Short Communication:

Title: Discrepancies in central review re-testing of patients with ER positive and HER2 negative breast cancer in the OPTIMA prelim randomised clinical trial.

Authors: SE Pinder[1], AF Campbell[2], JMS Bartlett[3], A Marshall[2], D Allen[4] M Falzon[5], JA Dunn[2], A Makris[6], L Hughes-Davies[7], RC Stein[8] on behalf of OPTIMA Trial Management Group.

1. Division of Cancer Studies, King's College London, Guy's Hospital, Great Maze Pond, London, UK. SE1 9RT. Tel. +44(0)20 7188 4260.
2. Warwick Medical School, University of Warwick, Gibbet Hill Campus, Coventry, UK. CV4 7AL. +44(0)24 7615 1948
3. Ontario Institute of Cancer Research, Toronto, ON, Canada
4. UCL-Advanced Diagnostics, University College London, London, UK
5. Department of Pathology, University College London Hospitals, London, UK
6. Department of Clinical Oncology Mount Vernon Cancer Centre, Mount Vernon Hospital, Northwood, Middlesex. UK. HA6 2RN. Tel: +44 203 826 2165.
7. Oncology Centre, Addenbrooke's Hospital, Cambridge University Hospitals NHS Foundations Trust, Cambridge, UK. Tel: +44 1223 245151
8. Department of Oncology, University College London Hospitals, London, UK .

Corresponding author:

Professor Sarah E Pinder

Professor of Breast Pathology

Division of Cancer Studies

King's College London

3rd Floor, Bermondsey Wing

Guy's Hospital

Great Maze Pond

LONDON

SE1 9RT

Tel: 020 7188 4260

Sarah.pinder@kcl.ac.uk

Running title: Discrepancies in receptors in OPTIMA prelim

Abstract:

Background: There is limited data on results of central re-testing of samples from patients with invasive breast cancer categorised in their local hospital laboratories as oestrogen receptor (ER) positive and human epidermal growth factor receptor homologue 2 (HER2) negative.

Methods: The Optimal Personalised Treatment of early breast cancer usIng Multiparameter Analysis preliminary study (OPTIMA prelim) was the feasibility phase of a randomised controlled trial to validate the use of multiparameter assay directed chemotherapy decisions in the UK National Health Service (NHS). Eligibility criteria included ER positivity and HER2 negativity. Central re-testing of receptor status was mandatory.

Results: Of the 431 patients tested centrally, discrepant results between central and local laboratory results were identified in only 19 (4.4%; 95% confidence interval 2.5%-6.3%) patients (with 21 tumours). On central review, seven patients had cancers that were ER negative (1.6%) and/or HER2 positive (13 (3.2%) patients with 15 tumours); including one tumour discrepant for both biomarkers.

Conclusion: Central re-testing of receptor status of invasive breast cancers in the UK NHS setting shows a high level of reproducibility in categorising tumours as ER positive and HER2 negative and raises questions regarding the costs and the value of central re-testing in this sub-group of breast cancers in this setting.

Keywords: Breast cancer; oestrogen receptor; HER2

Introduction

Oestrogen receptor (ER) and human epidermal growth factor receptor homologue 2 (HER2) are established biomarkers in invasive breast cancer and form the backbone of clinical decision making related to targeted therapies in the adjuvant setting. Although data from external quality assurance schemes (such as UK NEQAS ICC), successful participation in which is mandatory for UK laboratories, indicates good performance for testing these receptors nationally, there is relatively little published evidence comparing local results to central re-testing of local ER and HER2 expression in large clinical trial datasets. In particular, information from central laboratory testing/validation of series of invasive breast carcinomas that have been designated as ER positive and HER2 negative is limited; reports have largely described data from central re-testing of breast cancers which have been recorded as HER2 positive in local laboratories, such as in the Breast Intergroup Trial N9831 [Roche 2002][Suman 2006]. However, some of these publications have indicated alarming proportions of discrepancy in defining HER2 positivity. There are fewer publications comparing central repeat testing of hormone receptors from clinical trial samples but Viale et al examined 6291 of 8010 tumours from women in BIG1-98 and found that central review confirmed 97% of tumours were hormone receptor-positive (defined as ER and/or PgR > or = 10%) [Viale 2007]. Using tissue microarrays (TMAS) of tumours in the Tamoxifen and Exemestane Adjuvant Multinational (TEAM) trial, of 4325 cases with sufficient material only 42 were ER negative (0.99%), of these 28 were PgR positive and only 14 ER negative/PgR negative tumours were identified (0.3%) [Bartlett 2011].

The accuracy of defining hormone receptor positive and HER2 negative invasive breast cancer in local centres is clearly vital for patient management outside of the clinical trial setting, but also has significant resource and cost implications within randomised trials where ER and/or HER2 are critical components of eligibility. The question remains whether local biomarker results are sufficiently robust to allow trialists to avoid the costly re-analysis of biomarkers in central laboratories to confirm patient eligibility. To address this question we have examined data in the UK setting within OPTIMA prelim.

Material and methods:

The Optimal Personalised Treatment of early breast cancer usIng Multiparameter Analysis preliminary study (OPTIMA prelim) (ISRCTN42400492) was the feasibility phase of a randomised controlled trial designed to validate the use of multiparameter assay directed chemotherapy decisions in the UK National Health Service [Bartlett 2013][Stein 2016][Bartlett JM et al. 2016]. Patients were aged ≥ 40 years at entry with surgically treated ER positive, HER2 negative primary invasive breast cancer, with 1 to 9 involved axillary nodes or, if node negative, a tumour of at least 30mm in maximum dimension. Patients were randomised to standard care (chemotherapy followed by endocrine therapy) or an Oncotype DX® test (Genomic Health Inc., Redwood City, CA, USA) was performed on the surgically resected tumour to assign patients either to standard care (if 'recurrence score' (RS) was > 25), or to endocrine therapy alone (if RS was ≤ 25). In this feasibility study, ER and HER2 were both reassessed by a central laboratory (UCL Advanced Diagnostics) after registration into the trial to confirm eligibility prior to randomisation.

ER was assessed centrally on whole tissue sections by immunohistochemistry (6F11; Leica Biosystems) and an Allred score of 3 or more was regarded as positive, as per national guidance at that time [Harvey JM et al. 1999]. If central ER results were discordant with the local report, and there was any doubt, the assay was repeated with a second antibody (EP1, Dako). HER2 was re-assessed centrally with dual-color dual-hapten brightfield in situ hybridisation (DDISH) (Ventana Medical Systems) and, as per UK national guidelines, a ratio of *Her2* to chromosome 17 centromeric probe (CEP17) of 2.00 - 2.20 was considered to represent borderline/positive gene amplification, whilst a ratio of *Her2*:CEP17 of >2.20 was regarded as *Her2* gene amplification [Bartlett et al, 2011]. If DDISH proved unsuccessful, FISH was attempted using the HER2 PathVysion probe (HER2 PathVysion; Abbott Molecular). Her-2 Immunohistochemistry (4B5; Ventana Medical Systems) was applied in cases where no result was achievable by either HER2 ISH technique.

Results

Between October 2012 and August 2014, 442 patients were registered into OPTIMA prelim, but 11 patients were subsequently withdrawn prior to central testing. Thus a total of 431 patients had their tumours tested centrally. Nineteen patients with 21 tumours, showed discrepancies in receptor status between local and central laboratory results (4.4%; 95% confidence interval (CI) 2.5%-6.3%). The remaining 412 patients (95.6%) with concordant results went on to be randomised into OPTIMA prelim.

Seven tumours in 7 patients (1.6%) were found to be ER negative on central re-testing (Table 1). Two of the 7 were heterogeneous, with an uncommon admixture of ER negative and ER positive cells identified in the surgically excised tumour. Two appear to represent true errors in local laboratory tests; as local laboratory re-testing on the same sample found the tumours to indeed be ER negative (personal communications). In one case, an interpretive difference remained between the local and central testing; the core and the excision specimen were both reassessed locally as showing low level ER expression (Allred score 3 in the core biopsy) by the local pathologist. Unfortunately, despite liaison with the laboratories it has not been possible to discover whether ER status has been re-assessed locally for other two discrepant tumours.

In total 15 tumours in 13 patients (3.0%) from the total 431 patients tested centrally were discrepant for HER2 results (Table 2). One patient had one tumour that was centrally categorised as ER negative and also showed *Her2* amplification (ratio of *Her2*:CEP17 = 3.59). Seven others also showed *Her2* amplification (ratio of *Her2*:CEP17 ranged from 2.39-3.92). An additional patient had one tumour that was *Her2* amplified and one tumour that was borderline amplified (ratio of *Her2*:CEP17 = 2.78 and 2.11, respectively). The remaining four patients had tumours showing borderline *Her2* gene amplification (ratio between 2.00-2.20); including one patient with two tumours both showing borderline amplification. Only three of the 15 tumours demonstrated what some consider 'high-level' gene amplification (ratio >3.00) [Starczynski 2012] and none what others have described as 'high-grade' amplification (ratio >/=4.00) [Seol 2012].

It has not been possible to ascertain if there has been repeat HER2 testing (immunohistochemistry, or FISH or DDISH) on all of these 13 cases; for 5 women (with 7 tumours) the local team have managed the patient as per the central, HER2 positive, results without apparent retesting. In 3 further cases data has not been obtainable. In 4 cases local retesting has been undertaken: in 2 cases (one by FISH, other method uncertain) the local laboratory results have remained HER2 negative (both tumours borderline amplified by Her2:CEP17 ratio centrally), i.e. results remaining discrepant; one case was agreed to be HER2 positive by re-testing locally by FISH; the final case on local repeat testing had a Her2:CEP17 ratio of >2.00 but the local pathologist maintained that the tumour should be regarded as HER2 negative because of low average Her2 copy number. One case was negative immunohistochemically but showed Her2 gene amplification (3.92).

Discussion

Central re-testing of HER2 positive breast cancers has shown high levels of variability in some clinical trials; for example, HER2 positivity was only confirmed in 85.8% of 2,535 patients in the North Central Cancer Treatment Group N9831 intergroup adjuvant trial [Perez EA et al 2006]. Some of these trials, however, pre-date stringent guidelines for HER2 assessment and reporting and the reasons for discordance is often not clear. The value of central re-testing of breast cancers defined locally as HER2 negative as an eligibility criterion for other, more recent, clinical trials has not been well studied. Outwith clinical trials generally lower degrees of discrepancy, have been reported [Vani K 2008][Kaufman 2014]; for example, Kaufman et al identified that only 4% of 552 patients with metastatic HER2 negative carcinoma (defined locally) in a large observational cohort were HER2 positive on central re-testing [Kaufman 2014]. These data are essentially similar to the results in our UK clinical trial where 4.4% of tumours defined locally as HER2 negative were HER2 positive on central re-analysis.

These data highlight that approximately 3% of patients could be being excluded from HER2 directed therapies due to a potentially faulty local result in real-world testing in the UK. However, of note, we report here the proportion of cases that are *discordant* between local and central laboratory testing. Although for 2 cases, repeat re-testing of the same samples locally confirmed the tumour was ER negative (rather than ER positivity as initially reported), for others it is only possible to record that the other results were “discordant”. It is not per se the case that the central laboratory is correct and the local laboratory inaccurate, since both adhere to the same quality assurance and reporting guidelines.

Central repeat testing of hormone receptor status from clinical trial samples have reported similar, albeit slightly higher, levels of difference between local and central laboratories than we have found. Viale et al. [Viale 2007] examined 6291 of 8010 tumours from women in BIG1-98 and found that central review confirmed 97% of tumours were hormone receptor-positive, although this incorporated both ER and progesterone receptor and with different cut-offs than applied as routine in the UK (i.e. defined as ER and/or PgR > or =10%). Indeed, the authors note that, of 105 carcinomas that were reported locally as ER *negative*, 73 had >10%, and eight had 1% - 9% positive cells. This highlights the difficulty of non-standard definitions globally for hormone receptor positivity and the need for pathologists, as well as all other members of the multidisciplinary team, to be aware of study protocols and definitions.

Notwithstanding that these results compare favourably to the (albeit limited) published data, there are a number of possible explanations for discrepant results between local and central laboratories. Additional challenges include variation in methodology (for example, immunohistochemistry Vs fluorescence *in situ* hybridisation (FISH) Vs chromogenic *in situ* hybridisation (DDISH) for assessment of HER2 status), as well as differences in the antibody clones used, variation in the material assessed (cores Vs surgical excision specimens) and pathologist interpretation. It is well recognised that variation between core biopsy specimens and surgical excision is uncommon (<2% of cases showing heterogeneity) [Arnedos 2009][Lee 2012], although this clearly does occur and may potentially explain variations in receptor status if different specimens are submitted for central testing than examined locally. Indeed, this variation appears to explain at least 2 of the 7 cases with discrepant ER status in this study.

These potential discrepancies are all applicable even if the central review is undertaken in ‘real time’, i.e. prior to patient randomisation as in OPTIMA prelim. Nevertheless, particular care must be taken when analysing historical data on ER status, even in meta-analysis of clinical trials or when comparing to present day results; data extracted from local reports may be based on entirely different methodologies; Collins et al examined (on TMA) 1851 cases where tissue and histology reports were available and highlighted that in 82% of the cases the original assays were biochemical. Even where immunohistochemistry was applied both locally and centrally as the technique of choice, agreement was only 92% for ER status (310 of 336 specimens) [Collins 2008]. Again, the 1.6% difference seen in OPTIMA prelim compares favourably.

Despite all the potential technical and interpretive differences in biomarker analysis, the results from OPTIMA prelim indicate good concordance between local laboratories and central re-testing centre in the UK in classification of invasive breast cancers as ER positive and HER2 negative. Such re-testing in large randomised clinical trials recruiting thousands of patients is very expensive and, in the setting of this group of patients (as opposed to HER2 positive disease, for example, where discrepancies may be higher), the value is questionable.

References:

Arnedos M, Nerurkar A, Osin P, A'Hern R, Smith IE, Dowsett M (2009). Discordance between core needle biopsy (CNB) and excisional biopsy (EB) for estrogen receptor (ER), progesterone receptor (PgR) and HER2 status in early breast cancer (EBC). Ann Oncol. 20:1948-52

Bartlett JM, Starczynski J, Atkey N, Kay E, O'Grady A, Gandy M, Ibrahim M, Jasani B, Ellis IO, Pinder SE, Walker RA. **HER2** testing in the UK: recommendations for breast and gastric in-situ hybridisation methods. J Clin Pathol. 2011 Aug;64(8):649-53.

Bartlett JM, Brookes CL, Robson T, van de Velde CJ, Billingham LJ, Campbell FM, Grant M, Hasenburg A, Hille ET, Kay C, Kieback DG, Putter H, Markopoulos C, Kranenbarg EM, Mallon EA, Dirix L, Seynaeve C, Rea D. (2013) 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. 29:1531-8.

Bartlett, J, Canney, P, Campbell, A, Cameron, D, Donovan, J, Dunn, J, Earl, H, Francis, A, Hall, P, Harmer, V, Higgins, H, Hillier, L, Hulme, C, Hughes-Davies, L, Makris, A, Morgan, A, Mccabe, C, Pinder, S, Poole, C, Rea, D, Stallard, N & Stein, R. (2013) Selecting breast cancer patients for chemotherapy: the opening of the UK OPTIMA trial. *Clin Oncol (R Coll Radiol)* 25(2): 109-16.

Bartlett JM, Bayani J, Marshall A, Dunn JA, Campbell A, Cunningham C, Sobol MS, Hall PS, Poole CJ, Cameron DA, Earl HM, Rea DW, Macpherson IR, Canney P, Francis A, McCabe C, Pinder SE, Hughes-Davies L, Makris A, Stein RC; OPTIMA TMG (2016) Comparing Breast Cancer Multiparameter Tests in the OPTIMA Prelim Trial: No Test Is More Equal Than the Others. J Natl Cancer Inst 108(9). pii: djw050. doi: 10.1093/jnci/djw050.

Collins LC, Marotti JD, Baer HJ, Tamimi RM (2008) Comparison of estrogen receptor results from pathology reports with results from central laboratory testing. J Natl Cancer Inst. 100(3):218-21.

Harvey JM, Clark GM, Osborne CK, Allred DC (1999) Estrogen receptor status by immunohistochemistry is superior to the ligand-binding assay for predicting response to adjuvant endocrine therapy in breast cancer. J Clin Oncol. 17(5):1474-81

Kaufman PA, Bloom KJ, Burris H, Gralow JR, Mayer M, Pegram M, Rugo HS, Swain SM, Yardley DA, Chau M, Lalla D, Yoo B, Brammer MG, Vogel CL (2014). Assessing the discordance rate between local and central HER2 testing in women with locally determined HER2 negative breast cancer. Cancer. 120(17):2657-64

Lee AH, Key HP, Bell JA, Hodi Z, Ellis IO (2012). Concordance of HER2 status assessed on needle core biopsy and surgical specimens of invasive carcinoma of the breast. Histopathology. 60(6):880-4.

Perez EA, Suman VJ, Davidson NE, Martino S, Kaufman PA, Lingle WL, Flynn PJ, Ingle JN, Visscher D, Jenkins RB (2006) HER2 testing by local, central, and reference laboratories in specimens from the North Central Cancer Treatment Group N9831 intergroup adjuvant trial. J Clin Oncol. 24(19):3032-8

Roche PC, Suman VJ, Jenkins RB, Davidson NE, Martino S, Kaufman PA, Addo FK, Murphy B, Ingle JN, Perez EA (2002). Concordance between local and central laboratory HER2 testing in the breast intergroup trial N9831. J Natl Cancer Inst. 94(11):855-7

Hyesil Seol, Hyun Ju Lee, Yoomi Choi, Hee Eun Lee, Yu Jung Kim, Jee Hyun Kim, Eunyoung Kang, Sung-Won Kim and So Yeon Park (2012) Intratumoral heterogeneity of *HER2* gene amplification in breast cancer: its clinicopathological significance. Modern Pathology **25,** 938–948; doi:10.1038/modpathol.2012.36.

Stein RC, Dunn JA, Bartlett JM, Campbell AF, Marshall A, Hall P, Rooshenas L, Morgan A, Poole C, Pinder SE, Cameron DA, Stallard N, Donovan JL, McCabe C, Hughes-Davies L, Makris A (2016). OPTIMA prelim: a randomised feasibility study of personalised care in the treatment of women with early breast cancer. Health Technol Assess. 20(10):1-202. doi: 10.3310/hta20100.

Starczynski J, Atkey N, Connelly Y, O'Grady T, Campbell FM, di Palma S, Wencyk P, Jasani B, Gandy M, Bartlett JM; UKNEQAS (2012). HER2 gene amplification in breast cancer: a rogues' gallery of challenging diagnostic cases: UKNEQAS interpretation guidelines and research recommendations. Am J Clin Pathol. 137(4):595-605. doi: 10.1309/AJCPATBZ2JFN1QQC.

Vani K, Sompuram SR, Fitzgibbons P, Bogen SA (2008) National HER2 proficiency test results using standardized quantitative controls: characterization of laboratory failures. Arch Pathol Lab Med. 132(2):211-6]

Viale G, Regan MM, Maiorano E, Mastropasqua MG, Dell'Orto P, Rasmussen BB, Raffoul J, Neven P, Orosz Z, Braye S, Ohlschlegel C, Thürlimann B, Gelber RD, Castiglione-Gertsch M, Price KN, Goldhirsch A, Gusterson BA, Coates AS (2007). Prognostic and predictive value of centrally reviewed expression of estrogen and progesterone receptors in a randomized trial comparing letrozole and tamoxifen adjuvant therapy for postmenopausal early breast cancer: BIG 1-98. J Clin Oncol. 25(25):3846-52

Walker RA, Bartlett JM, Dowsett M, Ellis IO, Hanby AM, Jasani B, Miller K, Pinder SE (2008). HER2 testing in the UK: further update to recommendations. J Clin Pathol. 61(7):818-24

Acknowledgements: OPTIMA prelim (ISRCTN42400492) was funded by the National Institute for Health Research Health Technology Assessment Programme (project 10/34/01). The views and opinions expressed therein are those of the authors and do not necessarily reflect those of the Health Technology Assessment Programme, NIHR, NHS or the Department of Health. RCS was supported by the National Institute for Health Research, University College London Hospitals Biomedical Research Centre.

Conflicts of Interest: None to declare

Patient	CENTRAL ER RESULTS			
	ER Status	Allred score	% tumour cell positivity	Comment
A	Negative	0	0	Two clonality distinct tumours - part positive and part negative for ER. Original ER on core biopsy.
B	Negative	0	0	ER repeated in local laboratory on core biopsy using different antibody/clone and negative staining for ER confirmed.
C	Negative	0	0	ER retested locally and confirmed to be negative.
D	Negative	0	0	
E	Negative	0	0	Heterogenous tumour, at least focally ER negative.
F	Negative	0	0	
G	Negative	0	0	3 tumours: 2 eligible, 1 ineligible

Table 1: Details of the 7 patients (from 431 patients registered and tested centrally) with discrepant oestrogen receptor results.

Patient	CENTRAL HER2 RESULTS				
	HER2 STATUS	***Her2*: CEP17 ratio**	**Average CEP17 copy number per cell**	**Average *Her2* copy number per cell**	**Comments**
D	Amplified	3.59	1.10	3.95	Also ER negative on central testing
H	Borderline amplified	2.00	2.78	5.55	
I1*	Borderline amplified	2.14	1.65	3.53	
I2*	Borderline amplified	2.06	1.69	3.47	
J	Borderline amplified	2.20	1.43	3.14	
K	Amplified	2.70	2.00	5.40	
L	Amplified	2.39	1.55	3.70	
M	Amplified	2.81	1.35	3.80	Heterogeneous - testing of core (locally) and second block (centrally) showed Her2 non-amplified foci.
N	Amplified	3.23	2.80	9.05	
O1*	Amplified	2.78	1.35	3.75	
O2*	Borderline amplified	2.11	1.40	2.95	
P	Amplified	2.45	1.10	2.70	
Q	Amplified	2.64	1.08	2.83	
R	Amplified	3.92	1.85	7.25	
S	Borderline amplified	2.11	2.03	4.30	

Table 2: Details of the 15 discrepant tumours (13 patients) for human epidermal growth factor receptor homologue 2 (HER2) status.

Amplified: human epidermal growth factor receptor homologue 2 (HER2) to chromosome 17 centromeric probe (CEP17) ratio >2.20; borderline amplified: HER2 to CEP17 ratio 2.00-2.20. *G1 & G2, and N1 & N2, are tumours from the same patients respectively.