Clinico-pathologic relationships with Ki67 and its change with short-term aromatase inhibitor treatment in primary ER+ breast cancer: further results from the POETIC trial (CRUK/07/015)

Judith M Bliss[1], Holly Tovey[1], Abigail Evans[2], Chris Holcombe[3], Kieran Horgan[4], Elizabeth Mallon[5], Raghavan Vidya[6], Anthony Skene[7], Andrew Dodson[8], Margaret Hills[9], Simone Detre[9], Lila Zabaglo[9], Jane Banerji[1], Lucy Kilburn[1], James P Morden[1†], John F R Robertson[10], Ian Smith[11] and Mitch Dowsett[9] on behalf of the POETIC Trialists.

[1]*Clinical Trials and Statistics Unit (ICR-CTSU), The Institute of Cancer Research, London, UK;*

[2]*Poole Hospital, Poole, UK;*

[3]*Royal Liverpool University Hospital, Liverpool, UK;*

[4]*Dept. of Breast Surgery, St James's University Hospital, Leeds, UK;*

[5]*Western Infirmary, Glasgow, UK;*

[6] *Royal Wolverhampton NHS Trust, Wolverhampton, UK;*

[7]*Royal Bournemouth Hospital, Bournemouth, UK;*

[8]*UK NEQAS for Immunocytochemistry and In-Situ Hybridisation, London, UK;*

[9]*Ralph Lauren Centre for Breast Cancer Research, Royal Marsden Hospital, and Breast Cancer Now Centre, The Institute of Cancer Research, London, UK;*

[10]*University of Nottingham, Royal Derby Hospital, Derby, UK;*

[11] *Breast Unit, Royal Marsden Hospital, London, UK.*

† Deceased

All correspondence to Judith M Bliss

Email: Judith.bliss@icr.ac.uk

Phone: -+44-208-722-4297

Key words

Ki67; aromatase inhibitor; primary breast cancer

Abstract:

Purpose

Ki67 assessed at diagnosis ($Ki67_{baseline}$) is an important prognostic factor in primary oestrogen receptor positive (ER+) breast cancer. Proportional change in Ki67 after 2 weeks' ($\Delta Ki67_{2week}$) is associated with clinical benefit from endocrine therapies and residual Ki67 ($Ki67_{2week}$) with recurrence-free-survival. The aim was to define the association between $Ki67_{baseline}$, and after aromatase inhibitor (AI) exposure $\Delta Ki67_{2week}$ and $Ki67_{2week}$ with key prognostic and biologic factors utilising data from the POETIC study.

Patients and Methods

In POETIC 4480 postmenopausal patients with primary ER and/or PgR+ breast cancer were randomised 2:1 to 2 weeks' pre-surgical AI (anastrozole or letrozole) or no pre-surgical treatment (control). Ki67 was measured centrally in core-cut biopsies taken prior to AI and in core cuts or the excision biopsy at surgery. Relationships between the Ki67 and biologic factors were explored using linear regression.

Results

Established associations of $Ki67_{baseline}$ with biologic factors including PgR status, tumour grade, tumour size, histological subtype, nodal status, and vascular invasion were confirmed in the HER2- subpopulation. In the HER2+ subpopulation only grade and tumour size were significantly associated with $Ki67_{baseline}$. In control group $Ki67_{2week}$ was 18% lower than $Ki67_{baseline}$ ($p<0.001$) when $Ki67_{2week}$ was measured in excision biopsies but not when measured in core-cuts. Median suppression by AIs ($\Delta Ki67_{2week}$) was 79.3% (IQR: -89.9 - -54.6) and 53.7% (IQR: -78.9 - -21.1) for HER2-ve and HER2+ve cases respectively. Significantly less

suppression occurred in PgR- vs PgR+ and HER2+ vs HER2- tumours which remained apparent after adjustment for 2 week sample type.

Conclusions

The magnitude of this study allowed characterisation of relationships between $Ki67_{baseline}$, $\Delta Ki67_{2week}$ and $Ki67_{2week}$ with high degrees of confidence providing a reference source for other studies. Lower values of Ki67 occur when measured on excision biopsies and could lead to apparent but artefactual decreases in Ki67: this should be considered when either $\Delta Ki67_{2week}$ or $Ki67_{2week}$ are used in routine clinical practice to aid treatment decisions or in clinical trials assessing new drug therapies.

Background

The nuclear proliferation marker, Ki67, is measured in many malignancies including primary breast cancer(1). International efforts have shown progress in standardising its measurement such that its value for aiding clinical practise may be realised(2). Ki67 analysis in primary breast cancer is known to be a prognostic marker for the >80% of patients whose breast cancers are ER-positive(3) (ER+). Such an example is its licencing as a companion diagnostic for abemaciclib in the US(4). Yet, where an individual patient's Ki67 measurement sits within the distribution of the patient population with similar clinical and pathological characteristics is less well described. For example, how unusual is a Ki67 measurement >20% for a patient with lobular cancer, especially if this is residually high after short term exposure to an aromatase inhibitor (AI)? Optimising prognostic tools, which incorporate such biomarker results and illustrate the distribution of biomarkers according to classical clinical-pathological factors is therefore a high priority so that risk-based decisions can be estimated with confidence for the individual patient.

Short-term presurgical treatment of patients with primary breast cancer, particularly those with ER+ disease, has become popular to gain insights into drug activity but also for identifying groups of patients who may be candidates for response-adapted therapy(5). Ki67 is the primary endpoint for the large majority of these studies. The limited size of almost all these studies does not permit confident assessment of the relationship with clinicopathologic factors and commonly measured biomarkers or the impact of such on the pharmacologic effectiveness of presurgical therapy on Ki67.

In the large majority of primary ER+ breast cancer Ki67 is markedly suppressed by just 2 weeks' endocrine therapy(6). We and others have shown that the degree of suppression ($\Delta Ki67_{2week}$) is predictive of response to prolonged endocrine therapy (3, 7). For example, in the neoadjuvant IMPACT trial, the mean suppression of Ki67 by anastrozole was significantly greater than that by tamoxifen or the combination of anastrozole and tamoxifen at both 2 and 12 weeks(3). Similarly, in the parallel ATAC adjuvant trial, anastrozole reduced recurrence to a greater extent than tamoxifen or the combination(8). Given that the mean Ki67 suppression by each of the patient groups in IMPACT was only slightly more at 12 than at 2 weeks, and that 2 weeks is a common duration for the period between breast cancer diagnosis and surgery, the measurement of this biomarker change within what has become known as the presurgical "window of opportunity" has become a primary endpoint in pre-surgical studies of novel agents. The measurement of Ki67 after such presurgical treatment also has the potential to be used to triage patients away from endocrine treatment alone in the case of sub-optimal response(9). Of particular note regarding prognosis, the absolute level of Ki67 expression at 2 weeks ($Ki67_{2week}$) was shown to be more strongly related to recurrence-free survival than pre-treatment levels ($Ki67_{baseline}$)(10). This seems likely to be due to $Ki67_{2week}$ integrating the intrinsic prognostic value of $Ki67_{baseline}$ and the improvement in prognosis that is reflected by $\Delta Ki67_{2week}$. Some investigators advocate the estimation of complete cell cycle arrest (Ki67</=2.7%) for identifying patients with the best prognosis on endocrine therapy(11).

Evidence to inform whether the gain in prognostic insights from measuring $Ki67_{2week}$ is sufficient to merit routine administration of endocrine therapy prior to surgery has been

recently reported in the PeriOperative Endocrine Therapy for Individualised Care (POETIC) trial (**ISRCTN:** 63882543, CRUK/07/015)(12). This trial randomised over 4,400 UK postmenopausal women with hormone sensitive primary breast cancer to receive a non-steroidal AI (letrozole or anastrozole) for 2 weeks prior to and after surgery or no perioperative endocrine treatment (2:1). The study did not show that perioperative endocrine treatment improved long term outcomes but did show that $Ki67_{2week}$ <10% was associated with low risk of recurrence. Ki67 analyses from the trial used a scoring method that has formed the basis for international standardisation(13). We report here the relationship between $Ki67_{baseline,}$ $Ki67_{2week}$ and $\Delta Ki67_{2week}$ with key prognostic and biologic factors. While we have shown that the large majority of patients show a reduction in Ki67 after 2 weeks' treatment with an aromatase inhibitor, the degree of change differs markedly between patients. It is known that suppression is greater in tumours with high ER and PgR and in those negative for HER2(14) but the degree to which these relationships are independent of one another and of commonly measured clinicopathologic features could not be established in the modest sized studies to date. The number of patients included in POETIC enabled to address those issues. We also were able to determine if differences in Ki67 levels according to biopsy type were sufficiently substantial to impact on prognostic estimates, and to describe extent of Ki67 suppression achieved according to choice of AI, issues for which there was very limited information to date.

Methods

The primary clinical results and detailed methods for POETIC have already been reported(12). Details included here are those pertinent to the current report.

Patients and Procedures

POETIC was a phase III, multicentre, randomised trial for postmenopausal women with ER or PgR positive invasive breast cancer. Women were randomized (2:1 allocation ratio) to perioperative therapy with a non-steroidal AI (POAI), anastrozole (1mg/day) or letrozole (2.5mg/day) (AI choice determined by centre policy) for two weeks before and two weeks after surgery or no perioperative therapy (control). Subsequent therapy was according to local standard of care. Ki67 was evaluated as a biomarker in relation to its effect on predicting disease outcomes and as a secondary endpoint to assess changes between baseline and surgery. Full details of the design and statistical analysis methods of the main study are available in the main clinical paper(12).

Patients provided written consent for the use of core-cut biopsies taken at diagnosis or, if material was not available at diagnosis, for the taking of a core-cut for the purposes of the trial. Investigators were encouraged to take a further core-cut biopsy at the time of surgery but could alternatively provide a representative paraffin-embedded block. Provision of tissue sections was also acceptable at both baseline and surgery. All samples were fixed in formalin prior to paraffin embedding.

Ki67 methodology

Ki67 was assessed largely according to the method described in Zabaglo et al(15) that formed the basis for that method validated by the International Ki67 in Breast Cancer Working Group(13). Ki67 was visualized immunohistochemically using the MIB-1 monoclonal antibody (Dako UK Ltd) at a dilution of 1:50, staining was performed on an automated staining platform (Dako Autostainer, Dako UK Ltd). For scoring, all stained and unstained invasive tumour nuclei were counted in at least 5 high-power fields; the Ki67 staining index was calculated as the total number of stained nuclei counted/total number of all invasive nuclei counted. Only scores from samples in which there were at least 200 invasive cells in total were accepted. QCs consisting of a TMA of at least six cores in duplicate were included in each batch and batches were only accepted if the scores met specified criteria of acceptance. Paired baseline and surgical samples were stained in the same batch in almost all cases. Scoring was carried out centrally by a team of nine competency-approved technical staff who sought histopathologic advice as necessary and practised comparative quality assurance tests throughout the study; 86% of the scoring was conducted by 4 of the staff. Technicians scoring Ki67 were blinded to the treatment allocation. Fewer surgical samples from control patients were analysed because little extra value was expected from multiple samples in the absence of treatment. Initially all surgical samples were analysed but from early 2013 a subset of one third of remaining control patients were selected at random for analysis while all patients in the treatment group were analysed, this led to approximately 7/9 surgical samples from the whole trial being analysed.

Statistical analyses

Medians and interquartile ranges were used to summarise $Ki67_{baseline}$, $Ki67_{2week}$ and $\Delta Ki67_{2week}$. $\Delta Ki67_{2week}$ was calculated as $100*((Ki67_{2week}+0.1)$ -

177 ($Ki67_{baseline}$+0.1))/($Ki67_{baseline}$+0.1). The non-parametric sign-test was used to test whether
178 $\Delta Ki67_{2week}$ was different from zero in control group patients.
179
180 The relationship between each of $Ki67_{baseline}$, $Ki67_{2week}$ and $\Delta Ki67_{2week}$ and key prognostic and
181 biologic factors was assessed using linear regression. For $Ki67_{baseline}$ and $Ki67_{2week}$ an outcome
182 of ln(Ki67 + 0.1) was used. For categorical variables, the model coefficient β indicates the
183 mean difference in ln(Ki67 + 0.1) between a designated group and the reference group
184 (indicated by β=0). For continuous variables β indicates the mean difference in ln(Ki67 + 0.1)
185 per unit increase. For models of $\Delta Ki67_{2week}$ an outcome of log-fold change in Ki67 was used,
186 defined as ln(($Ki67_{2week}$+0.1)/(ln($Ki67_{baseline}$+0.1)). A positive value of β indicates a smaller drop
187 in Ki67 from baseline to 2 weeks for the designated group compared to the reference group.
188
189 Univariable models were fitted containing only the variable of interest. P-values given are for
190 a likelihood ratio test comparing this model with a null model containing no variables.
191 Multivariable models were fitted containing all known prognostic variables listed in the same
192 model. P-values given are from a likelihood ratio test comparing this model with a model
193 containing all variables except the one of interest. The multivariable models for $Ki67_{baseline}$ and
194 $Ki67_{2week}$ includes all factors listed. Multivariable models for $\Delta Ki67_{2week}$ additionally include
195 $Ki67_{baseline}$, and were subsequently adjusted for type of AI (letrozole vs anastrozole) and
196 surgical sample type (excision vs core-cut). Models were also repeated only including
197 variables identified as significant in univariable analyses but parameter estimates were not
198 significantly affected so full models are presented for completeness. No adjustment was
199 made to p values for multiplicity but for each multivariable model the adjusted critical value
200 for each term using a Benjamini-Hochberg correction is presented to assist interpretation.

Using this procedure, p-values are ranked and adjusted critical values are calculated based on the rank. P-values are compared to the adjusted critical values, the largest p-value which is smaller than its associated critical value and any p-values smaller than this are considered significant.

Analyses were based on the snapshot of the clinical data taken on 06/02/2018, consistent with the main clinical results paper. All analyses were performed using STATA 15.

Results

Of the 4480 women (POAI (n=2976); control (n=1504)) who entered POETIC, $Ki67_{baseline}$ data was available for 2610 (87.7%) and 1303 (86.6%) respectively; $Ki67_{2week}$ from 2551 (85.7%) and 692 (46.0%); and paired samples to allow calculation of $\Delta Ki67_{2week}$ from 2528 (84.9%) and 678 (45.1%), respectively. Figure 1 shows a consort diagram showing reasons for non-availability of data.

Ki67 assessed at diagnosis ($Ki67_{baseline}$)

In this population of 3913 women a highly skewed distribution of $Ki67_{baseline}$ was observed which could be normalised via a logarithmic transformation (Supplementary Figure S1(a) & S1(b)). The median $Ki67_{baseline}$ value was 15.2%; with an IQR of 8.6% to 26.0%; 69.2% of values were above the commonly used threshold of 10%. When considering relationships with common clinic-pathological factors clear evidence was observed of an association with HER2 status (median (IQR) HER2-ve 14.3 (8.2 - 24.6); HER2+ve 26.6 (17.0 - 37.4); Supplementary Figure S1(c)). Given this finding and the different treatment pathways followed by HER2-ve and HER2+ve patients all subsequent results are shown for the sub-populations split according to HER2 status, as shown for clinic-pathological factors (Figures 2(a), 3(a) and Supplementary Figure S2(a)).

Within the HER2-ve sub-population (n=3445) and in univariate analyses a relationship was seen between $Ki67_{baseline}$ and each of the clinic-pathological factors aside from age (Figure 3(a) & Table 1). In multivariable analyses a statistically significant association remained for all of

these factors (Table 1). This held regardless of whether tumour size was treated as continuous or categorical (additional data not shown).

(Insert table 1 here)

Within the smaller HER2+ve sub-population (n=413) in univariate analyses a relationship was observed between $Ki67_{baseline}$ and grade which remained significant in multivariable analysis. There was also a significant association between $Ki67_{baseline}$ and tumour size treated as ordinal or continuous but this did not remain significant in the multivariable analysis (Figure 3(a) & Table 1).

$Ki67_{2week}$ Control group

As expected the logarithmic distribution shown for $Ki67_{baseline}$ was maintained at 2 weeks for patients who were allocated not to receive perioperative AI therapy in both the HER2-ve and HER2+ve subgroups (Figure 2(b) & Supplementary Figure S2(b)). The median $Ki67_{2week}$ was 13.1% and 23.6% for HER2-ve and HER2+ve patients respectively.

$\Delta Ki67_{2week}$ Control group

In the control group for patients with HER2-ve tumours, there was a median fall of 14.6% (IQR: -40.8 – 18.3) in Ki67 ($\Delta Ki67_{2week}$) (Figure 2(C) & Supplementary Figure S2(C)); 100 patients (16.8%) had $Ki67_{baseline} \geq 10\%$ which dropped to <10% at 2 weeks. In multivariable analyses $\Delta Ki67_{2week}$ was associated with $Ki67_{baseline}$ and tumour grade (Supplementary Table S1). It was also associated with continuous tumour size but this was not significant in

multivariable analyses using the Benjamini Hochberg adjusted critical values and was not significant when categorised.

In HER2+ve patients, there was a median fall of 12.4% (IQR: -31.7 -7.1) in Ki67; 5 patients (7.1%) had $Ki67_{baseline}$≥10% which dropped to <10% at 2 weeks. In univariable analyses, $\Delta Ki67_{2week}$ was associated with $Ki67_{baseline}$ but this was not significant in multivariable analyses after Benjamini Hochberg adjustment to critical values. $\Delta Ki67_{2week}$ was not associated with any other clinic-pathological factors in this population (Supplementary Table S1).

In order to understand this apparent, potentially artefactual change, analyses of change in Ki67 were explored according to type of sample from which $Ki67_{2week}$ had been calculated. As previously alluded to in the main trial results paper(12) analysis of 679 control group patients with paired samples available (ie $Ki67_{baseline}$ and $Ki67_{2week}$) analyses indicated that where $Ki67_{2week}$ was taken from a core-cut sample the median proportional reduction was −4·1% (IQR −27·8 to 34·8), compared to −17·7% (IQR −44·2 to 12·7) when a surgical resection sample was used. This significant association between sample type and $\Delta Ki67_{2week}$ was observed in the sub-population of patients with HER2-ve tumours (Supplementary Figure S3(a)). However, adjusting for sample type in the multivariable model did not materially impact the effect of the clinic-pathological features on $\Delta Ki67_{2week}$ (Supplementary Table S2). No significant association was observed between sample type and $\Delta Ki67_{2week}$ in patients with HER2+ve tumours.

$Ki67_{2week}$ POAI group

Following this short exposure to AI treatment the distribution of $Ki67_{2week}$ looked very different to that observed at baseline (Figure 2(d) & Supplementary Figure S2(d)) and the level of Ki67 expression was significantly different. The median was 2.5% (IQR: 1.1 – 6.5) and 10.3% (IQR: 4.1 – 21.2) in HER2-ve and HER2+ve patients respectively with 17.5% of HER2-ve patients and 51.8% of HER2+ve patients now having $Ki67_{2week}$ above 10%.

In the HER2-ve cohort, the significant univariate relationships seen between grade, tumour size, histologic type (lobular vs ductal), nodal involvement, vascular invasion and $Ki67_{baseline}$ were all observed with $Ki67_{2week}$ (all $p<0.001$). Effect sizes were similar to those observed with $Ki67_{baseline}$(Figure 3(b) and Table 2). PgR negativity was also related to higher $Ki67_{2week}$ and this relationship was stronger than for $Ki67_{baseline}$. Similarly, the contribution of PgR status to the multivariable model was stronger with $Ki67_{2week}$ than with $Ki67_{baseline}$ (Table 2). Tumour size did not remain significant in the multivariable model while all other relationships were similar for Ki67 assessed at either time-point. This held regardless of whether baseline or surgical grade was used and whether tumour size was considered as categorical or continuous (additional data not shown).

(Insert table 2 here)

In the HER2+ve cohort significant univariate associations were observed between $Ki67_{2week}$ and PgR status and grade, both of which remain significant in multivariable analysis (Figure 3(b) & Table 2). There was also a significant association between $Ki67_{2week}$ and tumour size but this only remained significant in multivariable analysis when size was treated as categorical.

$\Delta Ki67_{2week}$ POAI group

The median suppression of Ki67 in relation to baseline was 79.3% (IQR: -89.9 - -54.6) and 53.7% (IQR: -78.9 - -21.1) for HER2-ve and HER2+ve cases respectively. The distribution of Ki67 change was logarithmic as shown in Figure 2(e). Only 11.0% of patients did not show a reduction of at least 10% (allowing for variability) after 2 weeks POAI treatment compared to at baseline (10.0% & 18.8% for HER2-ve and HER2+ve respectively).

For both the HER2-ve and HER2+ve cohorts no significant univariable or multivariable relationship with $\Delta Ki67_{2week}$ was observed for tumour size, nodal involvement, histologic subtype or vascular invasion (Figure 3(c) & Table 3). However, PgR status and tumour grade were significantly associated with $\Delta Ki67_{2week}$ and remained significant in multivariable analysis. Higher $Ki67_{baseline}$ was also significantly associated with a higher proportional change in Ki67 in both cohorts (Table 3). This did not alter following adjustment for sample type in the HER2-ve cohort (Supplementary Table S3).

(Insert table 3 here)

We also explored in what is a non-randomised comparison whether each of the AIs received was differentially associated with $\Delta Ki67_{2week}$. Of patients with paired $Ki67_{baseline}$ and $Ki67_{2week}$; 839 (33%) patients were known to have received anastrozole and 1689 (67%) letrozole. Although considerable change in Ki67 was seen for each AI the median suppression was observed to be slightly less with anastrozole than letrozole (75.6% vs 80.6%, $p<0.001$, respectively Supplementary Figure S3(b)) in HER2-ve patients but not in HER2+ve patients

where median suppression did not differ by type of AI (56.6 vs. 50.6 respectively, p=0.791). Upon further exploration, the association remained after adjustment for sample type but the difference appeared to be evident only within excision samples but not core-cuts (Supplemenrary Figure 3(b)). Inclusion of AI and sample type in multivariable models did not impact the association with other baseline characteristics (Supplementary Table S3).

Complete Cell Cycle Arrest (CCCA), AI group

Suppression of Ki67 to ≤2.7% has been used to define CCCA. Supplementary Table S4 shows the frequency of CCCA according to the choice of AI and surgical sample type by HER2 status. Similar to analyses of $\Delta Ki67_{2week}$, in HER2-ve patients there was a greater likelihood of recording CCCA if the surgical sample was an excision rather than a core-cut (55.4% vs 44.2%, respectively; p<0.001). There was no difference in the frequency of CCCA according to AI used for core-cuts at 2 weeks (anastrozole 44.8%, letrozole 44.1%). In patients with an excision at 2 weeks, CCCA was significantly less frequent with anastrozole than with letrozole (49.7% vs 59.1%, respectively; p<0.001). No differences were observed by AI or sample type in the HER2+ve population but sample size in this sub-cohort is restrictive.

Discussion

Ki67 is the most widely measured marker of proliferation in primary breast cancer. While there have been many reports of the association of Ki67 with clinico-pathologic parameters in breast cancer there have been very few large studies that focussed entirely on ER+ disease where its measurement has greatest impact. The magnitude of our study enabled us not only to confirm previously hypothesised relationships but also to quantify the degree of

independence of each relationship within a multivariable context. It also allowed us to discover with high levels of confidence other relationships that have remained either unknown or less clear in earlier studies. We were able to do so for 3 measurements with distinct clinical relationships with clinical outcome: (i) $Ki67_{baseline}$ which is related to prognosis in the absence of treatment(1); (ii) $Ki67_{2week}$ which relates to the prognosis of patients on adjuvant endocrine therapy otherwise known as residual risk(10, 12); (iii) $\Delta Ki67_{2week}$ which reflects the antiproliferative impact of estrogen deprivation with an AI and has been shown to predict the relative benefit of endocrine therapies given as adjuvant treatment(3, 7). While $Ki67_{baseline}$ is often measured in clinical practise for its prognostic information it is not currently considered to have sufficient clinical utility for that purpose to be mandated by authoritative guidelines. However, FDA has recently approved the use of the CDK4/6 inhibitor abemaciclib for use in early breast cancer patients with one of the conditions being that $Ki67_{baseline}$ is >20%. This enhances the relevance of the data we present here from our large cohort of baseline samples.

Other strengths of the study include the central analysis of Ki67 using a scoring method that was marginally modified prior to its analytical validation by the International Ki67 in Breast Cancer Working Group(13). Several scorers were involved with a rigorous internal QC program. The involvement of a large number of hospital sites with variability in collection and fixation procedures might be considered a weakness. On the other hand the authors view the large number of sites as a strength in that it enables interpretation within the context of routine conduct of Ki67 measurements and allowed the characterisation of an important difference in scores between biopsy types. The study involved only postmenopausal patients and may not be representative of premenopausal cases.

Relationships of $Ki67_{baseline}$ in an ER+ population with PgR and HER2 status are well known. We were also able to confirm results from our earlier much smaller patient series(16) that HER2 impedes the antiproliferative response (from approximately 80% to 50%) to AI but does not preclude it. Ellis et al similarly reported that Ki67 suppression by AIs was less in HER2+ cases(17). The size of the POETIC trial allows analyses to identify the molecular features that are associated with antiproliferative response or not within the HER2+ population that makes up only about 10% of ER+ breast cancer(18).

There was less proportional suppression of Ki67 in PgR- than PgR+ cases leading to the relative difference in $Ki67_{baseline}$ between these subsets also being seen at 2 weeks. This is consistent with our earlier report(14) and that of others and suggests that AIs may have greater relative benefit in PgR+ than PgR- patients. This has not been detected directly in adjuvant trials but the data from those trials relates to the comparative benefit from AIs versus tamoxifen(19). The lower value of $Ki67_{2week}$ in the PgR+ group is consistent with the substantially better prognosis of such patients on endocrine therapy(20, 21, 22). In contrast, lobular cancers showed a similar suppression of Ki67 compared to ductal cancers suggesting a similar biological response to AIs but better prognosis because of their lower $Ki67_{baseline}$ and $Ki67_{2week}$.

The poorer $\Delta Ki67_{2week}$ in higher grade tumours or those with high Ki67, similarly to that in PgR- and HER2+ tumours indicates that those with biologically more aggressive disease but not higher stage disease (cf the data on tumour size and nodal status) have a poorer biologic response to estrogen deprivation. In our report(23) of whole exome sequencing in samples

398 from POETIC those cases with high mutational load and/or TP53 mutation also had lower
399 $\Delta Ki67_{2week}$ and similarly would be enriched for cases with more aggressive disease.

400

401 While others have reported lower Ki67 values in excisions versus core-cuts of breast
402 cancers(24, 25) this has not been universally reported(26). The lack of difference between
403 Ki67 measured at baseline and 2-week in controls where core-cut biopsies were available
404 supports there being little overall impact of the procedures in the trial up to the point of
405 taking the 2-week sample. There may be a number of explanations for the finding that there
406 was a significant difference between Ki67 measured at baseline and then at 2 weeks in
407 controls where the 2-week sample was taken from the surgical resection specimen. Nuclear
408 integrity may be poorly preserved in routinely fixed excision specimens due to a delay in
409 formalin reaching the centre of the excision specimens where the tumour is situated, usually
410 surrounded by a margin of normal tissue which is variable from specimen to specimen. This
411 problem does not occur in core-cuts because of their smaller size. Also, under ultrasound
412 biopsy the needle is placed right at the edge of the tumour or even in it and therefore there
413 is much more rapid fixation of the tumour. Further explanation may be that core biopsies are
414 placed in fixative much more swiftly, indeed almost immediately and the tissue is therefore
415 not exposed to any ischaemic warm time. In contrast wide local excision specimens,
416 mastectomy and mastectomy and en-bloc axillary clearances have on average a greater warm
417 ischaemic time due to the increasing duration of surgical time and ischaemia of the tissues
418 resected. It is also possible that core-cuts may tend to sample more proliferative areas of the
419 tumour although that seems unlikely given that higher staining areas of Ki67 are more
420 commonly found at the tumour edge. Our scoring method involved selection of areas for

scoring to represent any heterogeneity in staining but it cannot be completely ruled out that this may also have contributed to the lower values in excisions. Whatever the cause(s) the relative difference of approximately 20% is important to consider and is highly preferable to avoid in pre-surgical studies. In the absence of a control arm a pre-surgical study in which excision specimens are used as the on-treatment sample may artifactually enhance the apparent antiproliferative impact of a treatment. For example, in our study, in the POAI group the median percentage change of Ki67 was -72.6% when the surgical sample was a core-cut compared to -79.3% in excisions. However, as a difference had been observed in the control arm, $Ki67_{2week}$ scores were adjusted for sample type prior to primary analysis by increasing $Ki67_{2week}$ scores derived from a resection sample by 15%. In addition such differences will be essential to consider in the application of cut-offs for Ki67. It is possible that some staining procedures may be more sensitive to differences to variability in fixation; it may therefore be prudent for pathologists to assess the impact of fixation quality on Ki67 analysis within their own practise. We have previously reported the impact of short-term AI therapy on grade and this should not be ignored(12). Where an AI has been given in the presurgical or neo-adjuvant setting preference may well be given to assessment of grade from a core rather than excision specimen to minimise this impact.

The suppression of Ki67 by AIs was similar to that reported previously(3, 7) but the suggestion of an apparent statistically significant difference between letrozole and anastrozole in the degree of suppression has not been previously reported. Although type of AI remains significant when adjusting for other clinic-pathological factors, it is important to note that this is not a randomised comparison but the choice of AI was centre dependent influenced by

local clinical practice. Given the difference is only observed in excision samples and not core-cuts and only in HER2-ve tumours, there is a high probability that this difference may be related to unmeasured or artefactual differences- eg in surgical procedures or processing of surgical specimens between centres. There was no difference in clinical outcomes between these two AIs in randomised clinical trials either in advanced breast cancer or in primary ER+ breast cancer(27, 28). There is therefore no evidence for a difference in clinical efficacy of these two agents in spite of a known small difference in estradiol suppression and the Ki67 data reported in this manuscript.

Conclusions

In conclusion, the magnitude of this study allowed assessment of relationships between clinic-pathological factors and $Ki67_{baseline}$, POAI induced and untreated $\Delta Ki67_{2week}$ and $Ki67_{2week}$ with high degrees of confidence. In particular, illustrating that POAI induced $\Delta Ki67_{2week}$ was independent of tumour size, nodal involvement, histology and vascular invasion but associated with both grade and PgR status. Lower values of Ki67 occur when measured on excision specimens rather than core-cut biopsies and both these factors should be considered when either $\Delta Ki67_{2week}$ or $Ki67_{2week}$ are used in routine clinical practice to aid treatment decisions or in clinical trials to assess new drug therapies. Our recommendation would be to use core-core comparisons where possible with the second core being taken in situ as soon as the tumour is excised to avoid this artefact.

Declarations

Ethics approval and consent to participate

POETIC was approved by the London South-East Research Ethics Committee (reference number 08/H1102/37). All patients provided written informed consent. The Clinical Trials and Statistics Unit at The Institute of Cancer Research, London, UK (ICR-CTSU), had overall responsibility for trial and data management.

Consent for publication

Not applicable.

Availability of data and materials

De-identified data will be made available to other researchers on request, subject to approval of a formal data access request in accordance with the ICR-CTSU data and sample access policy. Trial documentation including the protocol are available on request by contacting poetic-icrctsu@icr.ac.uk. The ICR-CTSU supports the wider dissemination of information from the research it does, and increased cooperation between investigators. Trial data is collected, managed, stored, shared, and archived according to ICR-CTSU Standard Operating Procedures in order to ensure the enduring quality, integrity, and utility of the data. Formal requests for data sharing are considered in line with the Institute of Cancer Research Clinical Trials and Statistics Unit (ICR-CTSU) procedures with due regard given to funder and sponsor guidelines. Requests are via a standard proforma describing the nature of the proposed research and extent of data requirements. Data recipients are required to enter a formal data sharing agreement which describes the conditions for release and requirements for data transfer, storage, archiving, publication and intellectual property. Requests are reviewed by the Trial Management Group (TMG) in terms of scientific merit and ethical considerations including patient consent. Data sharing is allowed if proposed projects have a

sound scientific or patient benefit rationale as agreed by the TMG and approved by the Trial Steering Committee as required. Restrictions relating to patient confidentiality and consent will be limited by aggregating and anonymising identifiable patient data. Additionally all indirect identifiers that might lead to deductive disclosures will be removed in line with Cancer Research UK Data Sharing Guidelines. Additional documents might be shared if approved by the TMG and Trial Steering Committee (eg, statistical analysis plan and informed consent form).

Competing interests

JMB reports grants from Cancer Research UK, during the conduct of the study; grants from Medivation; grants and non-financial support from AstraZeneca, Merck Sharp & Dohme, Puma Biotechnology, Clovis Oncology, Pfizer, Janssen-Cilag, Novartis, and Roche, outside the submitted work. LK reports grants from Cancer Research UK, during the conduct of the study. MD reports grants from Cancer Research UK, during the conduct of the study; and personal fees from Radius, Roche, Myriad, Orion, G1 Therapeutics, Nanostring, AbbVie, H3 Biomedicine, Lilly, and the ICR Rewards for Inventors Scheme, outside the submitted work. All other authors declare no competing interests.

Funding

POETIC was co-sponsored by The Institute of Cancer Research and The Royal Marsden NHS Foundation Trust. We are grateful for the support from the National Institute for Health Research (NIHR) Clinical Research Network and for the study grant from Cancer Research UK (CRUK/07/015 grant reference A8671). ICR-CTSU also receives programme grant funding from Cancer Research UK, grant number C1491–A15955.

514

515 ***Authors' contribution***

516 MB assisted with trial design, protocol development, statistical analysis, data interpretation,
517 and writing and was a Trial Management Group member. HT assisted with statistical analysis,
518 data interpretation, and writing. AE, CH, KH, RV and AS assisted with participant recruitment
519 and data collection and were Trial Management Group members. EM assisted with data
520 collection, data analysis, and data interpretation, and was a Trial Management Group
521 member. AD assisted with data analysis and data interpretation. MH, SD and LZ assisted with
522 data collection and analysis. JB assisted with trial management, data collection, and data
523 management, and was a Trial Management Group member. LK assisted with statistical
524 analysis and data interpretation, and was a Trial Management Group member. JPM assisted
525 with trial design, protocol development, statistical analysis, and data interpretation, and was
526 a Trial Management Group member. IS was chief investigator. IS and JR assisted with trial
527 design, protocol development, participant recruitment, data collection, data interpretation,
528 and writing and were Trial Management Group members. MD assisted with trial design,
529 protocol development, data analysis, data interpretation, and writing, and was a Trial
530 Management Group member. All authors reviewed the manuscript before submission.

531

532 ***Acknowledgements***

533 The POETIC trial represents independent research supported by the National Institute for
534 Health Research Biomedical Research Centre at The Royal Marsden NHS Foundation Trust
535 and the Institute of Cancer Research, London. The views expressed are those of the authors
536 and not necessarily those of the NIHR or the Department of Health and Social Care. We
537 thank all the patients and their families who participated in this study, all staff involved at

the 132 participating hospitals, the staff involved in the trial at ICR-CTSU, Ralph Lauren Centre for Breast Cancer Research at the Royal Marsden NHS Foundation Trust, and the Centre for Molecular Pathology at the Institute of Cancer Research. Finally, we thank the past and present colleagues on the POETIC Trial Management Group, and the POETIC Independent Data Monitoring Committee and Trial Steering Committee.

References

1. Urruticoechea A, Smith IE, Dowsett M. Proliferation marker Ki-67 in early breast cancer. J Clin Oncol. 2005;23(28):7212-20.
2. Dowsett M, Nielsen TO, A'Hern R, Bartlett J, Coombes RC, Cuzick J, et al. Assessment of Ki67 in breast cancer: recommendations from the International Ki67 in Breast Cancer working group. J Natl Cancer Inst. 2011;103(22):1656-64.
3. Dowsett M, Smith IE, Ebbs SR, Dixon JM, Skene A, Griffith C, et al. Short-term changes in Ki-67 during neoadjuvant treatment of primary breast cancer with anastrozole or tamoxifen alone or combined correlate with recurrence-free survival. Clin Cancer Res. 2005;11(2 Pt 2):951s-8s.
4. Administration FDAFaD. P210026: Ki-67 IHC MIB-1 pharmDx (Dako Omnis) Approval letter. 2021.
5. Nitz UA, Gluz O, Kummel S, Christgen M, Braun M, Aktas B, et al. Endocrine Therapy Response and 21-Gene Expression Assay for Therapy Guidance in HR+/HER2- Early Breast Cancer. J Clin Oncol. 2022:JCO2102759.
6. Dowsett M, Smith I, Robertson J, Robison L, Pinhel I, Johnson L, et al. Endocrine therapy, new biologicals, and new study designs for presurgical studies in breast cancer. J Natl Cancer Inst Monogr. 2011;2011(43):120-3.
7. Ellis MJ, Suman VJ, Hoog J, Lin L, Snider J, Prat A, et al. Randomized phase II neoadjuvant comparison between letrozole, anastrozole, and exemestane for postmenopausal women with estrogen receptor-rich stage 2 to 3 breast cancer: clinical and biomarker outcomes and predictive value of the baseline PAM50-based intrinsic subtype--ACOSOG Z1031. J Clin Oncol. 2011;29(17):2342-9.
8. M Baum AUB, J Cuzick, J Forbes, J H Houghton, J G M Klijn, T Sahmoud; ATAC Trialists' Group. Anastrozole alone or in combination with tamoxifen versus tamoxifen alone for adjuvant treatment of postmenopausal women with early breast cancer: first results of the ATAC randomised trial. The Lancet. 2002;359(9324):2131-9.
9. Ellis MJ, Suman VJ, Hoog J, Goncalves R, Sanati S, Creighton CJ, et al. Ki67 Proliferation Index as a Tool for Chemotherapy Decisions During and After Neoadjuvant Aromatase Inhibitor Treatment of Breast Cancer: Results From the American College of Surgeons Oncology Group Z1031 Trial (Alliance). J Clin Oncol. 2017;35(10):1061-9.

10. Dowsett M, Smith IE, Ebbs SR, Dixon JM, Skene A, A'Hern R, et al. Prognostic value of Ki67 expression after short-term presurgical endocrine therapy for primary breast cancer. J Natl Cancer Inst. 2007;99(2):167-70.
11. Ma CX, Gao F, Luo J, Northfelt DW, Goetz M, Forero A, et al. NeoPalAna: Neoadjuvant Palbociclib, a Cyclin-Dependent Kinase 4/6 Inhibitor, and Anastrozole for Clinical Stage 2 or 3 Estrogen Receptor-Positive Breast Cancer. Clin Cancer Res. 2017;23(15):4055-65.
12. Smith I, Robertson J, Kilburn L, Wilcox M, Evans A, Holcombe C, et al. Long-term outcome and prognostic value of Ki67 after perioperative endocrine therapy in postmenopausal women with hormone-sensitive early breast cancer (POETIC): an open-label, multicentre, parallel-group, randomised, phase 3 trial. Lancet Oncol. 2020;21(11):1443-54.
13. Leung SCY, Nielsen TO, Zabaglo L, Arun I, Badve SS, Bane AL, et al. Analytical validation of a standardized scoring protocol for Ki67: phase 3 of an international multicenter collaboration. NPJ Breast Cancer. 2016;2:16014.
14. Dowsett M, Ebbs SR, Dixon JM, Skene A, Griffith C, Boeddinghaus I, et al. Biomarker changes during neoadjuvant anastrozole, tamoxifen, or the combination: influence of hormonal status and HER-2 in breast cancer--a study from the IMPACT trialists. J Clin Oncol. 2005;23(11):2477-92.
15. Zabaglo L, Salter J, Anderson H, Quinn E, Hills M, Detre S, et al. Comparative validation of the SP6 antibody to Ki67 in breast cancer. J Clin Pathol. 2010;63(9):800-4.
16. Dowsett M, Harper-Wynne C, Boeddinghaus I, Salter J, Hills M, Dixon M, et al. HER-2 amplification impedes the antiproliferative effects of hormone therapy in estrogen receptor-positive primary breast cancer. Cancer Res. 2001;61(23):8452-8.
17. Ellis MJ, Tao Y, Young O, White S, Proia AD, Murray J, et al. Estrogen-independent proliferation is present in estrogen-receptor HER2-positive primary breast cancer after neoadjuvant letrozole. J Clin Oncol. 2006;24(19):3019-25.
18. Bergamino MA, Lopez-Knowles E, Morani G, Tovey H, Kilburn L, Schuster EF, et al. HER2-enriched subtype and novel molecular subgroups drive aromatase inhibitor resistance and an increased risk of relapse in early ER+/HER2+ breast cancer. EBioMedicine. 2022;83:104205.
19. Early Breast Cancer Trialists' Collaborative G. Aromatase inhibitors versus tamoxifen in early breast cancer: patient-level meta-analysis of the randomised trials. Lancet. 2015;386(10001):1341-52.
20. Dowsett M, Allred C, Knox J, Quinn E, Salter J, Wale C, et al. Relationship Between Quantitative Estrogen and Progesterone Receptor Expression and Human Epidermal Growth Factor Receptor 2 (HER-2) Status With Recurrence in the Arimidex, Tamoxifen, Alone or in Combination Trial. Journal of Clinical Oncology. 2008;26(7):1059-65.
21. Bartlett JM, Brookes CL, Robson T, van de Velde CJ, Billingham LJ, Campbell FM, et al. Estrogen receptor and progesterone receptor as predictive biomarkers of response to endocrine therapy: a prospectively powered pathology study in the Tamoxifen and Exemestane Adjuvant Multinational trial. J Clin Oncol. 2011;29(12):1531-8.
22. Viale G, Regan MM, Dell'Orto P, Mastropasqua MG, Maiorano E, Rasmussen BB, et al. Which patients benefit most from adjuvant aromatase inhibitors? Results using a composite measure of prognostic risk in the BIG 1-98 randomized trial. Ann Oncol. 2011;22(10):2201-7.
23. Gellert P, Segal CV, Gao Q, Lopez-Knowles E, Martin LA, Dodson A, et al. Impact of mutational profiles on response of primary oestrogen receptor-positive breast cancers to oestrogen deprivation. Nat Commun. 2016;7:13294.

24. Romero Q, Bendahl PO, Klintman M, Loman N, Ingvar C, Ryden L, et al. Ki67 proliferation in core biopsies versus surgical samples - a model for neo-adjuvant breast cancer studies. BMC Cancer. 2011;11:341.
25. Acs B, Leung SCY, Kidwell KM, Arun I, Augulis R, Badve SS, et al. Systematically higher Ki67 scores on core biopsy samples compared to corresponding resection specimen in breast cancer: a multi-operator and multi-institutional study. Mod Pathol. 2022.
26. Hadad SM, Jordan LB, Roy PG, Purdie CA, Iwamoto T, Pusztai L, et al. A prospective comparison of ER, PR, Ki67 and gene expression in paired sequential core biopsies of primary, untreated breast cancer. BMC Cancer. 2016;16(1):745.
27. Rose C, Vtoraya O, Pluzanska A, Davidson N, Gershanovich M, Thomas R, et al. An open randomised trial of second-line endocrine therapy in advanced breast cancer. comparison of the aromatase inhibitors letrozole and anastrozole. Eur J Cancer. 2003;39(16):2318-27.
28. Smith I, Yardley D, Burris H, De Boer R, Amadori D, McIntyre K, et al. Comparative Efficacy and Safety of Adjuvant Letrozole Versus Anastrozole in Postmenopausal Patients With Hormone Receptor-Positive, Node-Positive Early Breast Cancer: Final Results of the Randomized Phase III Femara Versus Anastrozole Clinical Evaluation (FACE) Trial. J Clin Oncol. 2017;35(10):1041-8.

Figure legends

Figure 1. Consort diagram of available samples

Fewer surgical samples from control patients were analysed because little extra value was expected from multiple samples in the absence of treatment. A subset of one third of control patients were selected at random for analysis while all patients in the treatment group were analysed, this lead to 7/9 samples from the whole trial being analysed.

Figure 2. Distribution of Ki67

Distribution of A. $Ki67_{Baseline}$ for all patients, B. $Ki67_{2week}$ in patients allocated control, C. percentage change Ki67 in patients allocated control, D. $Ki67_{2week}$ in patients allocated AI and E. percentage change Ki67 in patients allocated AI. Presented separately for HER2- and HER2+ patients.

Figure 3. Distribution of Ki67 by clinic-pathological factors

Distribution of A. $Ki67_{Baseline}$ for all patients, B. $Ki67_{2weeks}$ in patients allocated AI and C. log fold change Ki67 in patients allocated AI by clinic-pathological factors. Presented separately for HER2- and HER2+ patients.

Tables

Table 1. Univariable and multivariable linear regression results for $Ki67_{Baseline}^{\dagger}$ by HER2 status

		ER+ HER2-							ER+ HER2+						
		Univariable			Multivariable				Univariable			Multivariable			
		β	ci	p-value	β	ci	p-value	Adjusted critical value	β	ci	p-value	β	ci	p-value	Adjusted critical value
PgR Status	Positive	0	-	<0.001	0	-	**0.038**	**0.043**	0	-	0.980	0	-	0.366	0.029
	Negative	0.21	0.12 - 0.31		0.11	0.02 - 0.20			0.20	0.02 - 0.38		0.11	-0.06 - 0.29		
	Unknown	0.01	-0.05 - 0.07		0.00	-0.06 - 0.06			0.10	-0.08 - 0.28		0.09	-0.09 - 0.26		
Tumour grade (baseline)	1	0	-	<0.001	0	-	**<0.001**	**0.007**	0	-	<0.001	0	-	**<0.001**	**0.007**
	2	0.43	0.36 - 0.51		0.40	0.33 - 0.48			0.56	0.20 - 0.92		0.60	0.23 - 0.97		
	3	1.16	1.06 - 1.25		1.04	0.94 - 1.14			1.04	0.68 - 1.40		1.05	0.68 - 1.42		
	Not known	0.40	0.29 - 0.51		0.35	0.24 - 0.46			0.69	0.27 - 1.11		0.74	0.30 - 1.18		
Tumour size (baseline)	0-2cm	0	-	<0.001 (<0.001)*	0	-	**<0.001**	**0.021**	0	-	0.167 (0.043)*	0	-	0.561	0.036
	2-5cm	0.24	0.18 - 0.30		0.14	0.08 - 0.20			0.14	-0.01 - 0.29		0.07	-0.08 - 0.22		
	>5cm	0.28	0.08 - 0.49		0.11	-0.08 - 0.31			0.23	-0.46 - 0.91		-0.13	-0.80 - 0.54		
Histological type (baseline)	Ductal	0	-	<0.001	0	-	**<0.001**	**0.014**	0	-	0.286	0	-	0.798	0.043
	Lobular	-0.30	-0.38 - -0.22		-0.24	-0.31 - -0.16			-0.07	-0.40 - 0.25		0.07	-0.21 - 0.35		
	Other	-0.50	-0.65 - -0.35		-0.24	-0.39 - -0.10			-0.27	-1.03 - 0.49		-0.09	-0.54 - 0.36		
Nodal status	N0	0	-	<0.001	0	-	**0.008**	**0.036**	0	-	0.431	0	-	0.894	0.050

	N1-3	0.10	0.03 - 0.16	(<0.001)*	0.00	-0.06 - 0.06			0.06	-0.11 - 0.24	(0.211)*	0.02	-0.15 - 0.19		
	N4+	0.32	0.23 - 0.42		0.14	0.05 - 0.24			0.13	-0.08 - 0.33		-0.03	-0.25 - 0.19		
Age group	<60	0.02	-0.06 - 0.10	0.230	-0.03	-0.10 - 0.05	0.167	0.050	0.02	-0.18 - 0.22	0.052	0.07	-0.12 - 0.26	0.183	0.021
	60-69	0	-	(0.059)	0	-			0	-		0	-		
	70-79	0.04	-0.03 - 0.11		-0.01	-0.07 - 0.06			0.15	-0.03 - 0.34	(0.791)*	0.12	-0.06 - 0.30		
	80+	-0.07	-0.16 - 0.03		-0.10	-0.19 - -0.01			-0.25	-0.52 - 0.02		-0.17	-0.44 - 0.10		
Vascular invasion	Yes	0	-	<0.001	0	-	**<0.001**	**0.029**	0	-	0.068	0	-	0.117	0.143
	No	-0.34	-0.41 - -0.28		-0.15	-0.21 - -0.08			-0.17	-0.32 - -0.02		-0.16	-0.32 - 0.00		
	Not reported	-0.24	-0.38 - -0.10		-0.10	-0.23 - 0.03			0.06	-0.39 - 0.51		0.04	-0.40 - 0.47		

**Test for trend*

† Analysed as ln(Ki67 + 0.1)

Adjusted critical values calculated using Benjamini Hochberg method. Significant p-values following adjustment are highlighted in bold

Table 2. Univariable and multivariable linear regression results for $Ki67_{2week}^{\dagger}$ in patients allocated to AI by HER2 status

		ER+ HER2-							ER+ HER2+						
		Univariable			Multivariable				Univariable			Multivariable			
		β	ci	p-value	β	ci	p-value	Adjusted critical value	β	ci	p-value	β	ci	p-value	Adjusted critical value
PgR Status	Positive	0	-	<0.001	0	-	**<0.001**	**0.014**	0	-	<0.001	0	-	**0.004**	**0.014**
	Negative	0.70	0.53 - 0.88		0.49	0.33 - 0.65			0.75	0.40 - 1.09		0.50	0.19 - 0.81		
	Unknown	0.18	0.06 - 0.30		0.11	0.01 - 0.22			0.27	-0.06 - 0.60		0.09	-0.20 - 0.39		
Tumour grade (baseline)	1	0	-	<0.001					0	-	<0.001				
	2	0.44	0.29 - 0.59						0.82	0.13 - 1.52					
	3	1.51	1.32 - 1.70						1.66	0.95 - 2.36					
	Not known	0.63	0.41 - 0.85						1.36	0.55 - 2.16					
Tumour grade (2week)	1	0	-	<0.001 (0.001)*	0	-	**<0.001**	**0.007**	0	-	<0.001 (<0.001)*	0	-	**<0.001**	**0.007**
	2	0.50	0.37 - 0.63		0.52	0.38 - 0.66			0.70	0.03 - 1.37		0.41	-0.27 - 1.08		
	3	1.96	1.79 - 2.14		1.85	1.67 - 2.03			1.88	1.20 - 2.55		1.44	0.75 - 2.14		
Tumour size (baseline)	0-2cm	0	-	<0.001 (0.001)*					0	-	0.036 (0.066)*				
	2-5cm	0.33	0.22 - 0.44						0.36	0.07 - 0.65					
	>5cm	0.21	-0.19 - 0.60						-0.43	-1.81 - 0.96					
Tumour size (2week)	0-2cm	0	-	<0.001 (0.001)*	0	-	0.224	0.035	0	-	<0.001 (0.002)*	0	-	0.042	0.021
	2-5cm	0.23	0.12 - 0.34		0.02	-0.08 - 0.13			0.63	0.34 - 0.93		0.34	0.07 - 0.61		
	>5cm	0.08	-0.19 - 0.35		-0.19	-0.44 - 0.06			0.49	-0.19 - 1.18		0.28	-0.35 - 0.91		

Histological type (baseline)	Ductal	0	-	<0.001					0	-	0.393				
	Lobular	-0.34	-0.50 - -0.19						-0.41	-1.02 - 0.19					
	Other	-0.35	-0.65 - -0.06						-0.06	-0.98 - 0.86					
Histological type (2week)	Ductal	0	-	<0.001	0	-	**<0.001**	**0.021**	0	-	0.198	0	-	0.572	0.043
	Lobular	-0.40	-0.54 - -0.25		-0.30	-0.44 - -0.16			-0.40	-1.03 - 0.23		-0.18	-0.72 - 0.37		
	Other	-0.33	-0.59 - -0.07		-0.04	-0.28 - 0.21			-0.80	-2.00 - 0.39		-0.43	-1.45 - 0.60		
Nodal status	N0	0	-	0.005 (0.002)	0	-	0.782	0.050	0	-	0.058 (0.191)*	0	-	0.420	0.036
	N1-3	0.12	-0.01 - 0.24		0.00	-0.11 - 0.12			-0.11	-0.45 - 0.22		-0.07	-0.37 - 0.24		
	N4+	0.28	0.10 - 0.47		0.06	-0.12 - 0.24			0.40	0.01 - 0.79		0.19	-0.19 - 0.57		
Age group	<60	0.15	0.00 - 0.30	0.265 (0.775)*	0.10	-0.04 - 0.23	0.526	0.043	0.38	-0.00 - 0.77	0.100 (0.985)*	0.38	0.05 - 0.71	0.129	0.029
	60-69	0	-		0	-			0	-		0	-		
	70-79	0.05	-0.09 - 0.18		0.01	-0.11 - 0.13			0.37	0.02 - 0.71		0.15	-0.15 - 0.45		
	80+	0.05	-0.13 - 0.23		0.05	-0.12 - 0.21			0.10	-0.46 - 0.66		0.19	-0.29 - 0.68		
Vascular invasion	Yes	0	-	<0.001	0	-	**0.014**	**0.029**	0	-	0.101	0	-	0.670	0.050
	No	-0.40	-0.52 - -0.28		-0.13	-0.25 - -0.00			-0.30	-0.60 - -0.01		0.02	-0.27 - 0.30		
	Not reported	-0.09	-0.36 - 0.17		0.16	-0.08 - 0.41			0.12	-0.80 - 1.05		0.35	-0.45 - 1.15		

**Test for trend*

† Analysed as ln(Ki67 + 0.1)

Adjusted critical values calculated using Benjamini Hochberg method. Significant p-values following adjustment are highlighted in bold

Table 3. Univariable and multivariable linear regression results for change in Ki67 ($^{\dagger}\Delta Ki67_{2week}$) in patients allocated to AI by HER2 status

		ER+ HER2-							ER+ HER2+						
		Univariable			Multivariable				Univariable			Multivariable			
		β	ci	p-value	β	ci	p-value	Adjusted critical value	β	ci	p-value	β	ci	p-value	Adjusted critical value
Baseline Ki67 (log)		-0.23	-0.28 - -0.17	<0.001	-0.41	-0.47 - -0.35	**<0.001**	**0.006**	-0.35	-0.50 - -0.20	<0.001	-0.61	-0.77 - -0.46	**<0.001**	**0.006**
PgR Status	Positive	0	-	<0.001	0	-	**<0.001**	**0.019**	0	-	0.023	0	-	**0.008**	**0.019**
	Negative	0.49	0.34 - 0.65		0.45	0.30 - 0.60			0.46	0.13 - 0.78		0.45	0.15 - 0.75		
	Unknown	0.13	0.02 - 0.23		0.11	0.01 - 0.22			0.15	-0.17 - 0.46		0.07	-0.22 - 0.36		
Tumour grade (baseline)	1	0	-	<0.001					0	-	0.378				
	2	-0.01	-0.15 - 0.12						0.23	-0.46 - 0.93					
	3	0.30	0.13 - 0.48						0.45	-0.26 - 1.15					
	Not known	0.19	-0.01 - 0.39						0.38	-0.43 - 1.19					
Tumour grade (2week)	1	0	-	<0.001 (<0.001)*	0	-	**<0.001**	**0.013**	0	-	<0.001 (<0.001)*	0	-	**<0.001**	**0.013**
	2	0.01	-0.12 - 0.13		0.23	0.10 - 0.37			0.31	-0.39 - 1.01		0.32	-0.34 - 0.97		
	3	0.68	0.51 - 0.85		1.16	0.97 - 1.34			0.84	0.13 - 1.54		1.14	0.45 - 1.82		
Tumour size (baseline)	0-2cm	0	-	0.221 (0.015)*					0	-	0.094 (0.337)*				
	2-5cm	0.09	-0.01 - 0.19						0.20	-0.07 - 0.47					
	>5cm	0.05	-0.30 - 0.40						-0.95	-2.24 - 0.34					
	0-2cm	0	-	0.952	0	-	0.525	0.044	0	-	0.215	0	-	0.201	0.031

Tumour size (2week)	2-5cm	0.00	-0.10 - 0.10	(0.601)*	-0.03	-0.13 - 0.07			0.25	-0.03 - 0.53	(0.249)*	0.23	-0.03 - 0.50		
	>5cm	-0.04	-0.27 - 0.20		-0.13	-0.36 - 0.10			0.19	-0.46 - 0.84		0.19	-0.42 - 0.80		
Histological type (baseline)	Ductal	0	-						0	-					
	Lobular	-0.02	-0.16 - 0.11	0.452					-0.22	-0.78 - 0.34	0.639				
	Other	0.16	-0.10 - 0.42						0.22	-0.64 - 1.07					
Histological type (2week)	Ductal	0	-		0	-			0	-		0	-		
	Lobular	-0.10	-0.23 - 0.03	0.252	-0.14	-0.27 - -0.01	0.091	0.031	-0.21	-0.80 - 0.38	0.499	-0.16	-0.69 - 0.37	0.624	0.044
	Other	0.08	-0.16 - 0.31		0.06	-0.17 - 0.29			-0.54	-1.66 - 0.57		-0.37	-1.36 - 0.62		
Nodal status	N0	0	-		0	-			0	-		0	-		
	N1-3	0.03	-0.08 - 0.14	0.692 (0.773)*	0.01	-0.10 - 0.12	0.880	0.050	-0.16	-0.47 - 0.16	0.111 (0.400)*	-0.09	-0.38 - 0.21	0.252	0.038
	N4+	-0.05	-0.21 - 0.11		-0.03	-0.20 - 0.13			0.28	-0.09 - 0.64		0.23	-0.14 - 0.59		
Age group	<60	0.12	-0.01 - 0.25		0.10	-0.02 - 0.23			0.24	-0.13 - 0.60		0.34	0.02 - 0.66		
	60-69	0	-	0.166	0	-	0.292	0.038	0	-	0.285	0	-	0.130	0.025
	70-79	0.00	-0.12 - 0.12	(0.992)*	0.00	-0.11 - 0.12			0.14	-0.18 - 0.46	(0.929)*	0.08	-0.21 - 0.38		
	80+	0.12	-0.05 - 0.28		0.09	-0.07 - 0.24			0.45	-0.08 - 0.97		0.32	-0.15 - 0.79		
Vascular invasion	Yes	0	-		0	-			0	-		0	-		
	No	-0.05	-0.16 - 0.06	0.1	-0.06	-0.17 - 0.06	0.033	0.025	-0.11	-0.39 - 0.17	0.734	0.04	-0.24 - 0.32	0.809	0.050
	Not reported	0.18	-0.06 - 0.42		0.22	-0.00 - 0.45			-0.08	-0.94 - 0.79		0.24	-0.53 - 1.01		

**Test for trend*

$^{\dagger}\Delta Ki67_{2week} = \ln((Ki67_{2week}+0.1)/(\ln(Ki67_{baseline}+0.1))$

Adjusted critical values calculated using Benjamini Hochberg method. Significant p-values following adjustment are highlighted in bold