European Journal of Cancer 142 (2021) 48–62

Available online at www.sciencedirect.com

ScienceDirect

journal homepage: www.ejcancer.com

Clinical Trial

Bridging the age gap in breast cancer. Impacts of omission of breast cancer surgery in older women with oestrogen receptor positive early breast cancer. A risk stratified analysis of survival outcomes and quality of life

Lynda Wyld [a,*], Malcolm W.R. Reed [b], Jenna Morgan [a], Karen Collins [c], Sue Ward [d], Geoffrey R. Holmes [d], Mike Bradburn [e], Stephen Walters [e], Maria Burton [c], Esther Herbert [e], Kate Lifford [f], Adrian Edwards [f], Alistair Ring [g], Thompson Robinson [h], Charlene Martin [a], Tim Chater [e], Kirsty Pemberton [e], Anne Shrestha [a], Alan Brennan [d], Kwok L. Cheung [i], Annaliza Todd [a], Riccardo Audisio [j], Juliet Wright [b], Richard Simcock [b], Tracy Green [k], Deirdre Revell [k], Jacqui Gath [k], Kieran Horgan [l], Chris Holcombe [m], Matt Winter [n], Jay Naik [o], Rishi Parmeshwar [p], Julietta Patnick [q], Margot Gosney [r], Matthew Hatton [n], Alastair M. Thomson [s]

[a] *Department of Oncology and Metabolism, University of Sheffield Medical School, Beech Hill Road, Sheffield, S10 2RX, UK*
[b] *Brighton and Sussex Medical School, Falmer, Brighton, UK*
[c] *College of Health, Wellbeing and Life Sciences, Department of Allied Health Professions, Sheffield Hallam University, Collegiate Cresent Campus, Sheffield, UK*
[d] *Department of Health Economics and Decision Science, School for Health and Related Research, ScHARR, University of Sheffield, UK*
[e] *Clinical Trials Research Unit, School for Health and Related Research, University of Sheffield, UK*
[f] *Division of Population Medicine, Cardiff University, 8th Floor, Neuadd Meirionnydd, Heath Park, Cardiff, CF14 4YS, UK*
[g] *The Royal Marsden Hospital NHS Foundation Trust, London, UK*
[h] *Department of Cardiovascular Sciences and NIHR Biomedical Research Centre, University of Leicester, Cardiovascular Research Centre, The Glenfield General Hospital, Leicester, LE3 9QP, UK*
[i] *University of Nottingham, Royal Derby Hospital, Uttoxeter Road, Derby, DE22 3DT, UK*
[j] *University of Gothenberg, Sahlgrenska Universitetssjukhuset, 41345 Göteborg, Sweden*
[k] *Yorkshire and Humber Consumer Research Panel, UK*
[l] *Dept of Breast Surgery, Bexley Cancer Centre, St James's University Hospital, Leeds, LS9 7TF, UK*
[m] *Liverpool University Hospitals Foundation Trust, Prescott Street, Liverpool L7 8 XP, UK*
[n] *Weston Park Hospital, Whitham Rd, Sheffield S10 2SJ, UK*
[o] *Pinderfields Hospital, Mid Yorkshire NHS Foundation Trust, Aberford Rd, Wakefield, UK*
[p] *University Hospitals of Morecambe Bay, Royal Lancashire Infirmary Ashton Road, Lancaster, Lancashire, LA1 4RP, UK*
[q] *Cancer Epidemiology Unit, Nuffield Department of Population Health, University of Oxford, UK*

* *Corresponding author*:
E-mail address: l.wyld@sheffield.ac.uk (L. Wyld).

https://doi.org/10.1016/j.ejca.2020.10.015

[r] Royal Berkshire NHS Foundation Trust, Reading, RG1 5AN, UK
[s] Department of Surgery, Baylor College of Medicine, One Baylor Plaza, Houston, TX, USA

Received 18 June 2020; received in revised form 14 September 2020; accepted 6 October 2020

KEYWORDS
Adjuvant endocrine therapy;
Breast cancer;
Older women;
Patient-centred outcomes;
Primary endocrine therapy;
Propensity score matching;
Quality of life;
Surgery;
Survival

Abstract ***Background:*** Age-related breast cancer treatment variance is widespread with many older women having primary endocrine therapy (PET), which may contribute to inferior survival and local control. This propensity-matched study determined if a subgroup of older women may safely be offered PET.
Methods: Multicentre, prospective, UK, observational cohort study with propensity-matched analysis to determine optimal allocation of surgery plus ET (S+ET) or PET in women aged ≥70 with breast cancer. Data on fitness, frailty, cancer stage, grade, biotype, treatment and quality of life were collected. Propensity-matching (based on age, health status and cancer stage) adjusted for allocation bias when comparing S+ET with PET.
Findings: A total of 3416 women (median age 77, range 69–102) were recruited from 56 breast units—2854 (88%) had ER+ breast cancer: 2354 had S+ET and 500 PET. Median follow-up was 52 months. Patients treated with PET were older and frailer than patients treated with S+ET. Unmatched overall survival was inferior in the PET group (hazard ratio, (HR) 0.27, 95% confidence interval (CI) 0.23–0.33, $P < 0.001$). Unmatched breast cancer–specific survival (BCSS) was also inferior in patients treated with PET (HR: 0.41, CI: 0.29–0.58, $P < 0.001$ for BCSS). In the matched analysis, PET was still associated with an inferior overall survival (HR = 0.72, 95% CI: 0.53–0.98, $P = 0.04$) but not BCSS (HR = 0.74, 95% CI: 0.40–1.37, $P = 0.34$) although at 4–5 years subtle divergence of the curves commenced in favor of surgery. Global health status diverged at certain time points between groups but over 24 months was similar when adjusted for baseline variance.
Interpretation: For the majority of older women with early ER+ breast cancer, surgery is oncologically superior to PET. In less fit, older women, with characteristics similar to the matched cohort of this study (median age 81 with higher comorbidity and functional impairment burdens, the BCSS survival differential disappears at least out to 4–5 year follow-up, suggesting that for those with less than 5-year predicted life-expectancy (>90 years or >85 with comorbidities or frailty) individualised decision making regarding PET versus S+ET may be appropriate and safe to offer. The Age Gap online decision tool may support this decision-making process (https://agegap.shef.ac.uk/).
Trial registration number: ISRCTN: 46099296.

Introduction

Breast cancer is a common disease in older women with a third of cases occurring in women over 70 years of age. Overall mortality is higher among these women, but is commonly attributable to other causes than breast cancer. A US study found breast cancer to be the primary cause of death in just 23% of patients that died in their 80s, compared to 96% in women with breast cancer who died aged under 40. Breast cancer–specific mortality rates are also higher in older women, which may be due to later stage at diagnosis and suboptimal treatment. However, it is important to avoid overtreatment in the very frail for what may be an indolent disease in a patient with a very limited life expectancy [2]. Treatment needs to be tailored to health status, disease characteristics and patient preferences.

Surgery may be unnecessary for some frailer older women as short- and medium-term disease control may be achieved by use of anti-oestrogens (primary endocrine therapy, PET). In addition, the resilience of frailer older women to standard cancer treatments may be reduced, leading to a long-term deterioration of their functional capacity. Older women place a higher value on maintenance of independence and quality of life [3] compared to younger women, so they may prefer slightly less effective anticancer treatment to maintain disability-free life expectancy and quality of life [4].

A systematic review of randomised controlled trials (RCTs) evaluating the role of surgery versus PET

demonstrated no survival difference between the two treatments but inferior local control with PET [5]. Only one of the trials exclusively recruited women with ER+ breast cancer, and this trial demonstrated no survival advantage for surgery at 10 years [6]. An update of one of the trials with 28 years follow-up showed no survival advantage to surgery when all trial participants had died [7]. A recent patient-level meta-analysis of the data from these trials with longer-term follow-up demonstrated a significant survival benefit from surgery. The included trials were flawed, as they did not stratify patients according to age, fitness (all were fit for surgery under general anaesthesia) or tumour biology (particularly the ER), which may permit the identification of subgroups of women who may not benefit from surgery. They also did not restrict recruitment to women with ER+ cancer (the trials predated routine testing), so the PET group likely had 10–17% [1] of women who effectively had no active therapy, which will bias outcome towards surgery. Analysis of cohorts of women with strongly ER+ cancers treated with PET suggests that there is no survival advantage from surgery [8]. No study has looked at composite measures of health and tumour biology to select patients for surgery or PET.

A further complicating factor is that all of the RCTs used tamoxifen as the anti-oestrogen in both the S+ET and PET arms, whereas in modern practice aromatase inhibitors (AIs) are the preferred and more effective option [9]. PET may therefore be more efficacious if potential candidates are selected appropriately based on their health status and tumour biology and treated with AIs rather than tamoxifen. A previous randomised clinical trial (ESTEEM) attempted such health status stratification but failed to recruit due to lack of patient and clinician equipoise [10].

The purpose of this study was to use real-world, prospectively collected, observational data and adjust for allocation bias using propensity score matched analysis. This sought to identify whether PET may be appropriate for a subgroup of less fit, older women with ER+ breast cancer.

Methods

Ethics approval

Ethics (IRAS: 12 LO 1808) and research governance approvals were obtained. All patients (or their proxies, if cognitively impaired) gave written informed consent.

Study design

It was a prospective, multicentre, observational cohort study. Patients could participate at three levels: full, partial (no requirement to complete quality-of-life assessments) or proxy (third-party data collection if cognitively impaired). Study reporting is in line with STROBE guidelines [11].

Sites

Patients were recruited from 56 UK breast units in England and Wales (Supplemental Table ST1).

Inclusion criteria

Female patients aged ≥70 years at the time of breast cancer diagnosis. Primary unilateral or bilateral operable invasive breast cancer (TNM stages: T1-3 and some T4b, N0-2, M−0).

Exclusion criteria

Inoperable diseaseand previous breast cancer within 5 years were considered exclusion criteria. Patients lacking cognitive capacity as defined by the Mental Capacity Act were eligible if a friend or relative was willing to sign proxy consent. Surgery could be performed under local or general anaesthesia.

Baseline data collection

Women were recruited at the time of breast cancer diagnosis, before commencement of treatment.

At baseline women underwent health assessment using validated tools including:

1. Comorbidities (Charlson comorbidity index (CCI) [12]),
2. Nutrition (Abridged Patient Generated Subjective Global Assessment (aPG-SGA) [13]),
3. Physical functioning (Activities of Daily Living (ADL) [14]),
4. Complex physical functioning (Instrumental Activities of Daily Living (IADL) [15]),
5. Cognitive status (Mini Mental State Examination (MMSE [16], under licence),
6. Eastern Cooperative Oncology Group Performance Status (ECOG-PS) [17].
7. Medications.

In addition, quality of life was assessed at baseline using the European Organisation for Research and Treatment of Cancer Quality of Life Questionnaires (EORTC-QLQ)-C30 [18], a generic quality of life tool and health utility by the EQ-5D-5L [19] to monitor functional decline. (A more detailed analysis of quality-of-life outcomes are reported separately.)

Baseline cancer data were collected including: cancer type, grade, nodal status, primary size (clinical and on imaging), oestrogen, progesterone and Her-2 receptor status, and Oncotype DX score (if available). Staging for metastatic disease was performed if clinically indicated but otherwise presumed M0.

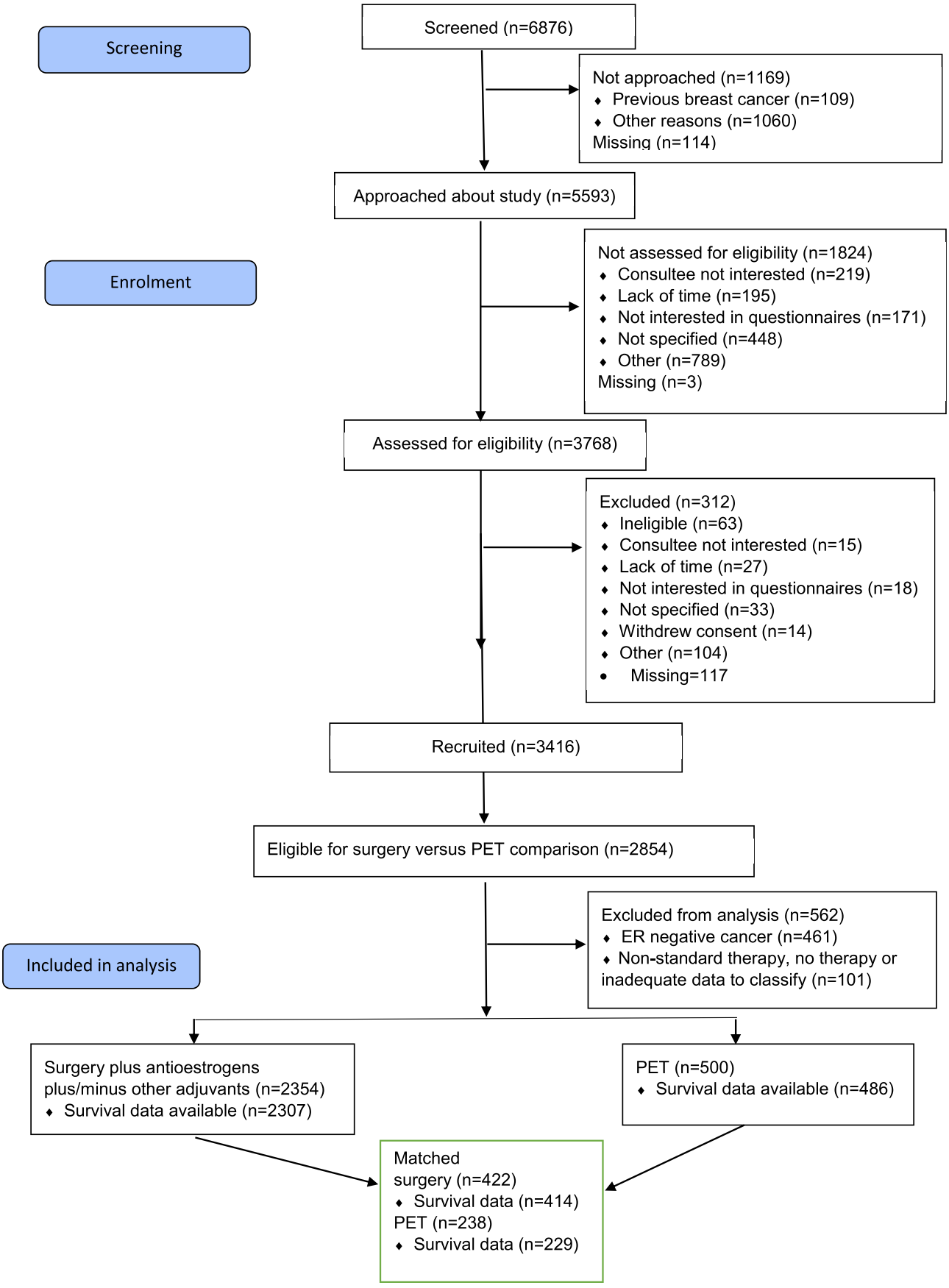

Fig. 1. STROBE diagram of patient recruitment and dispositions.

Table 1
Patient and cancer characteristics at baseline. Unmatched.

		PET	Surgery	Total
		N = 500	N = 2354	N = 2854
Age	n	500	2354	2854
	Mean (SD)	83.5 (6.5)	76.4 (5.1)	77.6 (6.0)
	Median (IQR)	84 (79, 88)	76 (72, 80)	77 (73, 82)
	Min, Max	70, 102	69, 94	69, 102
aPG-SGA score	n	322	2021	2343
	Mean (SD)	2.3 (3.1)	1.2 (2.2)	1.4 (2.4)
	Median (IQR)	1.0 (0.0, 3.0)	0 (0.0, 2.0)	0 (0.0, 2.0)
	Min, Max	0, 18	0, 17	0, 18
Barthel	n	399	2135	2534
ADL index	Mean (SD)	88.8 (16.6)	97.7 (6.2)	96.3 (9.3)
	Median (IQR)	95 (85.0, 100.0)	100 (100.0, 100.0)	100.0 (95.0, 100.0)
	Min, Max	5, 100	10, 100	5, 100
IADL index	n	382	2104	2486
	Mean (SD)	6.1 (2.1)	7.6 (0.9)	7.4 (1.3)
	Median (IQR)	7.0 (5.0, 8.0)	8.0 (8.0, 8.0)	8.0 (7.0, 8.0)
	Min, Max	0, 8	0, 8	0, 8
Modified CCI	n	459	2273	2732
	Mean (SD)	5.8 (2.0)	4.3 (1.4)	4.5 (1.6)
	Median (IQR)	6 (4.0, 7.0)	4 (3.0, 5.0)	4.0 (3.0, 5.0)
	Min, Max	3, 17	3, 13	3, 17
MMSE	n	273	1631	1904
	Mean (SD)	26.7 (3.7)	28.3 (2.5)	28.1 (2.8)
	Median (IQR)	28 (26.0, 29.0)	29 (28.0, 30.0)	29.0 (27.0, 30.0)
	Min, Max	10, 30	10, 30	10, 30
Number	n	450	2050	2500
of current medications	Mean (SD)	5.0 (3.0)	4.1 (2.6)	4.2 (2.7)
	Median (IQR)	5.0 (3.0, 7.0)	4.0 (2.0, 5.0)	4.0 (2.0, 6.0)
	Min, Max	0, 18	0, 18	0, 18
aPG-SGA	n	376	2109	2485
risk category	Low	291 (77.4%)	1848 (87.6%)	2139 (86.1%)
	Moderate	62 (16.5%)	227 (10.8%)	289 (11.6%)
	High	23 (6.1%)	34 (1.6%)	57 (2.3%)
ADL	n	408	2172	2580
risk	No dependency	191 (46.8%)	1684 (77.5%)	1875 (72.7%)
category	Mild dependency	53 (13.0%)	258 (11.9%)	311 (12.1%)
	Moderate/severe dependency	164 (40.2%)	230 (10.6%)	394 (15.3%)
IADL	N	403	2158	2561
risk category	No dependency	158 (39.2%)	1759 (81.5%)	1917 (74.9%)
	Mild dependency	64 (15.9%)	193 (8.9%)	257 (10.0%)
	Moderate/severe dependency	181 (44.9%)	206 (9.5%)	387 (15.1%)
MMSE	n	464	2286	2750
risk category	Normal function	339 (73.1%)	2037 (89.1%)	2376 (86.4%)
	Mild impairment	75 (16.2%)	206 (9.0%)	281 (10.2%)
	Moderate impairment	20 (4.3%)	29 (1.3%)	49 (1.8%)
	Severe	30 (6.5%)	14 (0.6%)	44 (1.6%)
How many clinically involved	n	483	2309	2792
nodes were detectable?	Mean (SD)	0.2 (0.6)	0,2 (0.8)	0.2 (0.7)
	Median (IQR)	0.0 (0.0, 0.0)	0.0 (0.0, 0.0)	0.0 (0.0, 0.0)
	Min, Max	0, 4	0, 20	0, 20
Size (mm)	n	487	2318	2805
	Mean (SD)	23.9 (12.0)	19.2 (12.3)	20.0 (12.4)
	Median (IQR)	21.0 (16.0, 30.0)	17.0 (11.0, 24.0)	18.0 (12.0, 25.0)
	Min, Max	0, 70	0, 150	0, 150
Nottingham Prognostic Index	n	456	2172	2628
	Mean (SD)	3.5 (0.8)	3.5 (0.8)	3.5 (0.8)
	Median (IQR)	3.4 (3.2, 3.9)	3.3 (3.2, 4.0)	3.4 (3.2, 4.0)
	Min, Max	2.1, 7	2, 6.7	2, 7
Side of primary tumour	n	500	2354	2854
	Right	223 (44.6%)	1084 (46.0%)	1307 (45.8%)
	Left	277 (55.4%)	1270 (54.0%)	1547 (54.2%)
Her 2 Score	n	359	1911	2270
	Negative	311 (86.6%)	1641 (85.9%)	1952 (86.0%)

Table 1 (*continued*)

		PET	Surgery	Total
		N = 500	N = 2354	N = 2854
	Inconclusive	14 (3.9%)	70 (3.7%)	84 (3.7%)
	Positive	34 (9.5%)	200 (10.5%)	234 (10.3%)
Provisional histological grade	n	484	2243	2727
	Grade 1	98 (20.2%)	399 (17.8%)	497 (18.2%)
	Grade 2	329 (68.0%)	1475 (65.8%)	1804 (66.2%)
	Grade 3	57 (11.8%)	369 (16.5%)	426 (15.6%)
Type of hormone therapy (for PET, at 6-weeks time point, for surgery at 6-months time point)	n	500	2354	2854
	Letrozole	414 (82.8%)	981 (41.7%)	1395 (48.9%)
	Anastrazole	31 (6.2%)	782 (33.2%)	813 (28.5%)
	Tamoxifen	22 (4.4%)	269 (11.4%)	291 (10.2%)
	Exemestane	7 (1.4%)	35 (1.5%)	42 (1.5%)
	Missing	26 (5.2%)	287 (12.2%)	313 (10.9%)

Abbreviations. aPG-SGA: abridged patient generated subjective global assessment. ADL: activities of global living. IADL: instrumental activities of global living. MMSE: mini mental state examination. CCI: Charlson comorbidities index. ECOG-PS: Eastern cooperative oncology group performance status. NPI: Nottingham prognostic index. SD: Standard deviation. IQR: interquartile range.

Type of breast surgery was categorised into breast conserving surgery (BCS) (wide excision, therapeutic mammoplasty) and mastectomy (+/− reconstruction).

Axillary surgery was categorised into no axillary surgery (NS), sentinel lymph node biopsy (SLNB, axillary sample) and axillary node clearance (ANC). Post-operative histology was recorded for patients undergoing surgery. Detailed reporting of surgery types and outcomes are reported separately [20].

Use of radiotherapy and systemic therapy (types and adverse events) was recorded when used in the S+ET patient cohort as part of standard care.

Survival and recurrence outcomes

Patients were followed up at 1.5, 6, 12, 18 and 24 months. All patients were assessed for evidence of local, regional or metastatic recurrence at each visit both clinically and using appropriate imaging and biopsy techniques depending on the location of recurrence.

Progression was defined using the RECIST criteria (response evaluation criteria in solid tumours [21]) for PET patients to allow us to accurately define when local failure occurred. The time to progression and the time to metastatic progression were both calculated starting

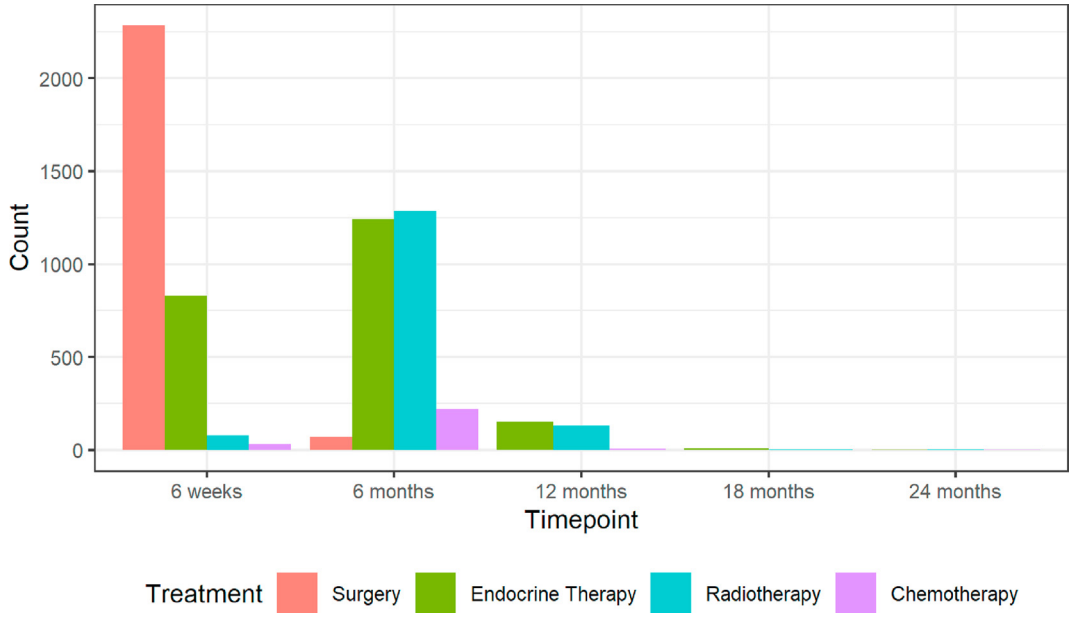

Fig. 2. **Temporal summary of surgically treated patients and the adjuvant therapies they received at each follow-up time point.** In line with UK treatment targets the majority of women had their surgery between baseline and 6 weeks, with many commencing their post-surgical adjuvant endocrine therapy at their 6-week follow-up visit. Radiotherapy was usually given between 6 weeks and 6 months. A small number of women had neoadjuvant chemotherapy starting between baseline and 6 weeks, but the majority of those having chemotherapy started it between 6 weeks and 6 months. These timelines are important for understanding the impact of therapies on quality of life and functional outcomes.

Table 2
Covariate balance in the final matched dataset: Surgery (plus adjuvant therapies) versus primary endocrine therapy (PET).

		Surgery	PET
		N = 422	N = 238
Age	n	422	238
	Mean (SD)	80.57 (5.36)	81.30 (5.94)
	Median (IQR)	81.00 (76.00, 84.00)	82.00 (77.00, 85.75)
	Min, Max	70, 94	70, 96
aPG-SGA	Low	343 (81.3%)	190 (79.8%)
	Moderate	62 (14.7%)	39 (16.4%)
	High	17 (4.0%)	9 (3.8%)
ADL	No dependency	271 (64.2%)	134 (56.3%)
	Mild dependency	54 (12.8%)	34 (14.3%)
	Moderate/severe dependency	97 (23.0%)	70 (29.4%)
iADL	No dependency	250 (59.2%)	124 (52.1%)
	Mild dependency	73 (17.3%)	38 (16.0%)
	Moderate/severe dependency	99 (23.5%)	76 (31.9%)
MMSE	Normal function	351 (83.2%)	191 (80.3%)
	Mild impairment	60 (14.2%)	39 (16.4%)
	Moderate impairment	8 (1.9%)	6 (2.5%)
	Severe	3 (0.7%)	2 (0.8%)
CCI	0	312 (73.9%)	161 (67.6%)
	2	110 (26.1%)	77 (32.4%)
ECOG-PS	Low	370 (87.7%)	194 (81.5%)
	Moderate	36 (8.5%)	27 (11.3%)
	High	16 (3.8%)	17 (7.1%)
Medications	3 or fewer	176 (41.7%)	100 (42.0%)
	4 or more	246 (58.3%)	138 (58.0%)
NPI	Moderate	217 (51.4%)	125 (52.5%)
	Good	194 (46.0%)	107 (45.0%)
	Poor	11 (2.6%)	6 (2.5%)

Abbreviations. aPG-SGA: abridged patient generated subjective global assessment. ADL: activities of global living. IADL: instrumental activities of global living. MMSE: mini mental state examination. CCI: Charlson comorbidities index. ECOG-PS: Eastern cooperative oncology group performance status. NPI: Nottingham prognostic index. SD: Standard deviation. IQR: interquartile range.

from the date of initial assessment. Patients with no recorded recurrence were censored at the date of last tumour assessment.

Deaths were adjudicated as breast cancer related or other causes blind to treatment. Survival outcomes were obtained directly via follow-up to 2 years and beyond 2 years via the UK cancer registry (with consent) and patient notes. Overall survival (OS) was defined as the time from initial assessment to death or censored at the date last known to be alive, and breast cancer–specific survival (BCSS) was defined as the time from initial assessment to breast cancer death, or censored either at the date of non-breast-cancer death or the date last known to be alive.

Adverse events

Complications were categorised using Common Terminology Criteria for Adverse Events [22] system (CTCAE) grouped into systemic (atelectasis, stroke, infarction, DVT/embolism, arrhythmia, allergic reaction and somnolence) and local (lymphoedema, neuropathy, functional difference, wound pain, wound necrosis, infection, haematoma/haemorrhage, seroma) complications.

Quality of life

Quality of life was recorded in fully participating patients at baseline and at 1.5, 6, 12, 18 and 24 months using the validated EORTC tools listed above which were scored according to the EORTC Scoring Manuals (3rd Edition [23]) and reference publications. Partially completed items were managed according to the EORTC manual recommendations. In addition, the EQ-5D-5L score was assessed at each visit to monitor health utility with particular emphasis on functional outcomes, which are of particular importance to women in this age group [24] as older women may lack resilience.

Functional resilience following surgery

Several domains in the quality-of-life tools reflect functional status including 3 in the EQ-5D and items 1–7 in the EORTC-QLQ-C30. These were assessed at baseline and at intervals after breast cancer treatment to determine the resilience of these older women to therapy. Data from patients with a complete series of scores at all-time points were compared at baseline and follow-up, to determine the functional impacts of treatments.

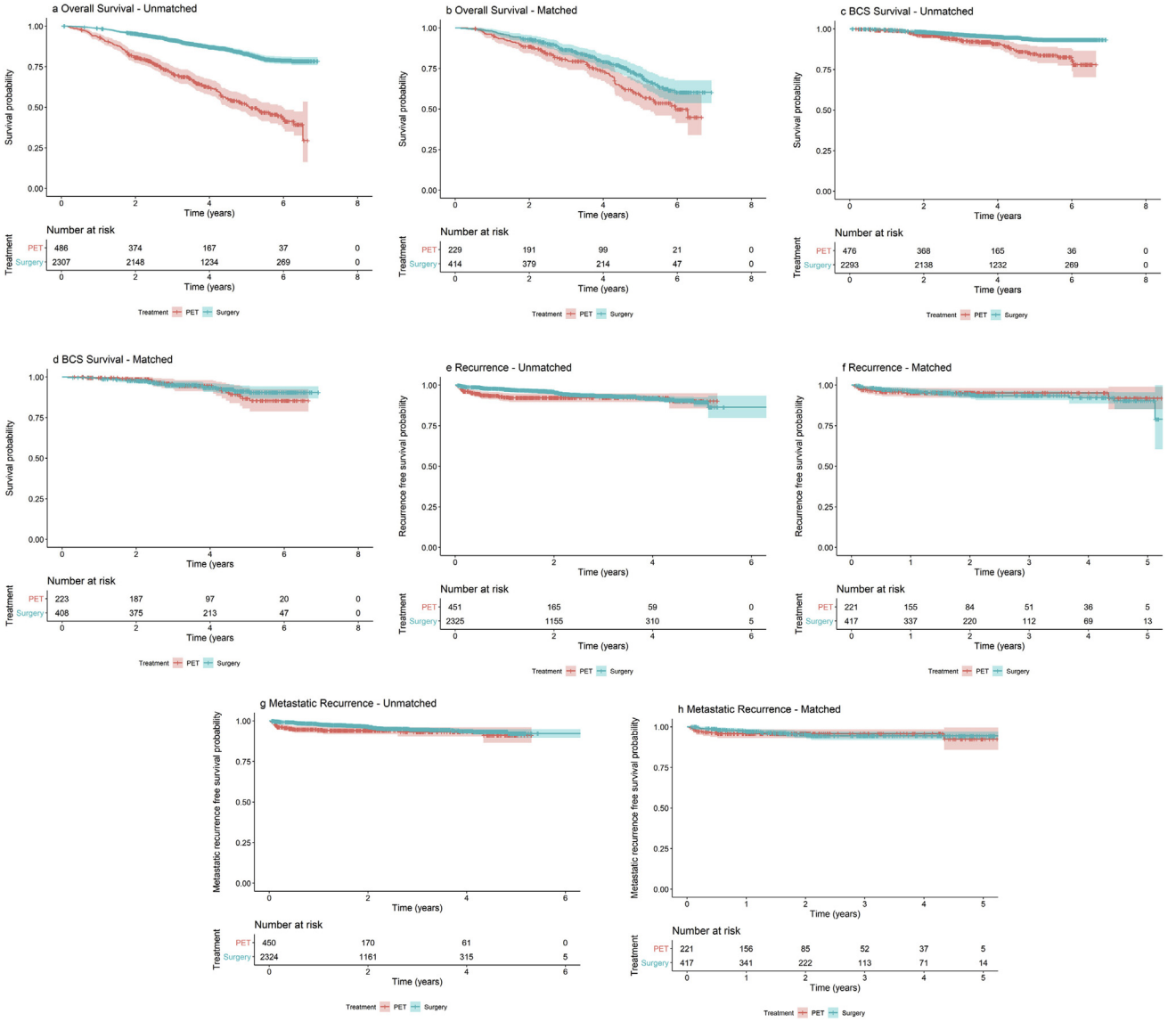

Fig. 3. a-h. Kaplan Meier overall survival (a and b), breast cancer–specific survival (c and d), recurrence-free survival curves (e and f) and metastatic recurrence-free survival (g and h) for unmatched (a, c, e and g) and matched (b, d, f and h) populations in women treated with surgery plus adjuvant endocrine therapy versus PET. Median follow-up of 52 months shown. Confidence intervals shown in pale blue or orange shading. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

Statistical analyses

Analyses were performed in IBM SPSS version 24 (IBM Corp, NY), R version 3.6.3 and Stata (StataCorp LLC).

Propensity matching

Logistic regression was used to calculate propensity scores for treatment allocation, which were used to match PET to S+ET patients. The covariates were measures of functionality (ADL, IADL, MMSE, ECOG), nutritional status (abridged PG-SGA), comorbidities (CCI, number of medications) and age.

The ratio and calliper widths of the propensity scores were chosen based on examination of propensity score overlaps for several combinations of ratios and callipers (to describe how closely matched the patients are). A 1:2 ratio for PET to surgery and a calliper of 0.25 times the propensity scores' standard deviation were used to optimally match quality and numbers. Participants were also matched by Nottingham Prognostic Index (NPI) [25] category (good $\leq$3.4, moderate 3.5–5.4, poor >5.4) to avoid fit participants with aggressive cancer being matched with frail participants with smaller cancers.

For both OS and BCSS, four Cox proportional hazard models were fitted to compare treatments on the entire unmatched cohort. The first two included all patients, both unadjusted and then adjusted for baseline

age, functionality and tumour characteristics, allowing us to give estimates for the whole group whilst still adjusting for these covariates. The last two included only matched patients and incorporated a shared frailty term (random effect) for matching, again both adjusted and unadjusted. The primary analysis was on the matched population as this related to a subgroup for whom either treatment may have been considered.

Results

Cohort description

The study recruited 3416 women between January 2013 and June 2018, with a median age of 77 years (range 69–102; 5 patients were recruited just before their 70th birthday and have been retained in the study). Of these 3315 were fully eligible for analysis (Fig. 1). The majority (2854; 86%) had ER+ breast cancer on biopsy and were therefore potential candidates for PET or S+ET. The following results relate to patients with ER+ breast cancer only.

Of the 2854 patients with ER+ cancer, 2354/2854 (82%) had S+ET (60% breast conservation and 40% mastectomy) and 500/2854 (18%) commenced PET (87% with letrozole) within 6 months of baseline assessment. Tumour, patient and treatment characteristics are summarised in Table 1. Patients undergoing S+ET were generally younger, fitter and had superior function compared to those undergoing PET.

Women in the S+ET group had various additional therapies including radiotherapy, chemotherapy and trastuzumab according to clinical indications at various time points (Fig. 2).

Propensity matching

Table 2 shows the balance of characteristics achieved in the final matched dataset, which found a suitable match for 238 (48%) of the PET patients, 184 of whom were matched with two surgery patients and 54 matched to only one. The remaining 262 PET patients comprised 201 patients with incomplete functional data (117 (58% were partial/consultee participants and therefore too frail to complete questionnaires) and a further 61 that were too dissimilar from patients undergoing surgery (again largely too frail and comorbid for any surgery). A summary of the matching process and matching quality are summarised in Supplemental Fig. SF1.

Overall survival and breast cancer–specific survival

Overall survival among all patients (unmatched analyses)
Overall mortality status at median 52 months follow-up was available for 2793/2854 (98%) patients in the S+ET and PET populations.

Of the 486 patients who received PET and with survival data, 203 (41.8%) died during follow-up, compared to 336/2307 (14.6%) of S+ET patients. Patients treated with PET had inferior overall survival compared with those treated with S+ET (unadjusted HR 0.27, 95% CI 0.23–0.33, $P < 0.001$) (Fig. 3 and Table 3) but, adjusting for case mix via multivariable Cox regression reduced this difference (adjusted HR = 0.83, 95% CI: 0.63–1.09, $P = 0.18$) (Fig. 3 and Table 3).

Breast cancer–specific survival (unmatched analyses)
A total of 45/476 (9.5%) patients died due to breast cancer in the PET group versus 113/2293 (4.9%) in the surgery group (unadjusted HR: 0.41, CI: 0.29–0.58, $P < 0.001$). Patients treated with PET had inferior BCSS compared with those treated with S+ET, but adjusting for case mix reduced this difference to a hazard ratio of 0.89 (95% CI: 0.52 to 1.53; $P = 0.68$; Fig. 3 and Table 3).

Overall survival (matched analyses)
Mortality status was available for 643/660 (97%) patients in the matched surgery versus PET population. Of the 229 patients who received PET and for whom we had adequate data for analysis, 79/229 (34.5%) died during follow-up from all causes, compared to 106/414 (25.6%) of surgery patients (matched, unadjusted HR:0.66, 95% CI: 0.49–0.90, $P = 0.008$). Matched patients treated with PET had inferior overall survival compared with those treated with S+ET. Further adjusting for case mix reduced but did not remove this difference (matched, adjusted HR: 0.72, 95% CI: 0.53–0.98, $P = 0.037$) (Fig. 3 and Table 3).

Breast cancer–specific survival (matched analyses)
In the matched breast cancer–specific mortality cohort there were 17/223 (7.6%) breast cancer–specific deaths for PET versus 27/408 (6.6%) for surgery (matched, unadjusted HR: 0.80, 95% CI: 0.43–1.47, $P = 0.46$). Adjusting for residual imbalance via multivariable regression provided similar findings (matched, adjusted HR: 0.74, 95% CI: 0.40–1.37, $P = 0.34$) (Fig. 3 and Table 3).

The various match/unmatched, adjusted and unadjusted hazard ratios are summarised in Supplemental Table ST2.

Recurrence and progression

Unmatched analyses
Rates of overall (locoregional and metastatic) recurrence in the unmatched cohort were higher in patients in the PET group 33/451 (7.3%) compared to the S+ET group 113/2325 (4.9%). Of these locoregional recurrences (or progression in the case of PET) there were 6/451 (1.33%) for PET and 25/2325 (1.07%) for S+ET (Table 3 and Fig. 3). After adjusting for age, baseline health status, physical function and NPI, surgery had no

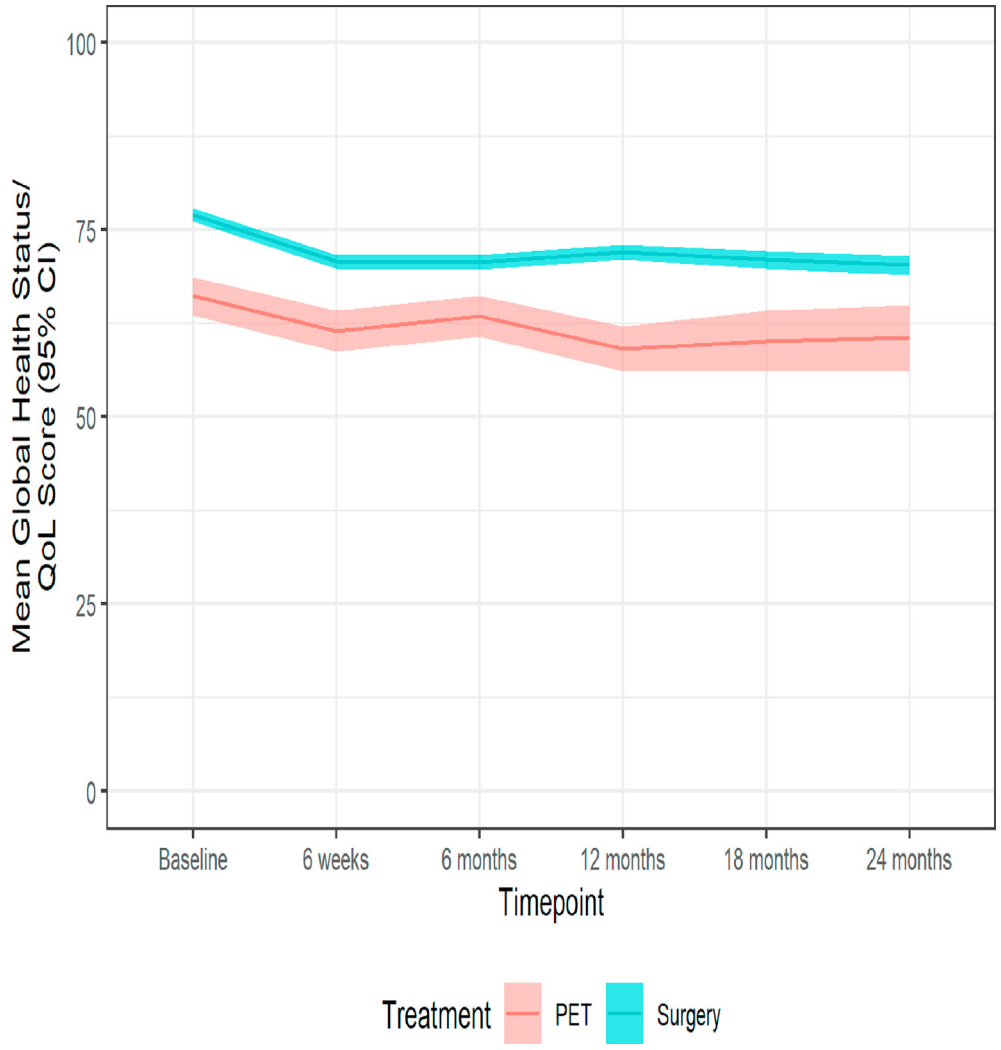

Fig. 4. Mean global health status and 95% confidence intervals for women treated with PET versus Surgery plus endocrine therapy in the unmatched cohort.

effect on the rate of recurrence (HR: 1.01; 95% CI: 0.52–1.95; $P = 0.981$).

Matched analysis

On matching, difference in rates of recurrence were not significant and the hazard ratio for recurrence between the two treatments was 1.11 (95% CI: 0.55–2.26), $P = 0.775$ (Table 3 and Fig. 3).

Quality of life

The global health score of the EORTC QLQ C30 (questions 29 and 30)

Significant variation in age and health status meant that baseline quality-of-life scores differed. Patients treated surgically reported higher QoL both at baseline and throughout the 24-month follow-up. Between baseline and 6 weeks (when the majority of patients started treatment with either surgery or PET, Fig. 2) a clinically significant reduction in mean global health status was observed from a mean of 66.2 (standard deviation 21.1) to 61.5 (21.4) for PET patients and from 77.1 (17.8) to 70.8 (18.5) for surgery patients. In neither group did levels recover to baseline following treatment even at 2 years (Table 4 and Fig. 4).

Functional independence after treatment

The EQ-5D-5L score was compared in the unmatched cohort. Baseline scores demonstrated significant variation. Both the overall score and the individual questions showed decreased health status across the 2-year period. For the PET group the fall was gradual whereas for the surgery group there was an early sharp fall between baseline and 6 weeks which then failed to recover, indicating a lack of resilience in this older population. This was particularly evident in the 'ability to perform usual activities score' shown in Fig. 5.

Table 3
Overall and breast cancer–specific survival and recurrence data for the unmatched and matched cohorts.

Total cohort		Unmatched		Matched	
		PET	Surgery	PET	Surgery
		N = 500	N = 2354	N = 238	N = 422
Overall survival	Data available	486	2307	229	414
	Alive	351 (72.2%)	2095 (90.8%)	150 (65.5%)	308 (74.4%)
	Died of any cause	135 (27.8%)	212 (9.2%)	79 (34.5%)	106 (25.6%)
Cause-specific survival	Data available	477	2297	223	408
	Alive or died of other causes	452 (94.8%)	2217 (96.5%)	206 (92.4%)	381 (93.4%)
	Died of breast cancer	25 (5.2%)	80 (3.5%)	17 (7.6%)	27 (6.6%)
Recurrence (all types)	Data available	451	2325	221	417
	No	418 (92.7%)	2212 (95.1%)	210 (95.0%)	392 (94.0%)
	Yes	33 (7.3%)	113 (4.9%)	11 (5.0%)	25 (6.0%)
Metastatic recurrence	Data available	450	2324	221	417
	No	423 (94.0%)	2236 (96.2%)	211 (95.5%)	399 (95.7%)
	Yes	27 (6.0%)	88 (3.8%)	10 (4.5%)	18 (4.3%)

Surgical morbidity and mortality

This is reported in detail elsewhere [20]. There was one death within 30 days of surgery which was due to metastatic breast cancer, though surgery cannot be ruled out as contributing. There was one further death at day 72, which was due to COPD and pneumonia. There were therefore no deaths directly attributable to surgery.

Surgical morbidity was moderate. There were 551/2854 (19%) complications although most were local wound complications such as seroma, haematoma and infection with only 59/2854 (2.1%) women suffering systemic complications such as cardiorespiratory, cerebrovascular or thrombotic events.

Discussion

The aim of the Age Gap study was to determine the outcomes of PET versus S+ET. As expected, patient characteristics differed significantly between groups with PET generally allocated to older, frailer women. There is significant variation in practice with respect to surgery rates across the UK [26]. This variance meant we were able to identify cohorts of women with similar characteristics who were treated with PET or S+ET permitting

Table 4
Global health score of the EORTC QLQ C30 score for matched and unmatched patients at each time point.

Unmatched			PET	Surgery	Total
Global health status/QoL	Baseline	n	258	1644	1902
		Mean (SD)	66.2 (21.1)	77.1 (17.8)	75.6 (18.7)
	6 weeks	n	230	1511	1741
		Mean (SD)	61.5 (21.4)	70.8 (18.5)	69.6 (19.2)
	6 months	n	199	1418	1617
		Mean (SD)	63.5 (19.5)	70.7 (19.3)	69.8 (19.5)
	12 months	n	149	1233	1382
		Mean (SD)	59.2 (18.6)	72.1 (18.4)	70.7 (18.9)
	18 months	n	109	1033	1142
		Mean (SD)	60.2 (21.7)	71.1 (19.3)	70.0 (19.8)
	24 months	n	82	902	984
		Mean (SD)	60.6 (20.6)	70.3 (19.6)	69.5 (19.8)
Matched			PET	Surgery	Total
Global health status/QoL	Baseline	n	151	293	444
		Mean (SD)	69.1 (19.5)	73.4 (18.5)	71.9 (18.9)
	6 weeks	n	139	259	398
		Mean (SD)	64.1 (20.9)	66.8 (19.1)	65.9 (19.8)
	6 months	n	122	236	358
		Mean (SD)	64.2 (18.7)	64.5 (19.4)	64.4 (19.1)
	12 months	n	96	206	302
		Mean (SD)	58.9 (18.9)	67.2 (18.4)	64.6 (19.0)
	18 months	n	68	175	243
		Mean (SD)	62.5 (22.0)	66.5 (20.4)	65.4 (20.9)
	24 months	n	50	147	197
		Mean (SD)	63.3 (19.6)	63.0 (21.3)	63.1 (20.8)

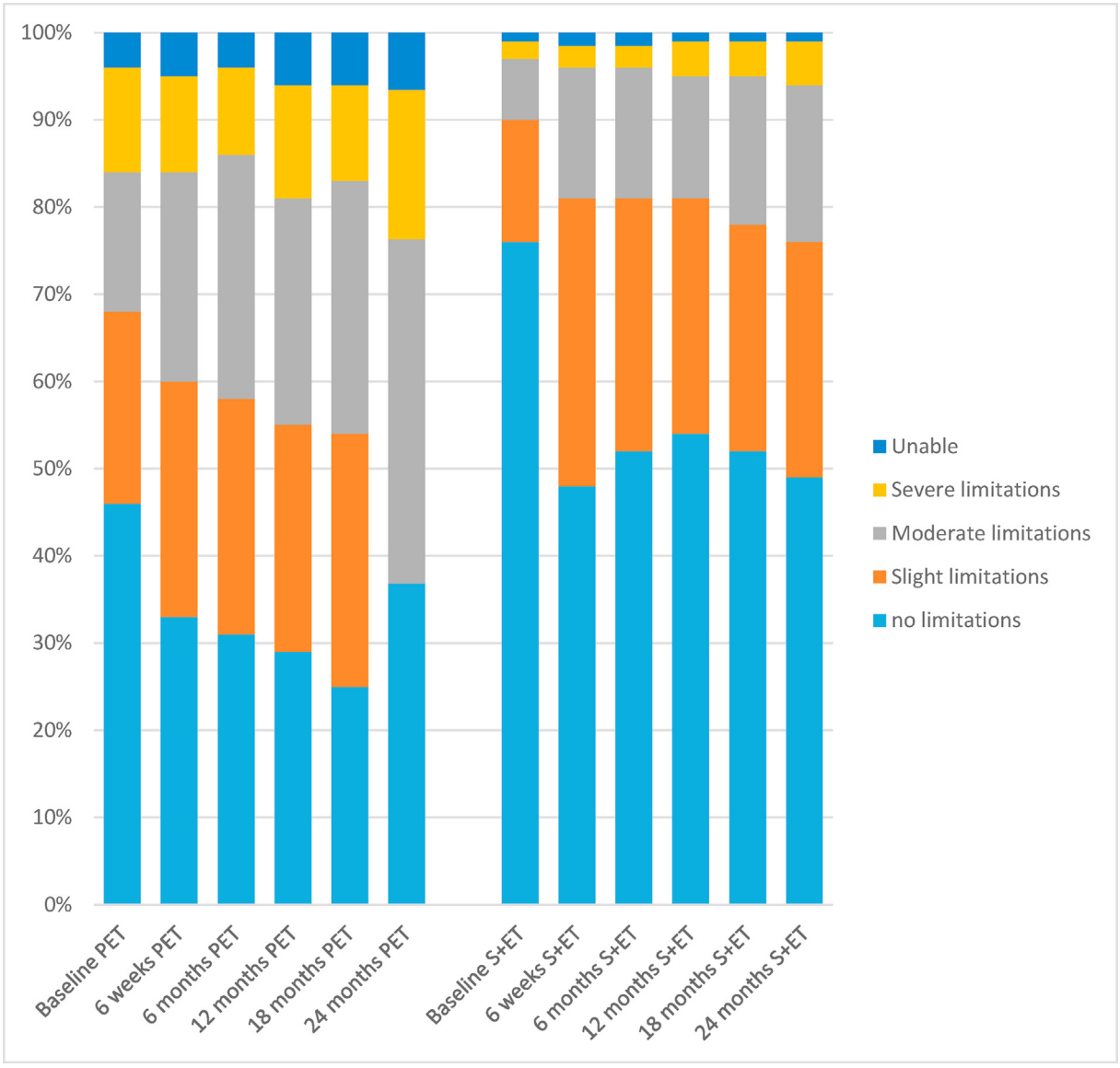

Fig. 5. Bar charts of the EQ-5D-5L 'ability to perform usual activities' score at each time point for the surgery versus PET population for the unmatched population.

us to create a matched cohort of quantitatively similar women who may have been candidates for either treatment. Some degree of bias inevitably persists due to imperfections in the matching process, but this study achieved a good quality match. As a result, whilst surgery is still associated with an overall survival benefit, this difference disappears in both BCSS and recurrence/progression-free survival out to 52 months follow-up. The rate of local recurrence/progression was low in both arms, and for the PET group the rate of progression is lower than other reported series. This may reflect lack of direct follow-up out to 52 months as directly collected data were only available to 24 months. It may also reflect the fact that the cohort selected for PET were all women with ER+ cancers and letrozole was mainly used, which is more effective than tamoxifen.

In this large cohort, surgery was the primary treatment modality in 83% and PET in 17%, which is slightly lower than a recent national audit (National Audit of Breast Cancer in Older People, NABCOP [27]) which reported a PET rate of 24%. This variance may reflect a difference in mix of unit practice, selective recruitment of the less frail into the trial or the technical challenge of recruiting women having PET where the recruitment window before treatment starts is narrow. As seen in Fig. 1, 156 eligible patients were missed due to commencing treatment before they could be recruited, the majority of whom would have been started on PET. The age range of patients within the study is not entirely representative of the UK population of women with breast cancer, with a slight excess of younger women and fewer of the very oldest. In addition, the rate of

study discontinuation was slightly higher in the older age groups. This reduces the generalisability of the study.

This study suggests that older women with a short life-expectancy of 5 years or less derive little or no survival benefit from undergoing surgery, but women likely to survive longer than 5 years may start to see endocrine resistance develop on PET. This accords with the RCTs where survival outcomes at 5 years were equivalent but over the longer term outcomes diverge. The data from our study have an advantage over the historic RCT data in that all women in this analysis had ER+ cancer, the majority were treated with aromatase inhibitors, which are more effective in the neoadjuvant and PET settings, and have the real-world age, comorbidity and frailty characteristics typical of PET patients in clinical practice.

Of great interest in this study was the impact of treatment on quality of life and physical function. In the unmatched analysis it is clear that women in the PET treatment group have worse baseline quality-of-life scores in most domains, reflecting their greater comorbidity and reduced physical function. Analysis of the global health status score shows that there was a slightly steeper decline between baseline and 6 weeks in the surgery group, reflecting the acute impact of surgery. Furthermore, these scores do not return to baseline thus demonstrating the longer term impact of a cancer diagnosis and its treatment. We also see a reduction in the number of women with no limitations to their 'usual activities' with the EQ-5D after surgery, which never recovers even after 24 months denoting a lack of resilience in this older age group. This should be taken into account when making clinical decisions concerning frailer, older women.

Previous work by this group has shown that older women value quality of life and independence highly in their treatment decision making [3,24]. These women perceive PET to offer a safe option to maintain the *status quo* and minimise risk.

There were no deaths directly attributable to surgery, in keeping with other national audits and severe surgical morbidity was low (2.1% had systemic morbidity such as cerebrovascular or thrombotic complications), reflecting the fact that breast surgeons may exclude the frailest older women from surgery if they have ER+ cancer by offering PET. Higher rates of adverse events are reported when frail older women have breast surgery. Two women died within 90 days of surgery of non-surgical causes and whilst these deaths cannot be attributed to surgery they probably represent surgical overtreatment.

For women in the matched cohort (generally older, less fit women), the benefits of surgery at 5 years were very small and diminish for women with greater age/frailty and comorbidity burdens (i.e. women too unfit to match to anyone in the surgery cohort).

One of the key aims of this work was to identify if a subgroup of women may safely be offered PET. The data suggest this is the case, although longer-term follow-up will be required. It is still likely, provided patients are fit enough, that surgery will give optimal survival for the majority of older women. Women who are of borderline fitness for surgery should be offered an informed choice, highlighting the potential difference in survival and adverse events. To this end, our group has developed the fully validated [28] Age Gap Decision Tool (https://agegap.shef.ac.uk/) using UK cancer registry data. The tool has been developed with older women for usability and acceptability [29,30]. This tool may support shared decision making in older women.

Author contributions

Substantial contributions to the conception of the study: Wyld L, Collins K, Reed MW, Brennan A, Burton M, Martin C, Morgan J, Walters S, Ring A, Patnick J. Design of the work: Wyld L, Collins C, Reed MW, Burton M, Morgan J, Walters S, Ring A, Martin CL, Bradburn M, Chater T, Pemberton K, Revill D, Green T, Gath J, Robinson T, Audisio R, Holmes GR. Acquisition of data: Wyld L, Martin CL, Morgan J, Burton M, Reed MW, Collins K, Shrestha A, Chater T, Pemberton K, Lifford K, Edwards A, Todd A, Cheung KL, Holcombe C, Winter M, Gosney M, Hatton M, Thompson A, Naik J, Parmeshwar R, Revill D, Ring A, Green T, Gath J, Robinson T, Audisio R, Horgan K. Analysis of data: Wyld L, Ward S, Bradburn M, Herbert E, Walters S, Martin CL, Morgan J, Burton M. Interpretation of data for the work: Wyld L, Walters S, Bradburn M, Herbert E, Morgan J. Data monitoring and ethics committee: Thompson A (Chair), Gosney M and Hatton M. AND. Drafting the work or revising it critically for important intellectual content: All. AND. Final approval of the version to be published: All. AND. Agreement to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved: All.

Funding

The study was funded by the UK National Institute for Health Research (NIHR).

Disclaimer

This paper presents independent research funded by the National Institute for Health Research (NIHR) under its Programme Grants for Applied Research Programme (grant reference number RP-PG-1209-10071).

The views expressed are those of the authors and not necessarily those of the National Health Service, the NIHR or the Department of Health. The NIHR has no oversight or control over the outputs from the study.

Data sharing

There are no plans for data sharing.

Disclosure

Professors Thompson Robinson and Stephen Walters are National Institute for Health Research (NIHR) Senior Investigators and Jenna Morgan is a NIHR Clinical Lecturer.

Conflict of interest statement

None declared.

Acknowledgements

Trial Sponsor: Doncaster and Bassetlaw Teaching Hospitals NHS Foundation Trust, Clinical Research Office, First Floor 'C' Block, Doncaster Royal Infirmary, Armthorpe Road, Doncaster, DN2 5LT, UK.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ejca.2020.10.015.

References

[1] Diab SG, Elledge RM, Clark GM. Tumor characteristics and clinical outcome of elderly women with breast cancer. J Natl Cancer Inst 2000;92(7):550–6.
[2] Tang V, Zhao S, Boscardin J, et al. Functional status and survival after breast cancer surgery in nursing home residents. JAMA Surg 2018;153(12):1090–6.
[3] Shrestha A, Martin C, Burton M, Walters S, Collins K, Wyld L. Quality of life versus length of life considerations in cancer patients: a systematic literature review. Psycho Oncol 2019;28(7):1367–80.
[4] Shrestha A, Martin C, Burton M, Walters S, Collins K, Wyld L. Quality of life versus length of life considerations in cancer patients: a systematic literature review. Psycho Oncol 2019;28(7):1367–80.
[5] Morgan JWL, Collins KA, Reed MW. Surgery versus primary endocrine therapy for operable primary breast cancer in elderly women (70 years plus). Cochrane Database Syst Rev 2014;5:CD004272.
[6] Johnston SJ, Kenny FS, Syed BM, et al. A randomised trial of primary tamoxifen versus mastectomy plus adjuvant tamoxifen in fit elderly women with invasive breast carcinoma of high oestrogen receptor content: long-term results at 20 years of follow-up. Ann Oncol 2012;23(9):2296–300.
[7] Gazet JC, Sutcliffe R. A randomised trial comparing tamoxifen vs. surgery in patients over the age of 70 with operable breast cancer–final results after 28 years of follow-up. Eur J Surg Oncol 2011;37(9):754–7.
[8] Syed BM, Al-Khyatt W, Johnston SJ, et al. Long-term clinical outcome of oestrogen receptor-positive operable primary breast cancer in older women: a large series from a single centre. Br J Canc 2011;104(9):1393–400.
[9] Ellis MJ, Coop A, Singh B, et al. Letrozole is more effective neoadjuvant endocrine therapy than tamoxifen for ErbB-1- and/or ErbB-2-positive, estrogen receptor-positive primary breast cancer: evidence from a phase III randomized trial. J Clin Oncol 2001;19(18):3808–16.
[10] Reed MW, Wyld L, Ellis P, Bliss J, Leonard R. Breast cancer in older women: trials and tribulations. Clin Oncol 2009;21(2):99–102.
[11] von Elm E, Altman DG, Egger M, et al. Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. BMJ 2007;335(7624):806–8.
[12] Charlson ME, Pompei P, Ales KL, MacKenzie CR. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. J Chron Dis 1987;40(5):373–83.
[13] Read JA, Crockett N, Volker DH, et al. Nutritional assessment in cancer: comparing the mini-nutritional assessment (MNA) with the scored patient-generated subjective global assessment (PGSGA). Nutr Canc 2005;53(1):51–6.
[14] Mahoney FI, Barthel DW. Functional evaluation: the barthel index. Md State Med J 1965;14:61–5.
[15] Lawton MP, Brody EM. Assessment of older people: self-maintaining and instrumental activities of daily living. Gerontol 1969;9(3):179–86.
[16] Folstein MF, Folstein SE, McHugh PR. "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. J Psychiatr Res 1975;12(3):189–98.
[17] Oken MM, Creech RH, Tormey DC, et al. Toxicity and response criteria of the eastern cooperative oncology group. Am J Clin Oncol 1982;5(6):649–55.
[18] Aaronson NK, Ahmedzai S, Bergman B, et al. The European Organization for Research and Treatment of Cancer QLQ-C30: a quality-of-life instrument for use in international clinical trials in oncology. J Natl Cancer Inst 1993;85(5):365–76.
[19] Herdman M, Gudex C, Lloyd A, et al. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). Qual Life Res 2011;20(10):1727–36.
[20] Morgan JWL, Collins KA, Reed MW, et al. Breast cancer surgery in older women: outcomes of the bridging age Gap in breast cancer study. Br J Surg 2020. https://doi.org/10.1016/j.jgo.2020.10.015. online.
[21] Therasse P, European Organisation for R, Treatment of Cancer Data C. Evaluation of response: new and standard criteria. Ann Oncol 2002;13(Suppl 4):127–9.
[22] Dueck AC, Mendoza TR, Mitchell SA, et al. Validity and reliability of the US national cancer institute's patient-reported outcomes version of the common Terminology criteria for adverse events (PRO-CTCAE). JAMA Oncol 2015;1(8):1051–9.
[23] Fayers PM, Aaronson NK, Bjordal K, Groenvold M, Curran D, Bottomley A. The EORTC QLQ-C30 scoring manual. 3rd ed. European Organisation for Research and Treatment of Cancer; 2001.
[24] Husain LS, Collins K, Reed M, Wyld L. Choices in cancer treatment: a qualitative study of the older women's (>70 years) perspective. Psycho Oncol 2008;17(4):410–6.

[25] Galea MH, Blamey RW, Elston CE, Ellis IO. The Nottingham Prognostic Index in primary breast cancer. Breast Canc Res Treat 1992;22(3):207–19.

[26] Morgan J, Richards P, Ward S, et al. Case-mix analysis and variation in rates of non-surgical treatment of older women with operable breast cancer. Br J Surg 2015;102(9):1056–63.

[27] Horgan KDD, Jauhari Y, Gannon M, Medina J, Cromwell D. Nationl audit of breast cancer in older people (NABCOP) annual report. 2018. https://www.nabcop.org.uk/reports/nabcop-2018-annual-report/. [Accessed December 2018].

[28] Ward SE, Richards PD, Morgan JL, et al. Omission of surgery in older women with early breast cancer has an adverse impact on breast cancer-specific survival. Br J Surg 2018;105(11):1454–63.

[29] Burton M, Collins KA, Lifford KJ, et al. The information and decision support needs of older women (>75 yrs) facing treatment choices for breast cancer: a qualitative study. Psycho Oncol 2015;24(8):878–84.

[30] Lifford KJ, Edwards A, Burton M, et al. Efficient development and usability testing of decision support interventions for older women with breast cancer. Patient Prefer Adherence 2019;13:131–43.