Long-term Outcome and Prognostic Value of Ki67 after Perioperative Endocrine Therapy in Postmenopausal Women with Hormone Sensitive Early Breast Cancer: The POETIC Randomised Trial

Authors: Ian Smith*, John Robertson*, Lucy Kilburn, Maggie Wilcox, Abigail Evans, Chris Holcombe, Kieran Horgan, Cliona Kirwan, Elizabeth Mallon, Mark Sibbering, Anthony Skene, Raghavan Vidya, Maggie Cheang, Jane Banerji, James Morden¥, Kally Sidhu, Andrew Dodson, Judith M Bliss**, Mitch Dowsett**

* Joint first and **joint last authors

¥ deceased

Affiliations:

The Institute of Cancer Research & The Royal Marsden NHS Foundation Trust, London, UK (Prof Ian Smith MD, Prof Mitch Dowsett PhD); The Institute of Cancer Research, London, UK (Lucy Kilburn MSc, James Morden MSc, Maggie Cheang PhD, Jane Banerji BSc, Prof Judith Bliss MSc); The Royal Marsden NHS Foundation Trust, London, UK (Kally Sidhu BSc, Andrew Dodson MPhil); University of Nottingham, Nottingham & Derby & Burton University Hospitals, UK (Prof John Robertson MD); Independent Cancer Patients Voice, London, UK (Maggie Wilcox); Poole Hospital NHS Foundation Trust, Poole, UK (Abigal Evans FRCS); Liverpool University Hospitals Foundation Trust, Liverpool, UK (Prof Chris Holcombe MD); Bexley Cancer Centre, Leeds, UK (Prof Kieran Horgan FRCS); University of Manchester and Manchester University NHS Foundation Trust (Prof Cliona Kirwan PhD); Queen Elizabeth University Hospital Glasgow, Govan, UK (Elizabeth Mallon MBChB); University Hospitals of Derby and Burton, Derby, UK (Mark Sibbering FRCS); Royal

Bournemouth and Christchurch NHS Foundation Trust, Bournemouth, UK (Anthony Skene FRCS); Birmingham University and Royal Wolverhampton NHS Trust, UK (Raghavan Vidya FRCS)

Address for correspondence

Prof Ian Smith

c/o ICR-CTSU

The Institute of Cancer Research

London SM2 5NG

POETIC-icrctsu@icr.ac.uk

Email: Ian.Smith@rmh.nhs.uk

POETIC-icrctsu@icr.ac.uk

Tel: +44 (0)208 661 3280

Abstract

Background: Pre/peri-operative aromatase inhibitor (POAI) therapy has potential to improve outcomes in women with operable ER-positive (ER+) primary breast cancer (BC). It has also been suggested that tumour Ki67 levels after two weeks ($Ki67_{2W}$) of POAI predicts individual patient outcome better than baseline Ki67 ($Ki67_{B}$). The POETIC trial tested these two hypotheses.

Methods: POETIC was an open-label randomised phase III trial (130 UK hospitals) where post-menopausal women age≥50 years and WHO performance status 0-1 with ER+ BC were randomised (2:1 allocation ratio) to POAI (letrozole 2.5mg/day or anastrozole 1mg/day orally) for 14 days before and following surgery or no POAI (Control). Adjuvant treatment was given as per UK standard local practice. Randomisation, performed centrally by computer-generated permuted block method (variable-block size six or nine), was stratified by hospital. Treatment allocation was not masked. The primary endpoint was Time to Recurrence (TTR). A key second objective explored association between Ki67 (dichotomised at 10%) and disease outcomes. The primary analysis for clinical endpoints was by modified Intention to Treat. For Ki67 biomarker association and endpoint analysis the evaluable population including all randomised patients who had Ki67 values available. This study is registered with ClinicalTrials.gov, number NCT02338310; the European Clinical Trials database, number EudraCT2007-003877-21; and the ISRCTN registry, number ISRCTN63882543. Recruitment is complete and long-term follow-up is ongoing.

Findings: Between 13/10/2008-16/04/2014, 4480 women (2976 POAI, 1504 Control) were recruited. Median age was 67 years (IQR 62-75), 54% had pathological tumour

size>2cm and 39% were node positive. On 06/02/2018, median follow-up was 62·9 months (IQR 58·1-74·1). 434 (9·7%) of the 4480 women had a TTR event (280 (9·4%) POAI, 154 (10·2%) Control): HR=0·92 (95%CI: 0·75-1·12); p=0·40 with the proportion TTR event-free at five years of 91.0% (95%CI: 89.9, 92.0) for patients allocated POAI and 90.4% (95%CI: 88.7, 91.9) for control. Within the POAI-treated HER2-negative sub-population, 5-year recurrence risk in women with low $Ki67_B$ and $Ki67_{2W}$ (L-L) was 4·3% (95%CI: 2.9-6.3), 8.4% (95%CI: 6.8, 10.5) with high $Ki67_B$ and low $Ki67_{2W}$ (H-L) and 21.5% (95%CI: 17.1, 27.0) with high $Ki67_B$ and $Ki67_{2W}$ (H-H). Within the POAI-treated HER2-positive sub-population, 5-year recurrence risk in L-L group was 10.1% (95%CI: 3.2-31.3), 7.7% (95%CI: 3.4, 17.5) in the H-L group and 15.7% (95%CI: 10.1, 24.4) in the H-H group.

Interpretation: POAI has not been shown to improve treatment outcome, but can be used without detriment to help select appropriate adjuvant therapy based on tumour Ki67. Most patients with low $Ki67_B$ or low POAI-induced $Ki67_{2W}$ do well with adjuvant standard endocrine therapy (giving consideration to clinical-pathological factors), while those whose POAI-induced $Ki67_{2W}$ remain high may benefit from further adjuvant treatment(s) and/or trials of new therapies.

Funding: Cancer Research UK

Introduction

The POETIC (Peri-Operative Endocrine Therapy – Individualising Care) Trial was designed to address two important hypotheses in the treatment of post-menopausal women with ER+ early breast cancer.

The first was that short duration pre-surgical endocrine therapy might improve clinical outcome. This hypothesis was plausible since two weeks' pre-operative endocrine therapy had been shown to markedly reduce proliferation in human breast cancer as measured by Ki67[1,2]. Longstanding experimental evidence had shown that the stimulatory effect of surgery on the growth of metastases in mice could be inhibited by peri-operative endocrine therapy[3,4]. Any improvement in long-term outcome following short exposure to pre/peri-operative endocrine therapy would be achieved with no additional toxicity or resource implications and be of considerable clinical importance.

The second hypothesis concerned identifying which patients with hormone receptor positive (HR+) early breast cancer have a sufficiently good prognosis such that standard of care medical treatment, often comprising adjuvant endocrine therapy alone, was sufficient and which group should be considered for additional therapies. Traditional approaches to this problem had used standard prognostic parameters including size, grade, nodal involvement, and age, often integrated into a prognostic tool (e.g. Nottingham Prognostic Index[5], Adjuvant Online[6], NHS PREDICT[7]), but these merely provided the predicted probability of benefit for a patient population with given tumour and demographic characteristics. More recently, genomic platforms have been developed aimed at providing more accurate prognostic and predictive information for

the individual patient[8,9]. These genomic tests however are expensive, by no means universally available, and differ[10] amongst themselves in the information provided.

A simple test which predicts outcome after short duration pre-operative endocrine therapy could therefore be helpful in accurately selecting appropriate treatment in the individual patient, having incorporated an *in vivo* response to AI. A small neo-adjuvant trial (IMPACT) had already suggested this may be feasible: it had shown that tumour Ki67 after two weeks ($Ki67_{2w}$) of endocrine treatment predicted outcome better than at baseline ($Ki67_{B}$), remaining statistically significant in multivariable analysis, whereas $Ki67_{B}$ did not[11,12]. Similar results have subsequently been reported from another small trial comparing letrozole with tamoxifen[13] and from a further trial comparing anastrozole, letrozole and exemestane with one another[14]. POETIC, with a much larger patient population, built on these findings to provide the definitive clinical evidence to inform future practice.

METHODS

Study design and participants

This multicentre, parallel group, randomised phase III trial recruited participants from 130 UK hospitals (protocol provided in appendix pp28-71). Eligible patients were post-menopausal women (aged ≥50 years with either amenorrhoea>12 months, bilateral oophorectomy or had a hysterectomy, or had been on HRT within previous 12 months and with FSH levels in postmenopausal range if aged <55 years) with ER and/or PR-positive (Allred≥3, H-score≥2, or ≥1% of positive cells, assessed in local pathology laboratories), HER2-positive or negative (assessed locally), operable primary breast cancer and no evidence of metastatic spread investigated according to local

guidelines. If palpable, a tumour of any size was sufficient, otherwise requiring an ultrasound size of ≥1·5cm. Women required WHO performance status 0-1 and an indication for standard adjuvant endocrine therapy. Required staging investigations, including laboratory investigations, were according to local practice with no additional trial specific investigations. Exclusion criteria were typical for this patient population. Previous endocrine therapy or chemotherapy was not allowed, nor was concurrent use of HRT or any other oestrogen-containing medication (within 4 weeks of randomisation). No previous use of oestrogen implants at any time, current, continuous long-term systemic steroid usage or treatment with an unlicensed or investigational drug within 4 weeks of randomisation was allowed. Patients with invasive malignancy diagnosed within the previous 5 years or any severe co-incident medical disease were ineligible (appendix p1).

Patients provided written informed consent before enrolment. POETIC was sponsored by The Institute of Cancer Research (ICR) and Royal Marsden NHS Foundation Trust and approved by the London – South East Research Ethics Committee (REC reference 08/H1102/37) and managed and analysed by the ICR Clinical Trials and Statistics Unit (ICR-CTSU) (appendix p1 for study oversight details).

Randomisation and masking

Participants were stratified by hospital and randomly allocated (2:1 ratio), to peri-operative aromatase inhibitor (POAI) treatment or no peri-operative treatment (control) by computer-generated permuted block method (variable-block size six or nine) derived centrally by ICR-CTSU using its dedicated randomisation system. To randomise a patient, staff at the recruiting site telephoned ICR-CTSU and thus had no

knowledge of future treatment assignment. The allocation ratio weighted trial information to study of biological peri-operative drug effects, in particular to assess how these effects relate to long-term outcome. No placebo was used; clinicians and patients were not blinded to treatment allocation, but central laboratory staff were blinded.

Procedures

POAI was a non-steroidal aromatase inhibitor (AI) in standard dosage (oral anastrozole 1mg/day or letrozole 2·5mg/day); choice of agent was declared by each participating hospital at trial outset. Prior to randomisation all patients had excisional surgery pre-booked for around two weeks (minimum ten days) later to ensure timing of surgery was not biased by treatment allocation. POAI was to commence immediately after randomisation allowing duration of treatment before surgery to be as close as possible to fourteen days. If surgery was delayed, the pre-treatment duration was extended and treatment continued without interruption until fourteen days after surgery.

All non-trial adjuvant therapy, laboratory investigations and disease staging were determined on clinical grounds according to standard-of-care local practice (appendix p1). All patients had pre-treatment mammography and breast ultrasound according to local diagnostic practice. Mid-trial the IDMC expressed caution relating to the potential influence of POAI therapy on tumour grade measured at surgery. In February 2011 a letter to investigators, followed by an approved protocol amendment, recommended that local multi-disciplinary teams gave due consideration to other factors, including

pre-treatment grade on diagnostic core where available, when considering use of adjuvant chemotherapy.

Follow-up data were submitted annually to ICR-CTSU; disease related events, second cancers and deaths were reported on occurrence. There was no specific safety endpoint. Adverse event data were restricted to three menopausal symptoms (hot flushes, sweating and musculoskeletal pain) at baseline, surgery and at follow-up two weeks post-surgery (assessed using National Cancer Institute Common Terminology Criteria for Adverse Events (CTCAE) version 3) as the safety profiles of the AIs used were well established. Serious adverse events were reported/recorded (as per protocol). Participants were able to withdraw from the trial at any time for any reason.

Formalin-fixed paraffin embedded (FFPE) tissue samples were required prior to randomisation (baseline) and at surgery. Baseline samples could be a core-cut diagnostic biopsy, a subsequent research core-cut biopsy, or sections from the diagnostic sample. At surgery samples could be either core biopsies or sections cut from the routine excision.

Tissue samples were processed, stored and analysed for Ki67 staining centrally in the Ralph Lauren Centre for Breast Cancer Research at the Royal Marsden NHS Foundation Trust. Ki67 was analysed immunohistochemically in a core biopsy taken at baseline ($Ki67_B$), and in either a core biopsy or the excision biopsy taken at surgery ($Ki67_{2W}$), and was estimated as the percentage of cancer cells staining positive. The primary antibody was MIB1 and detection was by the REAL EnVision system both from DAKO (Glostrup, Denmark until 2016; now Agilent Technologies, Didcot, UK).

Scoring was according to methodology including between-batch QC procedures as described by the International Ki67 in Breast Cancer Working Group Party[15]. Analysis of two-week samples from the control group was restricted to a randomly selected subset since minimal change from baseline was expected[16].

Outcomes

The primary endpoint was time to recurrence (TTR) (time from randomisation to local, regional, or distant tumour recurrence or death from breast cancer without prior notification of relapse) with second primary cancers and intercurrent deaths censored. Secondary clinical endpoints included relapse-free-survival (RFS, as per TTR but also including deaths from any cause as events), time to local recurrence (TTLR, time from randomisation to first confirmed local recurrence, censoring at prior distant recurrence, second primary cancer or death), time to distant recurrence (TTDR, time from randomisation to first confirmed distant recurrence or breast cancer death without prior relapse, censoring at second primary cancer or intercurrent death) and overall survival (OS, time from randomisation to death from any cause). Patients who did not have primary breast surgery as planned were censored at the date of that decision. Breast cancer free survival duplicated the definition of TTR, listed in the protocol in error.

The role of Ki67 in POETIC was two-fold; firstly to be evaluated as a biomarker in relation to its effect on predicting disease outcomes (one of the trial's two key objectives) and secondly as the molecular secondary endpoint to assess proliferation rate, at baseline ($Ki67_B$) and at surgery ($Ki67_{2w}$) thus assessing the impact of POAI. The additional molecular secondary endpoint of gene expression profile at core biopsy and at surgical excision is not reported here as data analysis is ongoing.

Statistical Considerations

The sample size assumed the proportion of patients with recurrence by five years would be low (approximately 10%) given known recurrence rates for similar populations[17,18]. With 4350 patients it would be possible to detect a 3% improvement in TTR at five years (10% to 7% recurrences) with 91% power (two-sided alpha of 5%). The sample size was increased originally from 4000 to 4350 patients in order to allow for underestimation of the relapse rate potentially due to patients dying from other causes prior to breast cancer relapse. This change was endorsed by the TSC and IDMC and managed via a protocol amendment approved on November 2012.

Analyses relating to clinical endpoints were according to a modified intention to treat (ITT) – removing patients who subsequently withdrew consent for use of data. For analyses which assessed the predictive value of $Ki67_B$ and $Ki67_{2W}$ the population was defined as all randomised patients who had paired Ki67 values available.

Descriptive statistics illustrated baseline demographic details, tumour characteristics, adjuvant treatment and Ki67 data. Protocol compliance between treatment groups (time from randomisation to surgery and number of inpatient days for surgery) was compared using Wilcoxon rank-sum tests; differences in tumour grade at surgery were assessed using a chi-square test for trend in pre-specified analyses. Worse grade of adverse events were summarised as were serious adverse reactions to POAI. $Ki67_B$ and $Ki67_{2W}$ was reported by HER2 status. Analysis of percentage change between $Ki67_B$ and $Ki67_{2W}$ used Wilcoxon sign rank tests within and Wilcoxon rank-sum test between treatment groups. In an unplanned exploratory analysis, following initial

planned analyses on the trial data, a multivariable logistic regression model was created, using a forward stepwise approach, to determine factors affecting chemotherapy use.

For survival-related endpoints, Kaplan-Meier curves were plotted and treatment groups compared with the log-rank test. Hazard ratios (HRs) and 95% CIs were calculated within Cox proportional hazards regression models, with HRs<1 taken to favour POAI. Proportional hazards assumption was assessed using Schoenfeld residuals and was found to hold. Comparisons between treatment groups were made with and without adjustment for PR status (positive, negative, unknown), HER2 status (positive, negative, unknown), pre-surgical tumour grade (G1, G2, G3), pathological tumour size (continuous), pre-surgical histological type (ductal, lobular, special type), nodal status (N0, N1-3, N4+), age at randomisation (continuous) and vascular invasion (yes, no). Subgroup analyses were conducted for baseline clinical characteristics and presented using a forest plot.

Associations between $Ki67_B$ and $Ki67_{2W}$ and TTR were undertaken separately in the POAI and control groups with the principal focus being to study the on-treatment effect of POAI. An analysis of all patients combined for $Ki67_B$ was included for completeness (appendix p24). Assessment of Ki67 in the control group was considered of limited additional value since i) patients were not exposed to peri-operative treatment and ii) the lack of association between POAI and TTR. Survival analysis included adjustment for clinical factors as mentioned above, except for HER2 status which was a stratifying factor. HER2-positive tumours have a different pattern of recurrence and were typically additionally treated with specific HER2 targeted therapy. To explore associations

between Ki67 and disease outcome in the POAI group Ki67 scores were dichotomized and patients divided into 4 groups as follows: Low-Low (L-L) ($Ki67_B$ and $Ki67_{2W}$<10%); High-Low (H-L) ($Ki67_B$≧10%, $Ki67_{2W}$<10%): High-High (H-H) ($Ki67_B$ and $Ki67_{2W}$ ≥10%); Low-High (L-H) ($Ki67_B$<10%, $Ki67_{2W}$≥10%). Few POAI patients were classified into the L-H group. These are reported for completeness but not further analysed as their apparent response is most likely due to measurement variability around the dichotomization cut-point. Post-hoc subgroup analyses explored associations between Ki67 and disease outcome by chemotherapy use and age with a view to avoid confounding of interpretation. In addition to the pre-defined 10% Ki67 dichotomisation, chosen to ensure consistency with other neoadjuvant trials[12,14], other cut-points were explored using Harrell's C coefficient[19] including that for complete cell cycle arrest (CCCA; Ki67≤2·7%[20]).

Analyses previously presented[21] of change in Ki67 in 679 control group patients with paired samples available indicated that in patients with a core-cut surgery sample the median proportional reduction was -4·1% (IQR: -27·8, 34·8) whereas in those with a resection sample at surgery the median proportional reduction in Ki67 between baseline and surgery was -17·7% (IQR: -44·2, 12·7) in contrast with an earlier small pilot study[16]. From this it was assumed that, for a given surgical sample, change in Ki67 score would be proportionally ≈15% less if the sample was core-cut than resection (e.g. 10% reduction with resection sample translated to 8·5% for core-cut). To account for this difference Ki67 data and the analyses linking Ki67 and TTR were performed with $Ki67_{2W}$ corrected according to surgical sample type. $Ki67_{2W}$ scores from resection samples were increased proportionally by 15%. This correction factor was derived (and used) in control participants and similarly applied to participants

allocated POAI. The correction was also made for patients for whom surgical sample type was unknown. For cases where $Ki67_{2W}$ was 0%, no adjustment was made.

This manuscript describes the primary endpoint analysis, TTR after a 5-year median follow-up for both hypotheses; firstly by randomised POAI allocation and secondly exploring the ability of Ki67 to predict disease outcome. No formal interim analyses were planned or conducted prior to the primary analysis. For this purpose a database snapshot was taken on February 6, 2018. All analyses were conducted using Stata (version 13·1). A p-value <0·05 was deemed to be statistically significant.

This study is registered with ClinicalTrials.gov, number NCT02338310; the European Clinical Trials database, number EudraCT2007-003877-21; and the ISRCTN registry, number ISRCTN63882543.

Role of the funding sources

The funder had no role in study design, data collection, data analysis, data interpretation, or report writing. The corresponding author had full access to study data and had final responsibility for the decision to submit for publication.

RESULTS

Between October 13, 2008 and April 16, 2014, 4486 patients were entered from 130 UK centres. Six patients subsequently withdrew consent for data to be used and therefore 4480 patients (2976 AI, 1504 Control) were available for a modified ITT analysis (Figure 1).

Median age at randomisation was 67·1 years (IQR: 61·5, 74·8), over half the patients 2536 (57%) of 4480 patients had a tumour size ≤2cm and all but 8 of 4480 patients (<1%) were confirmed locally to have HR+ tumours (Table 1). Twenty-three of 4480 patients (1%) did not have surgery as planned (16 allocated POAI and 7 control, Figure 1). Adherence to trial treatment and timelines was excellent. (appendix p14). 177 (6%) of 2976 patients did not have the protocol defined amount of POAI (pre-operatively <10 days or >21 days, post-operatively <10 days). The most common reasons were 63 (2%) had their surgery changed, 35 (1%) had less due to adverse events (16 were in the pre-surgical period) and 30 (1%) had less due to patient choice or omission. Surgical details and post-surgery tumour characteristics were well balanced between groups with the exception of pathological tumour grade, which was higher in the control group (p<0·0001) (Table 1).

Adjuvant radiotherapy and anti-HER2 therapy were given post-surgery with similar frequency for the two groups and in line with current UK standard-of-care. Adjuvant chemotherapy was given to 770 (26%) of 2930 patients allocated POAI and 460 (31%) of 1485 patients allocated control (appendix p15) with multivariable analyses attributing this to differences observed in post-surgical grade (appendix p16). Following surgery, most women were prescribed AI monotherapy (appendix p17).

With 62·9 months median follow-up (IQR 58·1 to 74·1), 434 (10%) of 4480 women had a breast cancer recurrence (POAI 280 (9%) of 2976 patients, control 154 (10%) of 1504 patients) (Table 2) with no significant difference observed between the treatment groups (HR=0·92, (95%CI: 0·75, 1·12); p=0·40, adjusted HR=0·96 (95%CI: 0·77, 1·19); p=0·70) (Figure 2a) with the proportion TTR event-free at five years of 91.0% (95%CI: 89.9, 92.0) for patients allocated POAI and 90.4% (95%CI: 88.7, 91.9)

for control. Subgroup analyses according to clinical characteristics including nodal status were consistent with the overall effect, (appendix p2). Likewise no significant differences were observed for RFS, TTLR and TTDR (Table 3). Second breast cancer primaries developed in 26 (<1%) of 2976 women allocated POAI compared with 24 (2%) of 1504 in the control group. 561 patients had died (POAI 365 (12%) of 2976, control 196 (13%) of 1504). Almost half of deaths were attributable to non-breast cancer cause; none were treatment-related (Table 2). There was no evidence of a difference in OS between treatment groups (unadjusted HR=0·94 (95%CI: 0·79, 1·12); p=0·50, adjusted HR=0·91 (95%CI: 0·75, 1·10); p=0·33) (Figure 2b).

Selected menopausal symptoms were assessed in 4201 (94%) of 4480 women, with higher symptom rates observed for POAI (appendix p18). Eleven patients each reported a single serious adverse reaction (appendix p19); all in the POAI group. The most common were pulmonary embolism (n=3) and musculoskeletal pain (n=3).

3913 (87%) of 4480 participants had $Ki67_B$ data available. 3206 (72%) of 4480 had paired $Ki67_B$ and $Ki67_{2W}$ data available of whom 2528 (85%) of 2976 participants were allocated to POAI and 678 (45%) of 1504 allocated control (Figure 1). In 2316 (72%) of the 3206 participants with paired Ki67 data the surgical sample was a resection (1834 (73%) of 2528 POAI, 474 (70%) of 678 control) or surgical sample type was unknown (n=6 POAI, n=2 control) and the $Ki67_{2W}$ scores for these were corrected as described above. 688 (27%) of 2528 POAI and 202 (30%) of 678 control patients' surgical sample type was core-cut biopsy.

The median $Ki67_B$ score in the 3913 of 4480 patients with a sample available was 15.2% (IQR: 8·6, 26·0). $Ki67_B$ values were different between HER2-negative and HER2-positive tumours (HER2-negative median 14·3% (IQR: 8·2, 24·6), HER2-positive median 26·6% (IQR: 17·0, 37·4); p-value for difference<0·0001). After two weeks of POAI, Ki67 was significantly suppressed compared with little change in the control group. $Ki67_{2W}$ was markedly lower in the HER2-negative tumours compared with a HER2-positive (appendix p3). In the control group, as expected given the little overall change, $Ki67_{2W}$ was again lower in the HER2-negative tumours than in those which were HER2-positive (appendix p3).

For those women who had HER2-negative cancers and allocated POAI (n=2235 of 2528 patients), 209 (9%) TTR events were reported. Women (n=732 (33%) of 2235 patients) with $Ki67_B$ <10% had a better prognosis than those (n=1503 (67%) of 2235 patients) with a $Ki67_B$ of ≥10% (appendix p20). Women whose $Ki67_{2W}$ remained high (H-H group) were significantly more likely to have a recurrence than those whose $Ki67_{2W}$ had dropped below 10% (H-L group) (unadjusted HR=2·59 (95%CI: 1·93, 3·47); p<0·0001, adjusted HR=2·10 (95%CI: 1·48, 2·98); p<0·0001) (Figure 3a). Adding a high vs low classification at two weeks segregated groups with relation to their baseline Ki67 (appendix p21).

In the HER2-negative POAI-treated sub-population exploratory analyses relating to the combined effects of age and chemotherapy use suggested that in patients with $Ki67_B$≥10% who did not receive adjuvant chemotherapy the residual $Ki67_{2W}$ (High or Low) conferred a differential impact on prognosis as assessed by TTR for both those aged <70 years and aged ≥70 (appendix pp4-9). Numbers were too small to fully

define effects for the corresponding group (i.e. $Ki67_B$≥10%) who did receive chemotherapy.

In patients with HER2-negative breast cancer allocated control 56 TTR events have been reported in the 597 of 678 patients for whom $Ki67_{2W}$ is available. There was no evidence of a difference in TTR between the H-H and H-L groups (appendix p10 and p22).

Exploratory analysis in the HER2-negative subgroup suggest an optimal cut-point around 15-20% for $Ki67_B$ and around 6-8% for $Ki67_{2W}$ and that using CCCA threshold for $Ki67_{2W}$ had prognostic discrimination. (appendix pp11-13).

In patients with HR+, HER2-positive breast cancer allocated to POAI (273 (10%) of 2528 patients) 33 TTR events have been reported. Women in the Ki67 H-H group (n=143) were more likely to have a recurrence than those in the H-L group (n=94) although the difference was not statistically significant in this small sub-population (unadjusted HR=2·08, (95%CI: 0·88, 4·90); p=0·093, adjusted HR=1·83, (95%CI: 0·71, 4·73); p=0·21) (Figure 3b). Similar to the HER2-negative group absolute risk of recurrence at 1, 3 and 5 years was higher in the H-H group (appendix p23). In the 70 women with HER2-positive breast cancer allocated control, only 9 TTR events were reported.

DISCUSSION

POETIC is by far the largest trial of its kind to assess the potential of POAI therapy in patients with postmenopausal HR+ early breast cancer and it did not observe any

significant long-term improvement in disease outcomes with this approach. This was despite preclinical experimental evidence in a mouse model suggesting the contrary.[3,4] A smaller phase 3 clinical trial, which reported after POETIC was initiated, randomised operable breast cancer patients (n=976, 50% HR+, 45% HR- and 5% HR unknown) to surgery or an intramuscular injection of depot hydroxyprogesterone 500mg 5-14 days before surgery; this showed no significant benefit in the overall population (HR=0·87 (95%CI 0·68, 1·09); p=0·23) but suggested a hypothesis generating potential DFS improvement in positive node subgroups (HR=0·72 (95%CI: 0·54, 0·97); p=0·02)[22]. In contrast, consistent with our overall finding, POETIC showed no suggestion of long-term outcome improvement with POAI in overall or in the node positive subgroup.

In POETIC the frequency of chemotherapy was slightly lower in patients allocated POAI than control. Multivariable regression supported the suggestion that this was likely due to MDTs being influenced by pathological tumour grade which was on average lower in that group (appendix p16). This absolute difference was small however (5%), and since the overall event rate was <20% would have had an imperceptible impact on outcome comparisons.

On a pragmatic note, it is common practice to start some patients on preoperative endocrine therapy if there has to be a significant delay in surgery for whatever reason. While not showing any statistical evidence of clinical benefit, our results provide reassurance that there is no detriment to this practice.

The second aim of this trial was to explore whether the measurement of tumour Ki67 two weeks after starting treatment could predict disease outcome better than baseline Ki67 alone thus providing the basis of a simple and inexpensive test to personalise adjuvant treatment in patients with HR+ HER2-negative breast cancer. Previously, IMPACT had showed that two week on-treatment Ki67 predicted outcome better than baseline and, unlike baseline, was significant in multivariable analysis[12]. POETIC has provided the definitive evidence for the clinical validity of on-treatment (AI) $Ki67_{2W}$ in addition to $Ki67_{B}$ to predict those with high residual risk of recurrence in spite of standard-of-care therapy. At the initiation of POETIC we believed the evidence was insufficient to withhold or direct therapy based on the $Ki67_{2W}$. Results provide an early indication of endocrine sensitivity or resistance including for the large number of patients who are not routinely considered for adjuvant chemotherapy.

Separate clearly defined adjuvant treatment pathways for HER2-positive and HER2-negative breast cancers now exist and thus we analysed these groups separately when considering prognostic risk. The HER2-positive subgroup was small with relatively few events. Focus for exploratory analysis was therefore on the HER2-negative subgroup which comprised approximately 90% of the POETIC population.

Previously it had been shown that patients with a low $Ki67_{B}$ have a better prognosis than those with a high $Ki67_{B}$ level[23]. POETIC confirmed this in a much larger prospective population, dichotomizing $Ki67_{B}$ at 10% with 5 year recurrence risk in HER2-negative patients allocated POAI of 4·4% and 11·8% respectively. To our knowledge this is the first large published dataset using the Ki67 scoring methodology recommended by the International Ki67 in Breast Cancer Working Group this result

served as a clinical validation of that methodology[15]. Patients whose $Ki67_B$ was low did well on current standard of care, with approximately 85% of those receiving endocrine therapy alone. It could be that if the patient's clinical pathological features led to chemotherapy being given this may have contributed to the good outcomes. But irrespective of adjuvant treatment it is reasonable to conclude that $Ki67_{2w}$ did not add significant prognostic or predictive information in this subgroup.

In contrast, for patients whose tumours had a high baseline Ki67 allocated POAI, 73% had a low $Ki67_{2w}$ two weeks after starting treatment; those patients had a better prognosis at 5 years than those who continued to have a high $Ki67_{2W}$ (8·4% vs. 21·5% 5 year recurrence risk). To what extent could this observation be applied to clinical practice?

The answer to this is influenced by the limitations of this trial The first concerns the optimal cut-off for Ki67, and we have shown that dichotomising for cut-offs other than 10% merit further exploration. The second limitation concerns interpreting the data in relation to age and chemotherapy usage. Older age has already been shown to be an independent prognostic factor in breast cancer[24] and in POETIC patients aged≥70 had poorer outcomes than those aged <70. Since a significant minority (26%) of POAI patients had adjuvant chemotherapy, there is a potential confounding in the interpretation of $Ki67_{2W}$ in relation to prognosis and prediction of the value of endocrine therapy alone. To address this, we repeated our analyses in patients according to their receipt of adjuvant chemotherapy. This confirmed a persisting significantly worse outcome for tumours H-H after 2 weeks of an AI compared with H-L in the 74% of patients not receiving chemotherapy. In the corresponding groups who received

chemotherapy, numbers were insufficient to determine a prognostic Ki67 effect or to define a plausible beneficial chemotherapy effect.

In the two-thirds of patients <70 not receiving chemotherapy the overall outcome was better, probably reflecting the choice of omitting chemotherapy for better prognosis patients. But the key point was that in this population of patients non-confounded by chemotherapy, 21% with high $Ki67_B$ remained high at surgery (H-H) and those had 11·2% 5-year recurrence risk (arguably meriting chemotherapy in addition), compared with the L-L and H-L groups where recurrence by 5 years was only 1·6% and 2·9% (indicating that additional chemotherapy would be of no clinically relevant benefit).

Similar findings were observed for patients aged ≥70. Only 59 of those patients received chemotherapy, too few to provide statistical confidence in the relationship between Ki67 and outcome. In those aged ≥70 years who did not receive chemotherapy, there was again a large difference in outcome between the H-L and H-H groups (5-year recurrence risk 12·3% vs 34·5%), again strongly supporting the discriminatory power of measuring Ki67 at two weeks, even though the absolute risks were greater.

The prespecified $Ki67_{2w}$ 10% cut-point was chosen for consistency with on-going clinical trials (ALTERNATE (NCT*01953588),* ADAPT (NCT 01779206)). The relationship of $Ki67_{2w}$ with recurrence risk is continuous and as illustrated by our analysis using CCCA, other cut-points may be selected if appropriate for a specific use (e.g. assessing the value of well-tolerated additional treatment).

In conclusion, in POETIC, giving perioperative endocrine therapy with an AI had no significant effect on long-term outcome. The trial also showed that using $Ki67_B$ and AI on-treatment $Ki67_{2w}$ could help guide adjuvant treatment decisions. First, we believe we have identified a subgroup with a low baseline Ki67 who have a sufficiently good prognosis that the majority will do well on standard endocrine therapy alone (except perhaps for a minority as dictated by other clinic-pathological factors) and who do not require a repeat two week biopsy. Second, giving POAI to the subgroup with high baseline Ki67 can differentiate two groups of patients according to their two week Ki67 value: those who convert to a low Ki67 may not need anything beyond adjuvant endocrine therapy (taking consideration of other clinic-pathological factors), while those with a Ki67 which has remained high should be considered for further adjuvant treatment(s) and trials. There are of course now several commercially available genomic platforms developed to provide the same kind of prognostic and predictive information for the individual patient[8,9]. But these tests are expensive, they often involve central testing of tissue which has to be sent long distances with inevitable time delay, and results can differ between the platforms. Ki67 as used in POETIC offers an inexpensive, easy and quick alternative where genomic testing is not readily available.

Contributors

IS- Chief Investigator, trial design, protocol development, participant recruitment, data collection, data interpretation, writing, Trial Management Group member

JR- trial design, protocol development, participant recruitment, data collection, data interpretation, writing, Trial Management Group member

LK- statistical analysis, data interpretation, writing, Trial Management Group member

MW- patient advocate, Trial Management Group member

AE- participant recruitment, data collection, Trial Management Group member

CH- participant recruitment, data collection, Trial Management Group member

KH- participant recruitment, data collection, Trial Management Group member

CK- data collection, data analysis, data interpretation, Trial Management Group member

EM- data collection, data analysis, data interpretation, Trial Management Group member

MS- participant recruitment, data collection, Trial Management Group member

AS- participant recruitment, data collection, Trial Management Group member

RV- participant recruitment, data collection, Trial Management Group member

MC- data analysis, data interpretation, Trial Management Group member

JBa- trial management, data collection, data management, Trial Management Group member

KS- data collection

JM- trial design, protocol development, statistical analysis, data interpretation, Trial Management Group member

AD- data analysis, data interpretation

MD- trial design, protocol development, data analysis, data interpretation, writing, Trial Management Group member

JBl- trial design, protocol development, statistical analysis, data interpretation, writing, Trial Management Group member

All authors reviewed the manuscript prior to submission.

Research in Context

Evidence before this study

Longstanding experimental evidence from 1989 led to the hypothesis that short duration pre-surgical endocrine therapy for early oestrogen-receptor positive breast cancer might improve clinical outcome. We carried out a PubMed search from 1989 until December 31, 2019 for relevant clinical studies using the terms “neoadjuvant endocrine”, “breast cancer” “clinical trial”, “presurgical and endocrine therapy”. No reasonably sized randomised trial addressed this issue by the time POETIC commenced recruitment in 2008. Subsequently, a randomised clinical trial reported that depot progesterone for 5-14 days before surgery improved outcome in node positive early breast cancer.

Before the initiation of POETIC two small clinical neo-adjuvant trials, IMPACT and Z1031, reported that tumour Ki67 2-4 weeks after starting preoperative endocrine treatment predicted outcome better than baseline Ki67. POETIC was designed to determine whether the gain in prognostic accuracy merited routine application of presurgical endocrine therapy for this purpose. An additional PubMed search was conducted with “Ki67” added to the above search terms. One small study of low dose tamoxifen was identified but this did not substantially add to the earlier evidence. Another modest sized trial used to triage patients with 2-4 week Ki67>10% to chemotherapy and reported the long term outcome for those less than 10%,has led to a larger on-going trial. One other large on-going trial applies 10% as a cut-off at two weeks of tamoxifen or an aromatase inhibitor for directing patients to different adjuvant

therapy. The concept of complete cell cycle arrest (CCCA) has been developed as an additional possible cut-off for on-treatment Ki67.

Added value of this study

POETIC indicated that that two weeks' preoperative endocrine therapy makes no perceptible improvement in long-term outcome, but was nevertheless a safe treatment practice. The trial confirmed the low risk of recurrence for those with a low baseline Ki67. In patients with a high baseline Ki67 value (>10%) that Ki67 levels on a biopsy two weeks after starting preoperative endocrine therapy provides additional clinical utility by predicting long term outcomes. The trial reliably documents the relationship of 2-week Ki67 with risk of recurrence for estimating whether the prognosis of individual patients is sufficiently good on endocrine therapy alone or whether additional treatment such as chemotherapy or new targeted therapies should be considered.

Implications of the available evidence

The data show no reason for short-term presurgical treatment to be applied for its direct therapeutic potential. but support prescribing an aromatase inhibitor with confidence for the short-term period before breast cancer surgery in ER+ tumours with a high proliferation rate to derive information on early endocrine responsiveness in that sub-group that can be used to predict a patients' 5 year prognosis on standard adjuvant therapy. The clinical manoeuvres to achieve this are straightforward and the measurement of Ki67 is inexpensive, making this an attractive, cost-effective approach to estimating the prognosis of the majority of early breast cancer patients.

Data sharing statement

The ICR-CTSU supports the wider dissemination of information from the research it conducts, 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 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 undertaken 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 may lead to deductive disclosures will be removed in line with Cancer Research UK Data Sharing Guidelines. Additional documents may be shared if approved by the TMG and Trial Steering Committee, e.g. statistical analysis plan and informed consent form.

Acknowledgements

We are grateful for the support from the National Institute for Health Research (NIHR) Cancer Research Network (CRN) Cancer 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 No. C1491/A15955.

The POETIC trial represents independent research supported by the National Institute for Health Research (NIHR) Biomedical Research Centre at The Royal Marsden NHS Foundation Trust and the Institute of Cancer Research, London. The views expressed are those of the author(s) and not necessarily those of the NIHR or the Department of Health and Social Care.

We 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 The Institute of Cancer Research Clinical Trials & Statistics Unit (ICR-CTSU), Ralph Lauren Centre for Breast Cancer Research at the Royal Marsden NHS Foundation Trust and 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.

POETIC is co-sponsored by The Institute of Cancer Research and The Royal Marsden NHS Foundation Trust.

References

1. Dowsett M, Smith IE, Ebbs SR, 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. Clinical cancer research : an official journal of the American Association for Cancer Research 2005; 11(2 Pt 2): 951s-8s.
2. Ellis MJ, Coop A, Singh B, et al. Letrozole inhibits tumor proliferation more effectively than tamoxifen independent of HER1/2 expression status. Cancer Res 2003; 63(19): 6523-31.
3. Fisher B, Gunduz N, Coyle J, Rudock C, Saffer E. Presence of a growth-stimulating factor in serum following primary tumor removal in mice. Cancer Res 1989; 49(8): 1996-2001.
4. Fisher B, Saffer E, Rudock C, Coyle J, Gunduz N. Effect of local or systemic treatment prior to primary tumor removal on the production and response to a serum growth-stimulating factor in mice. Cancer Res 1989; 49(8): 2002-4.
5. Haybittle JL, Blamey RW, Elston CW, et al. A prognostic index in primary breast cancer. Br J Cancer 1982; 45(3): 361-6.
6. Ravdin PM, Siminoff LA, Davis GJ, et al. Computer program to assist in making decisions about adjuvant therapy for women with early breast cancer. J Clin Oncol 2001; 19(4): 980-91.
7. Candido Dos Reis FJ, Wishart GC, Dicks EM, et al. An updated PREDICT breast cancer prognostication and treatment benefit prediction model with independent validation. Breast Cancer Res 2017; 19(1): 58.
8. van de Vijver MJ, He YD, van't Veer LJ, et al. A gene-expression signature as a predictor of survival in breast cancer. N Engl J Med 2002; 347(25): 1999-2009.
9. Paik S, Shak S, Tang G, et al. A multigene assay to predict recurrence of tamoxifen-treated, node-negative breast cancer. N Engl J Med 2004; 351(27): 2817-26.
10. Sestak I, Buus R, Cuzick J, et al. Comparison of the Performance of 6 Prognostic Signatures for Estrogen Receptor-Positive Breast Cancer: A Secondary Analysis of a Randomized Clinical Trial. JAMA Oncol 2018; 4(4): 545-53.
11. Dowsett M, Smith IE, Ebbs SR, et al. Proliferation and apoptosis as markers of benefit in neoadjuvant endocrine therapy of breast cancer. Clinical cancer research : an official journal of the American Association for Cancer Research 2006; 12(3 Pt 2): 1024s-30s.
12. Dowsett M, Smith IE, Ebbs SR, 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.
13. Ellis MJ, Tao Y, Luo J, et al. Outcome prediction for estrogen receptor-positive breast cancer based on postneoadjuvant endocrine therapy tumor characteristics. J Natl Cancer Inst 2008; 100(19): 1380-8.
14. Ellis MJ, Suman VJ, Hoog J, 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.
15. Leung SCY, Nielsen TO, Zabaglo L, et al. Analytical validation of a standardized scoring protocol for Ki67: phase 3 of an international multicenter collaboration. NPJ Breast Cancer 2016; 2: 16014.
16. Pinhel IF, Macneill FA, Hills MJ, et al. Extreme loss of immunoreactive p-Akt and p-Erk1/2 during routine fixation of primary breast cancer. Breast Cancer Res 2010; 12(5): R76.
17. Baum M, Budzar AU, Cuzick J, et al. 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. Lancet (London, England) 2002; 359(9324): 2131-9.
18. Thurlimann B, Keshaviah A, Coates AS, et al. A comparison of letrozole and tamoxifen in postmenopausal women with early breast cancer. N Engl J Med 2005; 353(26): 2747-57.
19. Harrell FE, Jr., Lee KL, Califf RM, Pryor DB, Rosati RA. Regression modelling strategies for improved prognostic prediction. Stat Med 1984; 3(2): 143-52.
20. Ma CX, Gao F, Luo J, 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. Clinical cancer research : an official journal of the American Association for Cancer Research 2017; 23(15): 4055-65.
21. Bliss J, Morden J, Evans A, et al. Abstract P2-05-01: Clinico-pathological relationships with Ki67 in POETIC (CRUK/07/015) – Critical lessons for assessing Ki67 for prognosis and as a pharmacodynamic marker. Cancer Research 2017; 77(4 Supplement): P2-05-1-P2--1.
22. Badwe R, Hawaldar R, Parmar V, et al. Single-injection depot progesterone before surgery and survival in women with operable breast cancer: a randomized controlled trial. J Clin Oncol 2011; 29(21): 2845-51.
23. Urruticoechea A, Smith IE, Dowsett M. Proliferation marker Ki-67 in early breast cancer. J Clin Oncol 2005; 23(28): 7212-20.
24. Host H, Lund E. Age as a prognostic factor in breast cancer. Cancer 1986; 57(11): 2217-21.

Declaration of Interest Statement

Prof. Dowsett reports grants from Cancer Research UK, during the conduct of the study; personal fees from Radius, personal fees from Roche, personal fees from Myriad, personal fees from Orion, personal fees from G1 Therapeutics, personal fees from Nanostring, personal fees from AbbVie, personal fees from H3 Biomedicine, personal fees from Lilly, personal fees from ICR Rewards for Inventors Scheme, outside the submitted work.

Prof. Bliss reports grants from Cancer Research UK, other from NanoString Technologies Inc, during the conduct of the study; grants and non-financial support from AstraZeneca, grants and non-financial support from Merck Sharp & Dohme, grants from Medivation, grants and non-financial support from Puma Biotechnology, grants and non-financial support from Clovis Oncology, grants and non-financial support from Pfizer, grants and non-financial support from Janssen-Cilag, grants and non-financial support from Novartis, grants and non-financial support from Roche, outside the submitted work.

Mrs Kilburn reports grants from Cancer Research UK, during the conduct of the study

The other authors declared no conflicts of interest.

List of headings for figures:

Figure 1: CONSORT diagram

Figure 2: Kaplan Meier survival curve by randomised treatment group for (A) time to recurrence and (B) overall survival

Figure 3: Kaplan Meier survival curve for time to recurrence by $Ki67_B$ and $Ki67_{2W}$ for patients allocated perioperative AI with (A) HR+ and HER2-negative breast cancer and (B) HR+ and HER2-positive breast cancer

Footnote: Patients in the L-H group are omitted from the figures (A) 28 patients, (B) 4 patients