sechopoulos, I. A review of breast tomosynthesis. Part I. The image acquisition process. *Med. Phys.* **40**, 014301 (2013).
13. Olgar, T., Kahn, T. & Gosch, D. Average glandular dose in digital mammography and breast tomosynthesis. *RoFo* **184**, 911–918 (2012).
14. Haas, J. S. *et al.* Disparities in the use of screening magnetic resonance imaging of the breast in community practice by race, ethnicity, and socioeconomic status. *Cancer* **122**, 611–617 (2016).
15. Berg, W. A. *et al.* Reasons women at elevated risk of breast cancer refuse breast MR imaging screening: ACRIN 6666. *Radiology* **254**, 79 (2010).
16. Evans, A. *et al.* Breast ultrasound: Recommendations for information to women and referring physicians by the European Society of Breast Imaging. *Insights Imaging* **9**, 449–461 (2018).
17. Duffy, M. J. Serum tumor markers in breast cancer: Are they of clinical value?. *Clin. Chem.* **52**, 345–351 (2006).
18. Klee, G. G. & Schreiber, W. E. MUC1 gene-derived glycoprotein assays for monitoring breast cancer (CA 15-3, CA 27.29, BR): Are they measuring the same antigen?. *Arch. Pathol. Lab. Med.* **128**, 1131–1135 (2004).
19. Ali, S. M. *et al.* Serum HER-2/neu and relative resistance to trastuzumab-based therapy in patients with metastatic breast cancer. *Cancer* **113**, 1294–1301 (2008).
20. Swellam, M. *et al.* Aberrant methylation of APC and RAR  $\beta$ 2 genes in breast cancer patients. *IUBMB Life* **67**, 61–68 (2015).
21. Gao, J., Zhang, Q., Xu, J., Guo, L. & Li, X. Clinical significance of serum miR-21 in breast cancer compared with CA153 and CEA. *Chin. J. Cancer Res.* **25**, 743–748 (2013).
22. Anderson, N. L. & Anderson, N. G. The human plasma proteome: History, character, and diagnostic prospects. *Mol. Cell. Proteom.* **1**, 845–867 (2002).
23. Moremen, K. W., Tiemeyer, M. & Nairn, A. V. Vertebrate protein glycosylation: Diversity, synthesis and function. *Nat. Rev. Mol. Cell Biol.* **13**, 448–462 (2012).
24. Tanaka, T. *et al.* Aberrant N-glycosylation profile of serum immunoglobulins is a diagnostic biomarker of urothelial carcinomas. *Int. J. Mol. Sci.* **18**, 1–14 (2017).
25. Higel, F. *et al.* N-glycans of complex glycosylated biopharmaceuticals and their impact on protein clearance. *Eur. J. Pharm. Biopharm.* **139**, 123–131 (2019).
26. Bas, M. *et al.* Fc sialylation prolongs serum half-life of therapeutic antibodies. *J. Immunol.* **202**, 1582–1594 (2019).
27. Adamczyk, B., Tharmalingam, T. & Rudd, P. M. Glycans as cancer biomarkers. *Biochim. Biophys. Acta Gen. Subj.* **1820**, 1347–1353 (2012).
28. Kailemia, M. J., Park, D. & Lebrilla, C. B. Glycans and glycoproteins as specific biomarkers for cancer. *Anal. Bioanal. Chem.* **409**, 395–410 (2017).
29. Dotz, V. & Wührer, M. N-glycome signatures in human plasma: Associations with physiology and major diseases. *FEBS Lett.* **593**, 2966–2976 (2019).
30. Reiding, K. R. *et al.* High-throughput serum N-glycomics: Method comparison and application to study rheumatoid arthritis and pregnancy-associated changes. *Mol. Cell. Proteom.* **18**, 3–15 (2019).
31. Reiding, K. R. *et al.* Human plasma N-glycosylation as analyzed by matrix-assisted laser desorption/ionization-fourier transform ion cyclotron resonance-MS associates with markers of inflammation and metabolic health. *Mol. Cell. Proteom.* **16**, 228–242 (2017).



32. Miura, Y. & Endo, T. Glycomics and glycoproteomics focused on aging and age-related diseases—glycans as a potential biomarker for physiological alterations. *Biochim. Biophys. Acta - Gen. Subj.* **1860**, 1608–1614 (2016).
33. Knežević, A. *et al.* Effects of aging, body mass index, plasma lipid profiles, and smoking on human plasma N-glycans. *Glycobiology* **20**, 959–969 (2010).
34. Blaschke, C. R. K., Black, A. P., Mehta, A. S., Angel, P. M. & Drake, R. R. Rapid N-glycan profiling of serum and plasma by a novel slide-based imaging mass spectrometry workflow. *J. Am. Soc. Mass Spectrom.* **31**, 2511–2520 (2020).
35. Blaschke, C. R. K. *et al.* Direct N-glycosylation profiling of urine and prostatic fluid glycoproteins and extracellular vesicles. *Front. Chem.* **9**, 5398 (2021).
36. Blaschke, C. R. K. & Drake, R. R. MALDI mass spectrometry imaging and glycomics. In *MALDI Mass Spectrometry Imaging: From Fundamentals to Spatial Omics* (ed. Porta Siegel, T.) 207–233 (The Royal Society of Chemistry, 2021).
37. McDowell, C. T., Lu, X., Mehta, A. S., Angel, P. M. & Drake, R. R. Applications and continued evolution of glycan imaging mass spectrometry. *Mass Spectrom. Rev.* **2021**, 1–32 (2021).
38. Angel, P. M. *et al.* A rapid array-based approach to N-glycan profiling of cultured cells. *J. Proteome Res.* **18**, 3630–3639 (2019).
39. Black, A. P., Angel, P. M., Drake, R. R. & Mehta, A. S. Antibody panel based N-glycan imaging for N-glycoprotein biomarker discovery. *Curr. Protoc. Protein Sci.* **98**, 257 (2019).
40. Siu, A. L. Screening for breast cancer: US preventive services task force recommendation statement. *Ann. Intern. Med.* **164**, 279–296 (2016).
41. McDowell, C. T. *et al.* Imaging mass spectrometry and lectin analysis of n-linked glycans in carbohydrate antigen-defined pancreatic cancer tissues. *Mol. Cell. Proteom.* **20**, 100012 (2021).
42. de Vroome, S. W. *et al.* Serum N-glycome alterations in colorectal cancer associate with survival. *Oncotarget* **9**, 30610–30623 (2018).
43. Song, T., Aldredge, D. & Lebrilla, C. B. A method for in-depth structural annotation of human serum glycans that yields biological variations. *Anal. Chem.* **87**, 7754–7762 (2015).
44. Harvey, D. J. *et al.* Proposal for a standard system for drawing structural diagrams of N- and O-linked carbohydrates and related compounds. *Proteomics* **9**, 3796–3801 (2009).
45. Storey, J. D. A direct approach to false discovery rates. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **64**, 479–498 (2002).
46. Storey, J. D. & Tibshirani, R. Statistical significance for genomewide studies. *Proc. Natl. Acad. Sci. USA* **100**, 9440–9445 (2003).
47. Cleveland, W. S. Robust locally weighted regression and smoothing scatterplots. *J. Am. Stat. Assoc.* **74**, 829–836 (1979).
48. Wang, J. *et al.* Fc-glycosylation of IgG1 is modulated by B-cell stimuli. *Mol. Cell. Proteom.* **10**, 1–12 (2011).
49. Mackiewicz, A., Ganapathi, M. K., Schultz, D. & Kushner, I. Monokines regulate glycosylation of acute-phase proteins. *J. Exp. Med.* **166**, 253–258 (1987).
50. Wu, Y. *et al.* Probing serum N-glycan patterns for rapid and precise detection of Crohn's disease. *Chem. Commun.* **57**, 11362–11365 (2021).
51. Ozcan, S. *et al.* Serum glycan signatures of gastric cancer. *Cancer Prev. Res.* **7**, 226–235 (2014).
52. Kim, J. H. *et al.* Mass spectrometric screening of ovarian cancer with serum glycans. *Dis. Markers* **2014**, 1–9 (2014).
53. Gizaw, S. T., Ohashi, T., Tanaka, M., Hinou, H. & Nishimura, S. I. Glycoblotting method allows for rapid and efficient glycome profiling of human Alzheimer's disease brain, serum and cerebrospinal fluid towards potential biomarker discovery. *Biochim. Biophys. Acta Gen. Subj.* **1860**, 1716–1727 (2016).
54. Comunale, M. A. *et al.* Total serum glycan analysis is superior to lectin-FLISA for the early detection of hepatocellular carcinoma. *Proteom. Clin. Appl.* **7**, 690–700 (2013).
55. Saldova, R., Reuben, J. M., Abd-Hamid, U. M., Rudd, P. M. & Cristofanilli, M. Levels of specific serum N-glycans identify breast cancer patients with higher circulating tumor cell counts. *Ann. Oncol.* **22**, 1113–1119 (2011).
56. Vreeker, G. C. M. *et al.* Serum N-glycan profiles differ for various breast cancer subtypes. *Glycoconj. J.* **38**, 387–395 (2021).
57. Kyselova, Z. *et al.* Breast cancer diagnosis and prognosis through quantitative measurements of serum glycan profiles. *Clin. Chem.* **54**, 1166–1175 (2008).
58. Ju, L. *et al.* Elevated level of serum glycoprotein bifucosylation and prognostic value in Chinese breast cancer. *Glycobiology* **26**, 460–471 (2016).
59. Abd-Hamid, U. M. *et al.* A strategy to reveal potential glycan markers from serum glycoproteins associated with breast cancer progression. *Glycobiology* **18**, 1105–1118 (2008).
60. de Leoz, M. L. A. *et al.* High-mannose glycans are elevated during breast cancer progression. *Mol. Cell. Proteom.* **10**, 1–9 (2011).
61. Pierce, A. *et al.* Levels of specific glycans significantly distinguish lymph node-positive from lymph node-negative breast cancer patients. *Glycobiology* **20**, 1283–1288 (2010).
62. Saldova, R. *et al.* Association of N-glycosylation with breast carcinoma and systemic features using high-resolution quantitative UPLC. *J. Proteome Res.* **13**, 2314–2327 (2014).
63. Lee, T. B. Comparison of breast cancer screening results in Korean middle-aged women: A hospital-based prospective cohort study. *Osong Public Heal. Res. Perspect.* **4**, 197–202 (2013).
64. DeSantis, C. E. *et al.* Breast cancer statistics, 2019. *CA Cancer J. Clin.* **69**, 438–451 (2019).
65. Pfeiffer, R. M., Webb-Vargas, Y., Wheeler, W. & Gail, M. H. Proportion of US trends in breast cancer incidence attributable to long-term changes in risk factor distributions. *Cancer Epidemiol. Biomark. Prev.* **27**, 1214–1222 (2018).
66. Ding, N. *et al.* Human serum N-glycan profiles are age and sex dependent. *Age Ageing* **40**, 568–575 (2011).
67. Ogden, C. L. *et al.* Trends in obesity prevalence by race and hispanic origin—1999–2000 to 2017–2018. *JAMA* **324**, 1208–1210 (2020).
68. Ellulu, M. S., Patimah, I., Khazaai, H., Rahmat, A. & Abed, Y. Obesity and inflammation: The linking mechanism and the complications. *Arch. Med. Sci.* **13**, 851–863 (2017).
69. Gaudet, M. M. *et al.* Pooled analysis of nine cohorts reveals breast cancer risk factors by tumor molecular subtype. *Cancer Res.* **78**, 6011–6021 (2018).
70. Chen, Z., Huang, J. & Li, L. Recent advances in mass spectrometry (MS)-based glycoproteomics in complex biological samples. *Trends Anal. Chem.* **118**, 880–892 (2019).
71. Black, A. P. *et al.* A novel mass spectrometry platform for multiplexed N-glycoprotein biomarker discovery from patient biofluids by antibody panel based N-glycan imaging. *Anal. Chem.* **91**, 8429–8435 (2019).

## Acknowledgements

This research was partially funded by the Biostatistics Shared Resource, Hollings Cancer Center, Medical University of South Carolina (P30 CA138313 to E.G.H.) and the National Institutes of Health (U01 CA242096 to R.R.D.).

## Author contributions

C.R.K.B., C.L., and R.R.D. conceived the experiments. C.L. supervised clinical sample collection and clinical data. C.R.K.B. conducted the experiments with consultation with R.R.D., P.M.A., and A.S.M. C.R.K.B. and E.G.H. analyzed the results. All authors reviewed the manuscript.

### Competing interests

The authors declare no competing interests.

### Additional information

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1038/s41598-022-25401-0>.

**Correspondence** and requests for materials should be addressed to R.R.D.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022