Evaluating the role of Magnetic Resonance Imaging post Neoadjuvant Therapy for Breast Cancer in the NEONAB trial.

Caitlin Murphy¹,²* Medical Oncologist
Violet Mukaro¹,² Clinical Trial Co-ordinator/Research Fellow
Robert Tobler¹ Clinical Trial Co-ordinator
Rebecca Asher⁵ Statistician
Emma Gibbs⁵ Statistician
Linda West³ Radiologist
Bruno Giuffre⁴ Radiologist
Sally Baron-Hay⁴ Medical Oncologist
Mustafa Khasraw¹,²,⁴,⁵ Medical Oncologist

1. University Hospital Geelong, Ryrie St, Geelong, VIC, Australia
2. Department of Medicine, Deakin University, Geelong, VIC, Australia
3. Lake Imaging, Geelong, VIC, Australia
4. Royal North Shore Hospital, Reserve Road, St Leonards, NSW, Australia
5. National Health and Medical Research Council Clinical Trials Centre, University of Sydney, NSW, Australia

Key words: breast cancer, MRI, neoadjuvant systemic therapy

*corresponding author:
Dr Caitlin Murphy
Medical Oncologist

This is the author manuscript accepted for publication and has undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1111/imj.13617
Abstract Word Count: 250

Manuscript Word Count: 2320

Acknowledgements & Funding:

This work was supported by an independent research grant from Specialised Therapeutics Australia (STA). Genomic Health conducted the Oncotype DX Breast Recurrence Score at no cost for this trial.
**Background**

Magnetic Resonance Imaging (MRI) accuracy after neoadjuvant systemic therapy (NST) for breast cancer varies according to hormone receptor (HR), human epidermal growth factor receptor type-2 (HER2) subtype and Ki67 proliferation index. Whether MRI accuracy varies by genomic signatures is unknown. We examined accuracy of MRI in the NEONAB trial (Clinicaltrials.gov #: NCT01830244).

**Methods**

Patients with stage II-III breast cancer received sequential epirubicin, cyclophosphamide and nab-paclitaxel and trastuzumab if HER2+. Sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV) were calculated to assess utility of preoperative MRI to predict pathological complete response (pCR). Bland-Altman plots were used to assess agreement between MRI and pathological assessment of residual disease.

**Results**

MRI correctly predicted pCR in 64.1% of the cohort. Sensitivity and specificity were 52% and 78%, PPV 73% and NPV 58%. MRI predicted pCR most accurately in HER2-positive patients; sensitivity 58%, specificity 100%, PPV 100% and NPV 38%. MRI had higher PPV and NPV in tumours with Ki67 ≥ 15% than tumours with Ki67 < 15%, 75% versus 50% and 57.5% versus 50% respectively. In this study, MRI underestimated residual tumour size by 1.65mm (limits of agreement: 43.07mm-39.77mm).

**Conclusions**

MRI appears more accurate for predicting pCR in HER2+ disease than other subtypes and in cancers with Ki67 ≥15% compared to those with Ki67 <15%. Accuracy of MRI in our HR+, RS ≥ 25 cohort is comparable to previous reports of unselected HR+ disease. MRI post-NST should be interpreted in conjunction with HER2 status and Ki67 index of the primary.
Figure 1: Bland-Altman plot for the entire cohort
Figure 2: Bland-Altman plot coded by disease subtype
Figure 3. Bland-Altman plots coded by Ki67 status.
Evaluating the Role of Magnetic Resonance Imaging post Neoadjuvant Therapy for Breast Cancer in the NEONAB trial.

Manuscript type: original article

Key words: breast cancer, MRI, neoadjuvant systemic therapy
INTRODUCTION

Accurate radiological assessment of response and residual disease is essential for achieving optimal surgical outcomes after neoadjuvant systemic therapy in breast cancer. Magnetic resonance imaging (MRI) is the most sensitive imaging modality used in the detection and staging of breast cancer(1). Guidelines suggest MRI can also be used to assess response to neoadjuvant therapies(2). Although several studies, including a meta-analysis and a systematic review, have shown MRI has greater accuracy compared to ultrasound (US), mammography and clinical examination, there is no consensus regarding its utility in the neoadjuvant setting(3-7). MRI still has the potential to both under and overestimate residual disease. More recent data suggests the accuracy of MRI varies according to pre-treatment breast cancer subtype defined by the presence or absence of the oestrogen receptor (ER), progesterone receptor (PR) and human epidermal growth factor receptor type 2 (HER2)(8-15). These data suggest MRI is most accurate in highly proliferative cancers, specifically triple negative breast cancer (TNBC) and HER2 positive, HR negative tumours. In addition, the accuracy of MRI may also vary depending on radiological phenotype and other biomarkers such as Ki67 proliferation index (12, 16-18).

Current evidence suggests MRI findings post neoadjuvant therapy should be interpreted in combination with baseline tumour characteristics(19). Existing biomarkers such as hormone receptor (HR) and HER2 profiles have proved useful but data regarding the utility of molecular biomarkers in this setting are scarce. Gene expression profiles are used to predict benefit from adjuvant chemotherapy after surgery for early breast cancer(20, 21). Oncotype DX recurrence score (RS) is FDA approved for this indication and its utility in the neoadjuvant setting is under investigation. Other gene expression profiles in use include MammaPrint and Prosigna(21, 22). MRI post neoadjuvant therapy is least accurate in patients with HR positive disease but it is unknown if a tumour’s gene expression profile may alter this interpretation. We report the accuracy of MRI in predicting pathological response to neoadjuvant epirubicin, cyclophosphamide and nab-paclitaxel in the primary cohort of NEONAB (Clinicaltrials.gov #: NCT01830244), a phase 2 study of tailored neoadjuvant therapy for breast cancer. We
include in our analysis a subset of patients with HR-positive, HER2-negative tumours and RS ≥25 in addition to subsets defined by HR, HER2 and Ki67.

MATERIALS AND METHODS
NEONAB was a single arm phase II clinical trial of tailored neoadjuvant chemotherapy in stage II-III breast cancer. This trial was conducted across three centres in two states; University Hospital Geelong and South West Health Care Warrnambool in Victoria and Royal North Shore Hospital in New South Wales.

Patients
Between April 2013 and December 2015, patients presenting to one of the recruitment sites with stage II or III, unilateral, histologically confirmed invasive breast cancer were eligible. Prior to chemotherapy all tumours were assessed for oestrogen receptor (ER) and progesterone receptor (PR) status, considered positive if ≥1% of tumour cells stained for ER and/or PR. HER2 status was also assessed by immunohistochemistry and in situ hybridization. Ki67 was assessed and recorded as equal to or greater than 15% or less than 15%. Patients with HER2-negative, HR-positive cancers had RS assay performed. Patients were analysed by three receptor subgroups; patients with HR-positive, HER2-negative, RS ≥25 tumours (hereafter designated ‘HR-positive’), patients with HER2-positive tumours and patients with ER, PR and HER2-negative tumours i.e. triple negative breast cancers. Patients with HR-positive, HER2-negative, RS <25 cancers were treated as a separate cohort and excluded from this study. A separate subgroup analysis was performed on the total cohort divided into two groups defined by Ki67 status; Ki67 ≥15% and Ki67 <15%. This study was reviewed and approved by the relevant ethics committees. All patients signed written informed consent.

Treatment
Patients received epirubicin 90 mg/m² and cyclophosphamide 600 mg/m² (EC) every 3 weeks for 12 weeks; followed by nab-paclitaxel 125 mg/m² days 1, 8, and 15 every 4 weeks for 12 weeks. Trastuzumab 8mg/kg loading dose then 6mg/kg every 3 weeks was added to nab-paclitaxel for HER2-positive patients and continued post-operatively for a total of 12 months.
Patients underwent mastectomy or breast conserving surgery and sentinel node biopsy or axillary dissection after completion of neoadjuvant therapy and reassessment of residual disease with MRI.

**Radiology & Pathology**

Magnetic resonance imaging of the breast was performed before and 2 weeks after primary systemic therapy. The MR imaging was performed on a Siemens 3T Verio Magnetom using dedicated software and a 16 channel AL dedicated breast coil except 3 studies which were performed on a Siemens 1.5T Aera with a 16 channel closed breast specific coil. Studies were performed utilising a dynamic contrast enhanced technique with axial acquisition, high resolution 3 plane reconstructions and 7.5ml Gadovist (gadobutrol) contrast agent injected manually according to a set timing protocol. The 3 North Shore Private Hospital studies utilised 10ml Multihance (gadobenate dimeglumine) in place of Gadovist. The sequences were: a fat suppressed sequence, a non-fat suppressed T2 sequence, the dynamic rapid sequence fat suppressed T1 sequence and a high resolution fat suppressed T1 sequence. For the 3 studies performed at North Shore Private Hospital this last sequence was acquired in the sagittal plane. The data was further analysed using a CAD program to allow multiplanar and maximum intensity projection analysis, measurement and dynamic contrast enhancement analysis. Due to variation in methods of volumetric measurement of tumour-size, we measured the tumour in 3 planes and used the longest diameter of tumour in millimetres as the size of the tumour on MRI. Multiple target lesions were recorded separately. Diffusion studies were performed but not used in this trial. Radiological response was determined by comparison of target lesions, a complete response was defined as resolution of all target lesions. All MR studies in Geelong were double-reported by two breast radiologists, one of which was the central trial radiologist, LEW, who has expertise in breast imaging and MRI. MR studies from other centres were all reviewed by LEW. Pathological response was assessed by experienced local pathologists. Pathological complete response was defined as no residual invasive disease in breast with residual *in situ* disease permitted, ypT0/is.

**Statistical Analysis**
To evaluate the role of MRI in predicting pCR, diagnostic measures; sensitivity, specificity, PPV and NPV were calculated. The analysis was conducted on the entire cohort and by each disease subtype (HR-positive, HER2-positive and TNBC) and Ki67 status.

The Bland-Altman method was used to quantify the agreement between the measurements of size of residual lesions from the MRI and the measurements of the size of the tumours in the resection specimen. The difference between the two measurements from each patient was plotted against the average of the two measurements and the corresponding mean difference and limits of agreement (mean ± 1.96SD) were calculated. As the group sizes were too small to construct Bland-Altman plots for each subgroup, the plots were reproduced coding the observations by subgroup to visually assess the differences.

RESULTS

Patients and materials

Patients’ baseline characteristics are described in Table 1. Forty patients were enrolled, one patient did not undergo radiological assessments therefore data were available in thirty-nine; 14 TNBC, 15 HER2-positive and 10 HR-positive patients. Ki67 was evaluable in 34 patients, 26 Ki67≥15% and 8 Ki67<15%. Patients with Ki67<15% were present in all receptor subgroups. Radiological complete response was observed in 15 patients (38%); 6 TNBC, 7 HER2-positive, 2 HR-positive and pCR in 21 patients (53%); 6 TNBC, 12 HER2-positive and 3 HR-positive. Four patients had radiological complete response on MRI with residual disease on pathological examination; 3 TNBC and 1 HR-positive (10%).

Accuracy of MRI for predicting pCR

MRI correctly classified 64.1% of patients in the primary cohort. MRI sensitivity and specificity were 52% and 78%, with corresponding PPV 73% and NPV 58%. MRI was most useful for predicting pCR in the HER2-positive subgroup where sensitivity was 58%, specificity 100%, PPV 100% and NPV 38% Table 2. In HR-positive and TNBC patients MRI was less useful with NPV 75% and 62.5% respectively and PPV of 50% in both subgroups Table 2. Analysis of the cohort
by Ki67 status showed MRI had higher PPV and NPV in tumours with Ki67 ≥ 15% than tumours with Ki67 < 15%, 75% versus 50% and 57.5% versus 50% respectively Table 3.

**Accuracy of MRI for predicting size of residual disease**

In the entire cohort the mean difference between MRI and pathology measurements was -1.65mm (SD 21.1); on average, MRI underestimated the residual tumour size. The lower and upper limits of agreement were -43.07mm and 39.77mm Figure 1. Figure 2 shows the same Bland-Altman plot as in but Figure 1 by disease subtype. From inspection, there does not appear to be any pattern in the difference in measurements across the disease subtypes. The mean differences were 1.67mm (SD 10) for HER2-positive, -4.72mm (SD 28.1) for HR-positive and -4.02mm (SD 26.3) for TNBC. Although this indicates MRI overestimated residual disease in HER2-positive cancers and underestimated disease in HR-positive and TNBC, these results should be interpreted with caution due to the small numbers in each of the subgroups.

Figure 3 shows the Bland-Altman plot coded by Ki67 status which does not show any pattern in the difference in measurements between groups. The mean difference for Ki67 <15% was 1.35mm (SD 9.5) and Ki67≥15% was -3.08mm (SD 25.9) but again, the small number of tumours in each group limits interpretation of these findings.
DISCUSSION

Assessment of residual disease after neoadjuvant therapy for breast cancer can involve several modalities including clinical examination, ultrasonography, mammography, functional imaging and MRI. MRI is deemed ‘optional’ by international guidelines, reflecting the lack of consensus regarding its utility in this setting(2). Despite the controversies, in the clinic, MRI is often used to re-evaluate tumours and informs surgical plan in almost fifty percent of patients(23). It is therefore imperative that the accuracy of this imaging tool is optimised and understood in context of individual tumour characteristics and biology. Our analysis of the agreement between MRI and pathologic measurements of residual disease demonstrated both under and overestimation occurred. Subgroup analyses were limited by small numbers but inspection of Bland-Altman plots did not reveal any biases. Overall, in keeping with previous studies, approximately 65% of our patients were correctly classified by MRI (3, 15, 24). Of those incorrectly classified, 10% of patients with radiological complete response had residual invasive tumour evident on pathologic examination. These patients are theoretically at risk of inadequate margins and the need for further surgery, however, overtreatment may also be a risk. When used for staging primary breast cancer, MRI often leads to more extensive surgery (1). In the neoadjuvant setting, McDermott et al found despite having a higher rate of pCR, patients who underwent post treatment MRI had a significantly lower rate of breast conserving surgery compared to those who did not. Twenty-five percent of our patients had residual lesions on MRI but achieved a pCR, overestimations that could have led to overtreatment.

It has become evident that the accuracy of post-neoadjuvant therapy MRI is improved when analysed in the context of pre-treatment breast cancer subtype(6, 9, 16, 25). Further improvements in accuracy may be possible if interpretation is coupled to a better understanding of tumour biology incorporating markers like Ki67 index, genomic profile and radiological phenotype. The findings of this study are limited by the small number of patients in each subgroup. More definitive results would require a larger sample size to enable robust statistical analyses of the differences between these subgroups. This limitation is present throughout the literature on this subject, a 2015 meta-analysis collated individual data from only three hundred patients across eight studies(3). Inclusion of MRI in large prospective neo-
adjunct trials, exemplified by the I-SPY analysis, will provide more definitive answers in future research (16). Nevertheless, our results add to the evidence that post-neoadjuvant therapy MRI is likely more useful for predicting pCR in HER2-positive tumours, with PPV 100% in this subgroup. Recent evidence also indicates MRI has higher accuracy in TNBC but this was not seen in our analysis (14, 16, 17, 24, 26). This is likely a reflection of the small number in our subgroups but examination of the individual cases also underlines the limitations of MRI in detecting microscopic residual disease. Notably, of the 3 TNBC patients with radiological complete response but pathological residual disease, 2 had near complete pathological response with very small foci of invasive cancer, 1.1mm and <5mm respectively. Our analysis measured the accuracy of MRI in predicting pCR whereas other studies have assessed MRI’s ability to predict a ‘clinically meaningful response’ defined as regression of tumour to a size permitting breast conserving surgery (16). Whilst MRI was not useful in predicting pCR in our TNBC cohort, its utility in predicting ‘clinically meaningful response’ was not assessed and cannot be excluded. In our cohort of patients with HR-positive, HER2-negative, Oncotype DX RS ≥ 25 tumours, the accuracy of MRI was similar to other studies’ HR-positive cohorts (16, 18, 24). Given our small sample size and the problem of cross trial comparison, these findings should be interpreted with caution, however, in this study, the addition of genomic information did not appear to increase the accuracy of MRI compared to other unselected HR-positive cohorts.

Previous reports have shown Ki67 index, independent of receptor subtype, can be used to identify tumours in which MRI has greater accuracy to predict pCR (15, 17, 18). The proportion of cells staining for Ki67 provides an indication of a tumour’s rate of proliferation. In HR-positive cancers, a Ki67 index < 15% is occasionally used as a surrogate for molecularly defined luminal A subtype (27, 28). In our analysis, MRI appeared to be more accurate for predicting pCR in tumours with a higher proliferative index; PPV 75% for Ki67 ≥ 15% versus 50% for Ki67 < 15%. This is a lower cut off than previous studies but still indicates the rate of proliferation of the primary tumour maybe relevant when interpreting MRI findings post-neoadjuvant therapy.

There is no standard approach to the assessment of residual breast cancer after neoadjuvant treatment but the use of MRI can be refined by incorporating the individual patient’s tumour subtype and biology into the analysis. Recent studies suggest measures of tumour volume and
contrast kinetic thresholds can also be used to increase the accuracy of post-NST MRI over traditional measurements of tumour diameter(29, 30). We propose the development of algorithms that incorporate receptor subtype, tumour proliferation and radiological phenotype into the interpretation of post-neoadjuvant therapy MRI(16, 31). Our findings support the use of HER2 status and Ki67 index in such an assessment.

CONCLUSIONS

The role of MRI in the assessment of residual breast cancer post-neoadjuvant therapy is evolving. Our study suggests MRI maybe more accurate for predicting pCR in HER2+ disease than other subtypes and in cancers with higher Ki67 proliferative index. Whether more detailed knowledge of tumour biology, such as that gleaned from genomic biomarkers, could further inform the assessment of residual disease is a question for future research.

ACKNOWLEDGEMENTS

This work was supported by an independent research grant from Specialised Therapeutics Australia (STA). Genomic Health conducted the Oncotype DX Breast Recurrence Score at no cost for this trial.
REFERENCES


Figure Legend

Figure 1. Bland-Altman plot for the entire cohort: agreement between MRI & pathological measurement of residual disease

Figure 2. Bland-Altman plot coded by disease subtype

Figure 3. Bland-Altman plot coded by Ki67 level

Table 1. Summary of Patient Baseline Characteristics

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>n</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Age, years</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Median</td>
<td>51</td>
<td></td>
</tr>
<tr>
<td>Range (35–77 years)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Subtype</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HER2+</td>
<td>15</td>
<td>38</td>
</tr>
<tr>
<td>HR+, HER2-, RS ≥ 25</td>
<td>10</td>
<td>26</td>
</tr>
<tr>
<td>TNBC</td>
<td>14</td>
<td>36</td>
</tr>
<tr>
<td><strong>Ki67</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>≥ 15%</td>
<td>26</td>
<td>67</td>
</tr>
<tr>
<td>&lt; 15%</td>
<td>8</td>
<td>21</td>
</tr>
<tr>
<td>unknown</td>
<td>5</td>
<td>13</td>
</tr>
<tr>
<td><strong>Multifocal</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>21</td>
<td>54</td>
</tr>
<tr>
<td>No</td>
<td>18</td>
<td>46</td>
</tr>
<tr>
<td><strong>ECOG</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>38</td>
<td>98</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td><strong>Ethnic group</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Caucasian</td>
<td>36</td>
<td>93</td>
</tr>
<tr>
<td>Asian</td>
<td>3</td>
<td>7</td>
</tr>
<tr>
<td><strong>Tumour stage</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2a</td>
<td>10</td>
<td>26</td>
</tr>
<tr>
<td>2b</td>
<td>11</td>
<td>28</td>
</tr>
<tr>
<td>3a</td>
<td>14</td>
<td>36</td>
</tr>
<tr>
<td>3b</td>
<td>3</td>
<td>8</td>
</tr>
<tr>
<td>3c</td>
<td>1</td>
<td>2</td>
</tr>
</tbody>
</table>

* HR status deemed negative if <1% nuclei staining, + American Joint Committee on Cancer Breast Cancer Staging v.7, †Multifocal as per MRI Assessment

ECOG, Eastern Cooperative Oncology Group; HER2, human epidermal growth factor receptor type 2
### Table 2. Accuracy of MRI for predicting pCR according to subtype

<table>
<thead>
<tr>
<th></th>
<th>HER2-positive</th>
<th></th>
<th>TNBC</th>
<th></th>
<th>HR-positive, RS ≥ 25</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>95% CI</td>
<td>95% CI</td>
<td></td>
<td>95% CI</td>
<td></td>
<td>95% CI</td>
</tr>
<tr>
<td>Sensitivity</td>
<td>58.3%</td>
<td>(27.7, 84.8)</td>
<td>50%</td>
<td>(11.8, 88.2)</td>
<td>33.3%</td>
<td>(0.84, 90.6)</td>
</tr>
<tr>
<td>Specificity</td>
<td>100%</td>
<td>(29.2, 100)</td>
<td>62.5%</td>
<td>(24.5, 91.5)</td>
<td>85.7%</td>
<td>(42.1, 99.6)</td>
</tr>
<tr>
<td>Positive predictive value</td>
<td>100%</td>
<td>(59, 100)</td>
<td>50%</td>
<td>(11.8, 88.2)</td>
<td>50%</td>
<td>(1.26, 98.7)</td>
</tr>
<tr>
<td>Negative predictive value</td>
<td>37.5%</td>
<td>(8.5, 75.5)</td>
<td>62.5%</td>
<td>(24.5, 91.5)</td>
<td>75%</td>
<td>(34.9, 96.8)</td>
</tr>
</tbody>
</table>

### Table 3. Accuracy of MRI for predicting pCR according to Ki67 status

<table>
<thead>
<tr>
<th></th>
<th>Ki67 &lt; 15%</th>
<th></th>
<th>Ki67 ≥ 15%</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>95% CI</td>
<td>95% CI</td>
<td>95% CI</td>
<td>95% CI</td>
</tr>
<tr>
<td>Sensitivity</td>
<td>25%</td>
<td>(0.63, 80.6)</td>
<td>60%</td>
<td>(32.3, 83.7)</td>
</tr>
<tr>
<td>Specificity</td>
<td>75%</td>
<td>(19.4, 99.4)</td>
<td>72.7%</td>
<td>(39, 94)</td>
</tr>
<tr>
<td>Positive predictive value</td>
<td>50%</td>
<td>(1.26, 98.7)</td>
<td>75%</td>
<td>(42.8, 94.5)</td>
</tr>
<tr>
<td>Negative predictive value</td>
<td>50%</td>
<td>(11.8, 88.2)</td>
<td>57.1%</td>
<td>(28.9, 82.3)</td>
</tr>
</tbody>
</table>
Minerva Access is the Institutional Repository of The University of Melbourne

Author/s:
Murphy, C; Mukaro, V; Tobler, R; Asher, R; Gibbs, E; West, L; Giuffre, B; Baron-Hay, S; Khasraw, M

Title:
Evaluating the role of magnetic resonance imaging post-neoadjuvant therapy for breast cancer in the NEONAB trial

Date:
2018-06-01

Citation:

Persistent Link:
http://hdl.handle.net/11343/285113