THE LANCET
Oncology

Supplementary appendix

This appendix formed part of the original submission and has been peer reviewed. We post it as supplied by the authors.

Supplement to: Early Breast Cancer Trialists' Collaborative Group (EBCTCG). Aromatase inhibitors versus tamoxifen in premenopausal women with oestrogen receptor-positive early-stage breast cancer treated with ovarian suppression: a patient-level meta-analysis of 7030 women from four randomised trials. *Lancet Oncol* 2022; published online Feb 3. https://doi.org/10.1016/S1470-2045(21)00758-0.

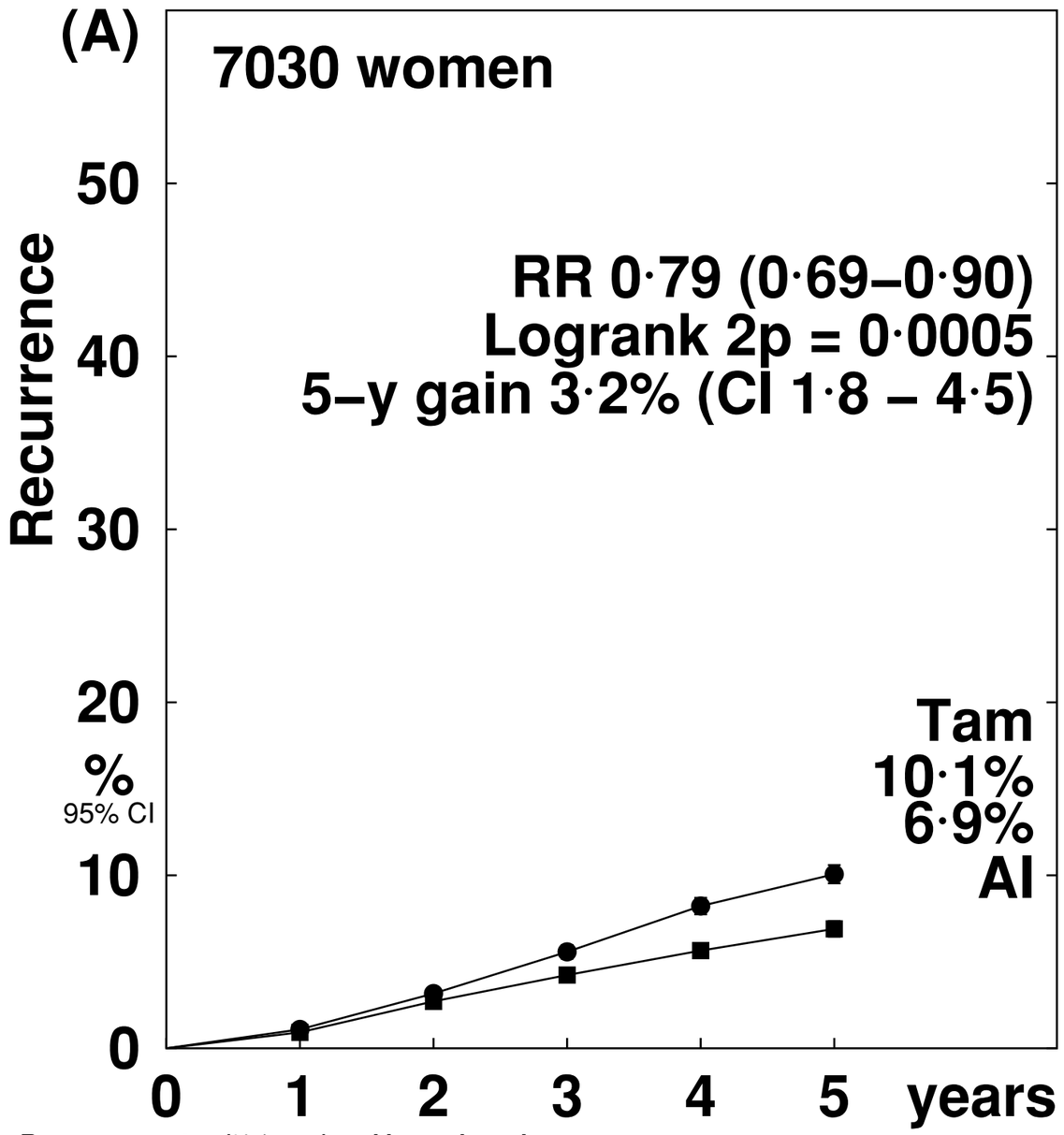
P1: Supplementary appendix for:

Aromatase inhibitors versus tamoxifen in premenopausal women with oestrogen receptor-positive early-stage breast cancer treated with ovarian suppression: a patient-level meta-analysis of 7030 women from four randomised trials

P2: Webappendix: Supplementary figures and tables for "Aromatase inhibitors versus tamoxifen in premenopausal women with oestrogen receptor-positive early-stage breast cancer treated with ovarian suppression: a patient-level meta-analysis of 7030 women from four randomised trials"

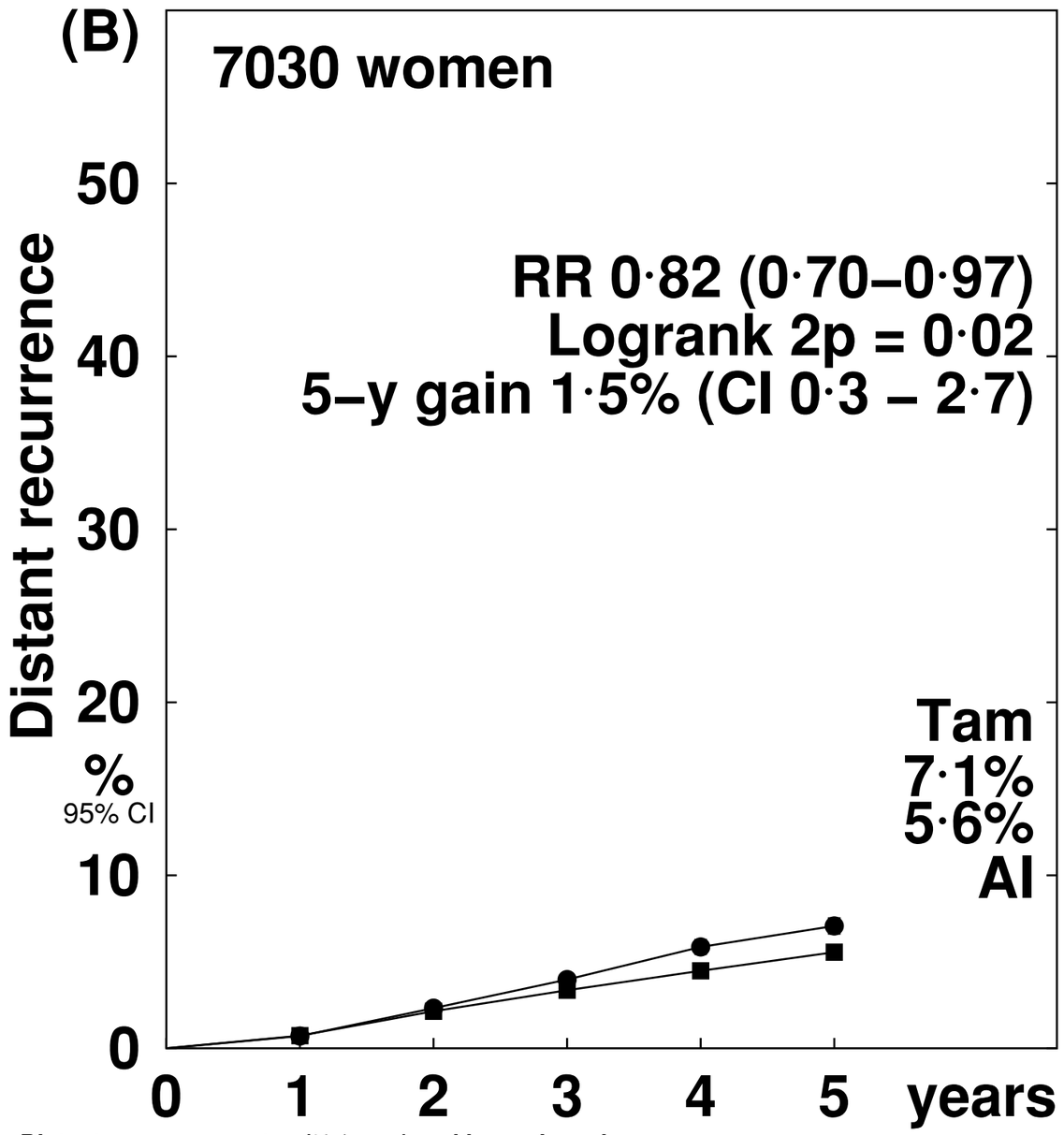
CONTENTS LIST (Click on any item with a page number to jump to it)

P3: 5 year cumulative risk of (A) any recurrence, (B) distant recurrence at any time, (C) isolated local recurrence as first event, (D) contralateral recurrence as first event in trials of AI versus tamoxifen in premenopausal women



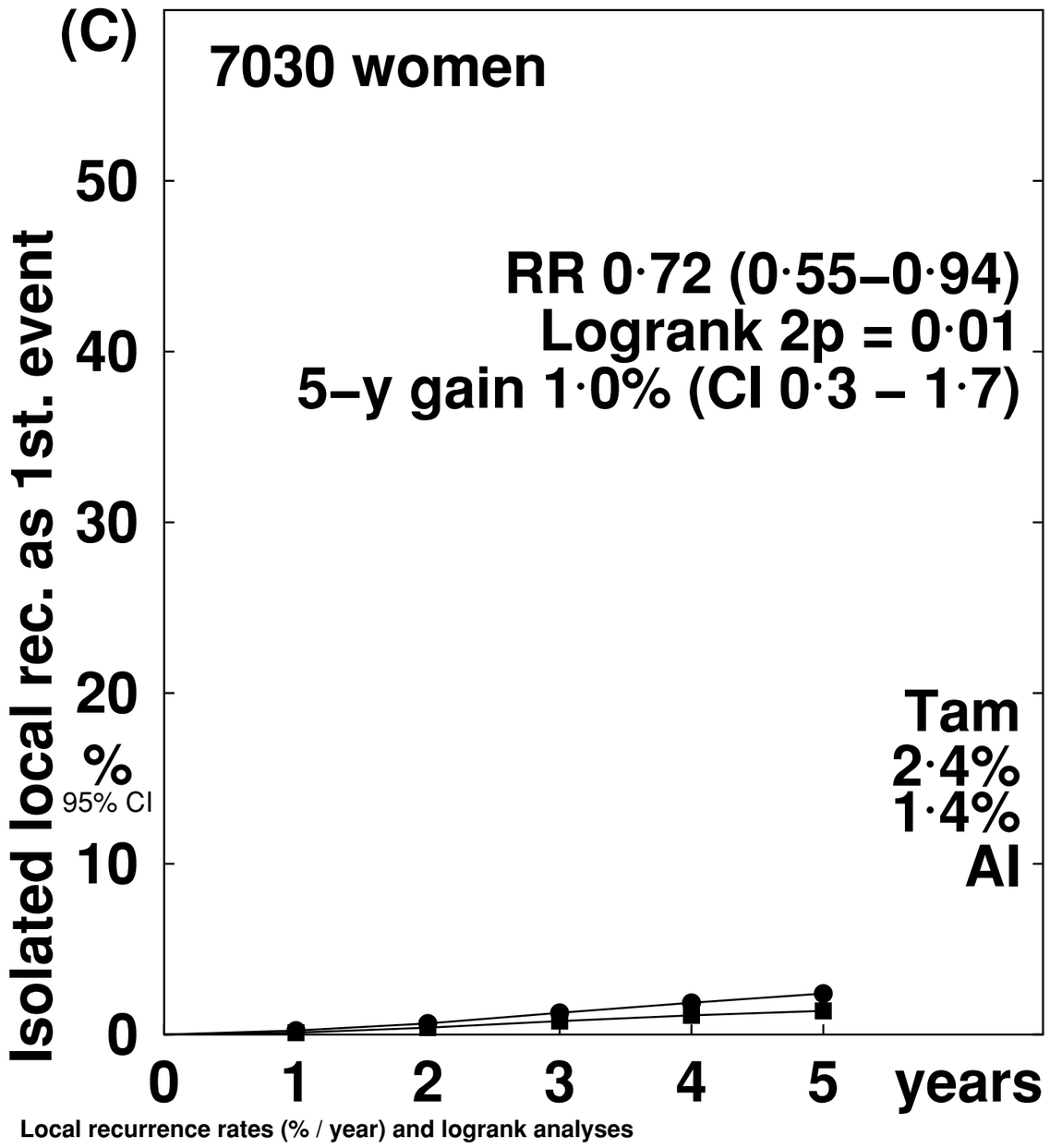
Recurrence rates (% / year) and logrank analyses

Allocation	Years 0 – 1	Years 2 – 4	Year 5+
AI	1·39 (95 / 6854)	1·51 (142 / 9421)	1·76 (167 / 9476)
Tam	1·59 (108 / 6805)	2·45 (224 / 9152)	1·67 (152 / 9098)
Rate ratio, from	0·85 CI 0·65 – 1·13	0·60 CI 0·49 – 0·74	1·03 CI 0·82 – 1·28
(O–E) / V	–7·7 / 49·0	–45·2 / 88·2	1·9 / 77·1



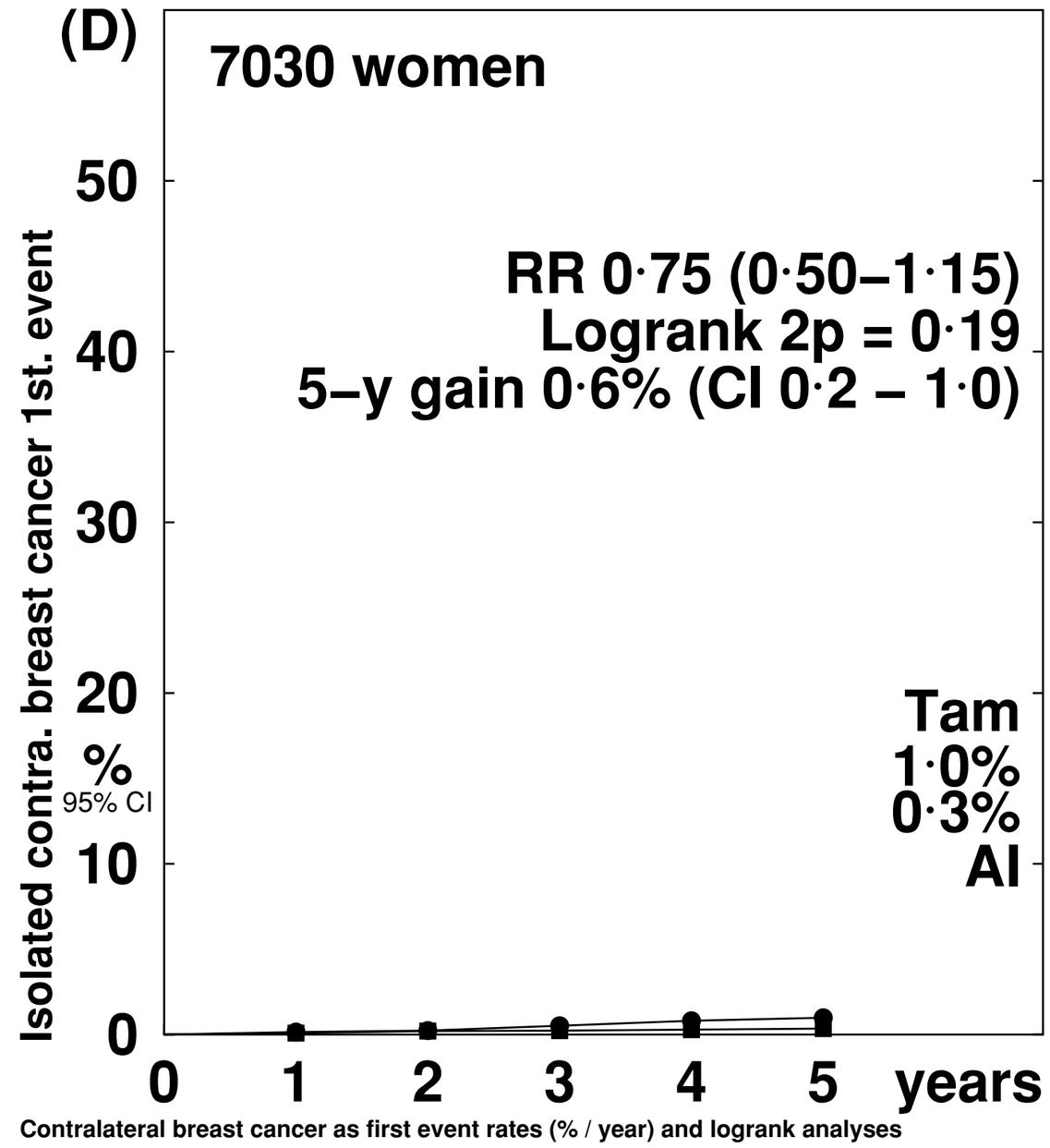
Distant recurrence rates (% / year) and logrank analyses

Allocation	Years 0 – 1	Years 2 – 4	Year 5+
AI	1·09 (75 / 6870)	1·21 (115 / 9516)	0·99 (96 / 9724)
Tam	1·16 (79 / 6831)	1·65 (154 / 9337)	1·02 (97 / 9507)
Rate ratio, from	0·92 CI 0·67 – 1·27	0·71 CI 0·56 – 0·91	0·92 CI 0·69 – 1·23
(O–E) / V	–3·0 / 37·2	–22·1 / 64·8	–3·8 / 46·8



Local recurrence rates (% / year) and logrank analyses

Allocation	Years 0 – 1	Years 2 – 4	Year 5+
AI	0·20 (14 / 6854)	0·34 (32 / 9421)	0·53 (50 / 9476)
Tam	0·32 (22 / 6805)	0·59 (54 / 9152)	0·57 (52 / 9098)
Rate ratio, from	0·62 CI 0·32 – 1·20	0·57 CI 0·37 – 0·87	0·92 CI 0·62 – 1·37
(O–E) / V	–4·2 / 8·9	–12·1 / 21·3	–2·0 / 25·0

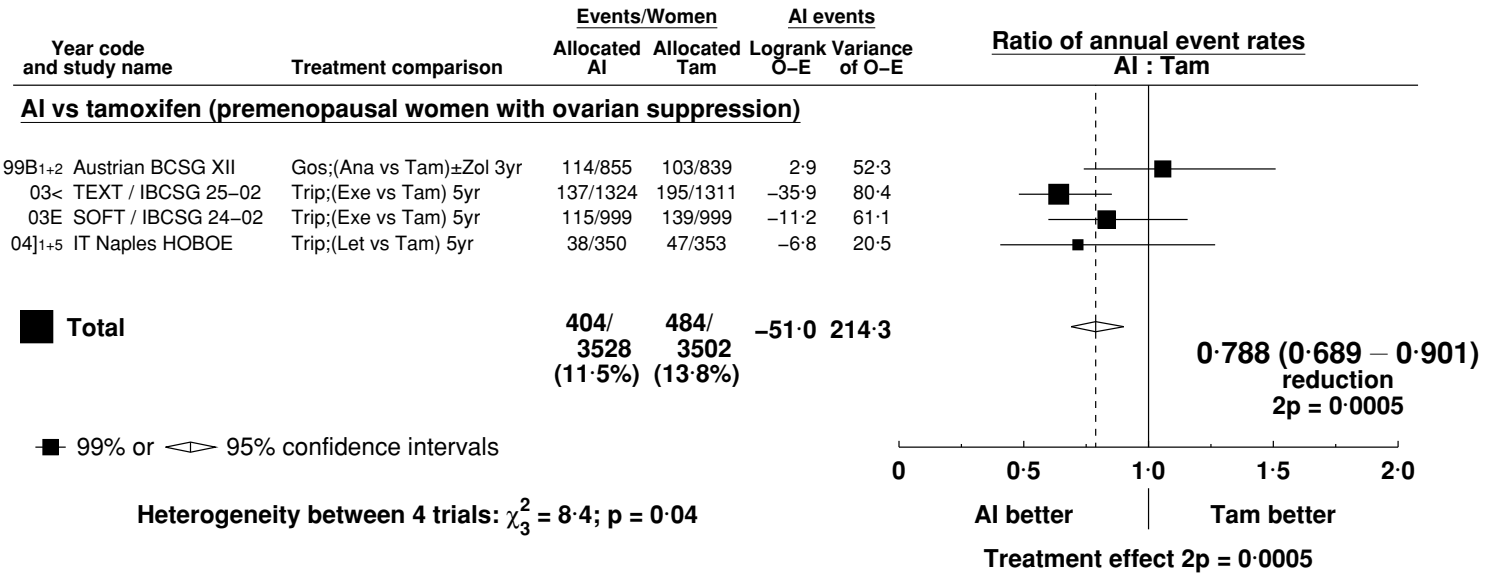


Contralateral breast cancer as first event rates (% / year) and logrank analyses

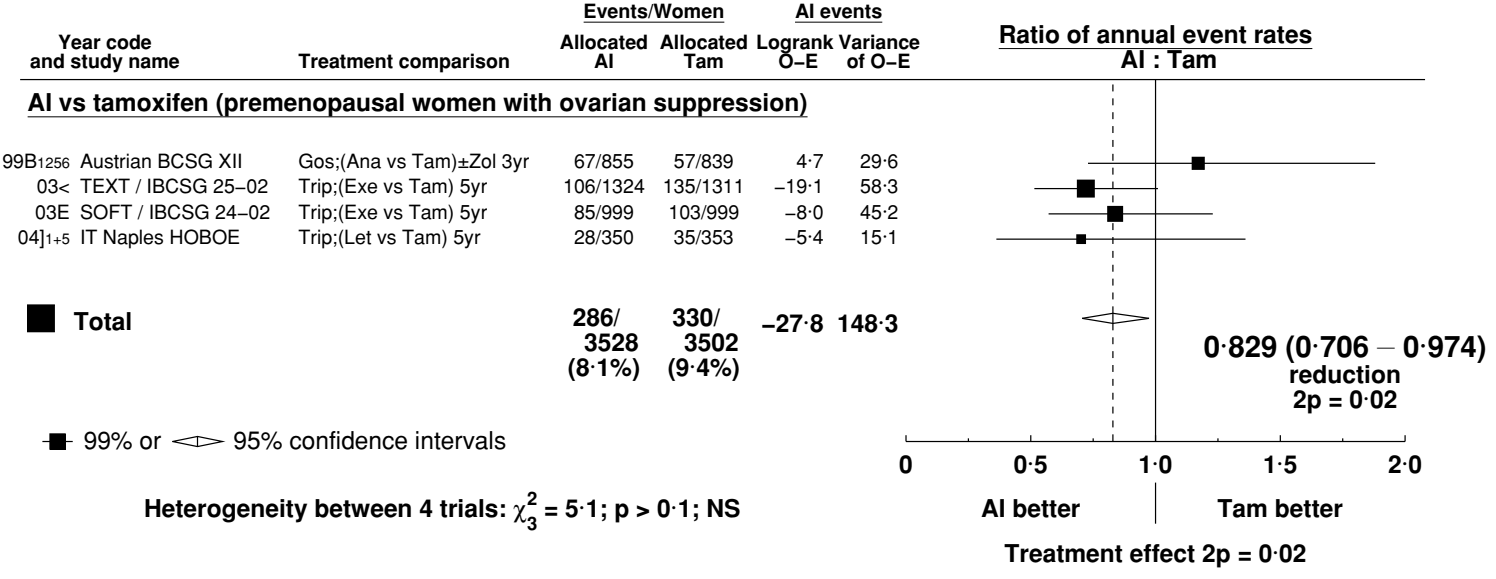
Allocation	Years 0 – 1	Years 2 – 4	Year 5+
AI	0·10 (7 / 6854)	0·05 (5 / 9421)	0·28 (27 / 9476)
Tam	0·12 (8 / 6805)	0·25 (23 / 9152)	0·20 (18 / 9098)
Rate ratio, from	0·87 CI 0·32 – 2·41	0·25 CI 0·12 – 0·54	1·41 CI 0·78 – 2·53
(O–E) / V	–0·5 / 3·7	–9·5 / 6·9	3·8 / 11·2

P4: Recurrence and distant recurrence at any time in trials testings AI versus tamoxifen in premenopausal women

Recurrence

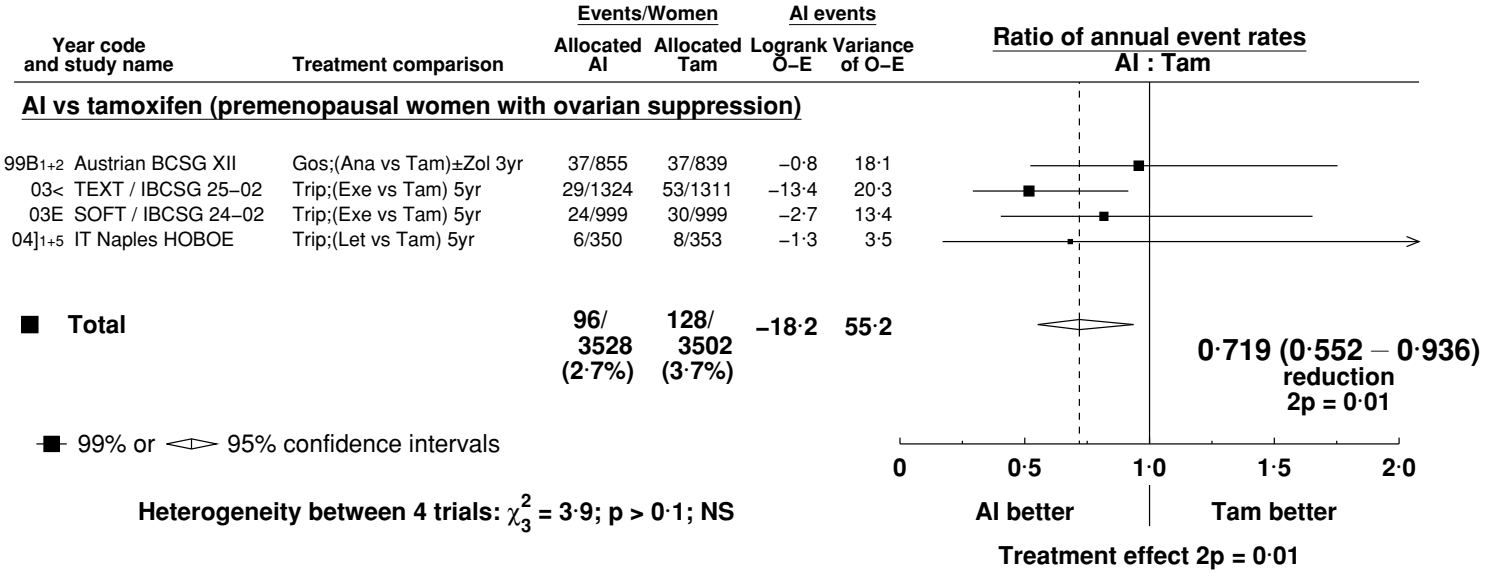


Distant recurrence at any time

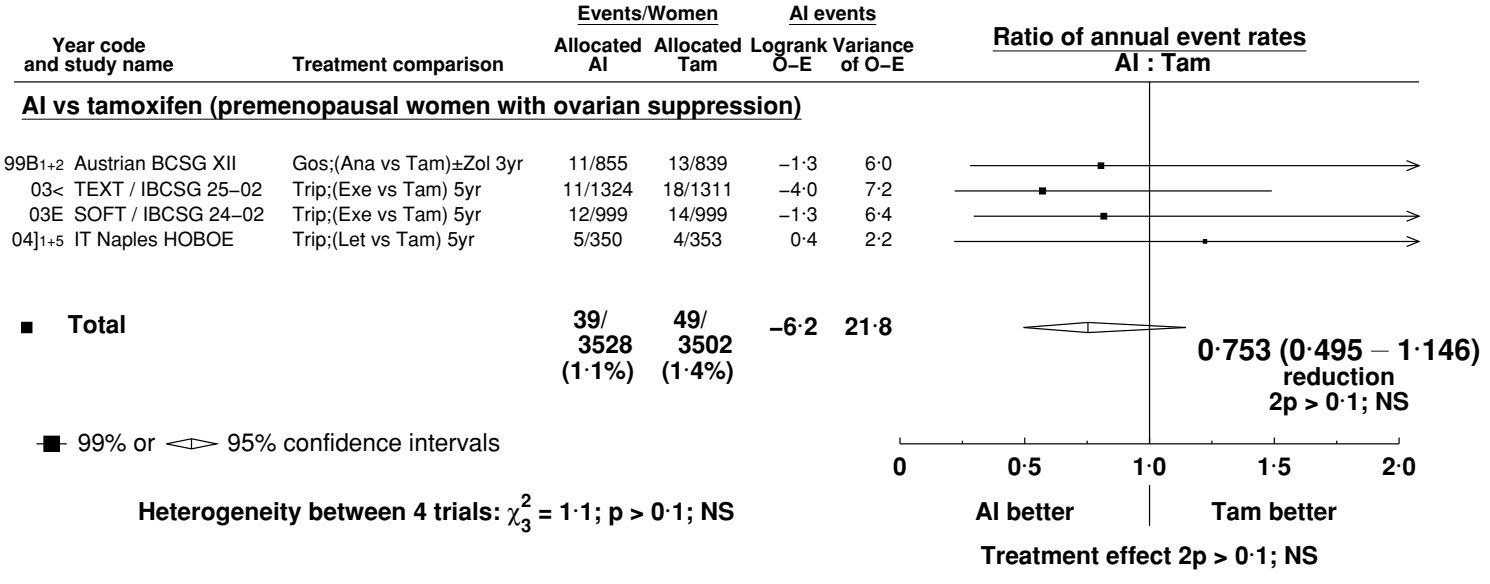


P5: Local recurrence and first event and contralateral recurrence as first event in trials testings AI versus tamoxifen in premenopausal women

Local recurrence as first event

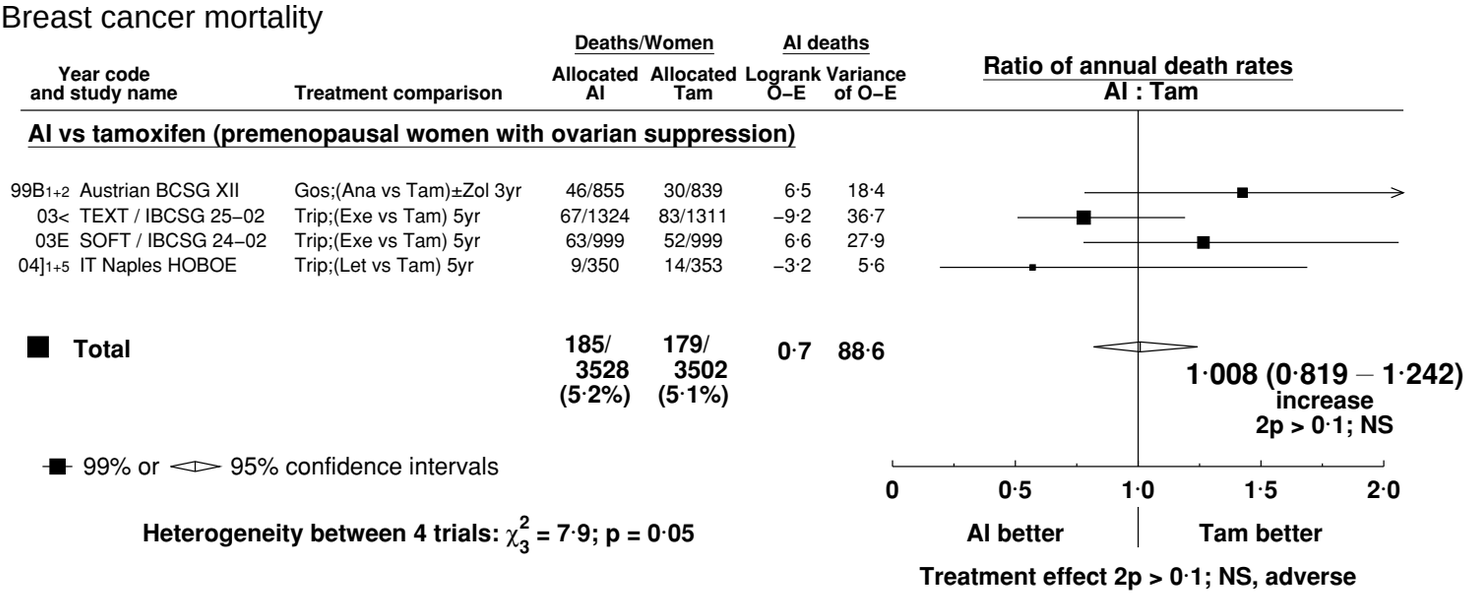


Contralateral recurrence as first event

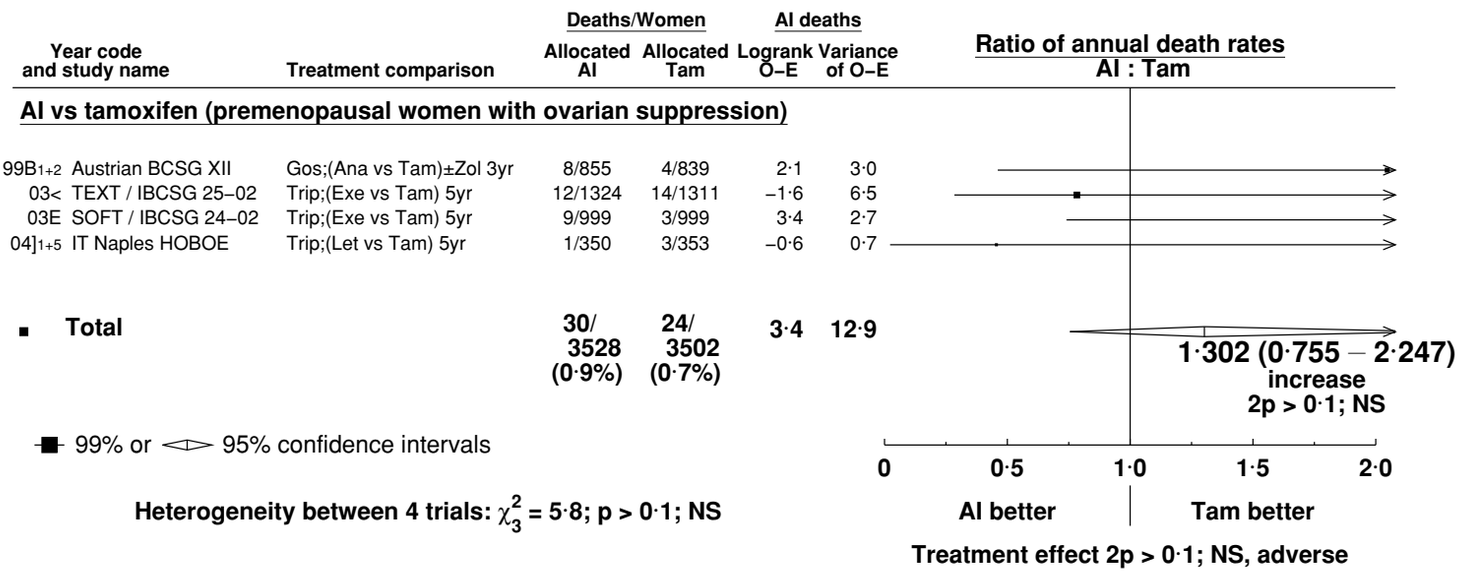


P6: Breast cancer mortality, death without recurrence and all-cause mortality in trials testings AI versus tamoxifen in premenopausal women

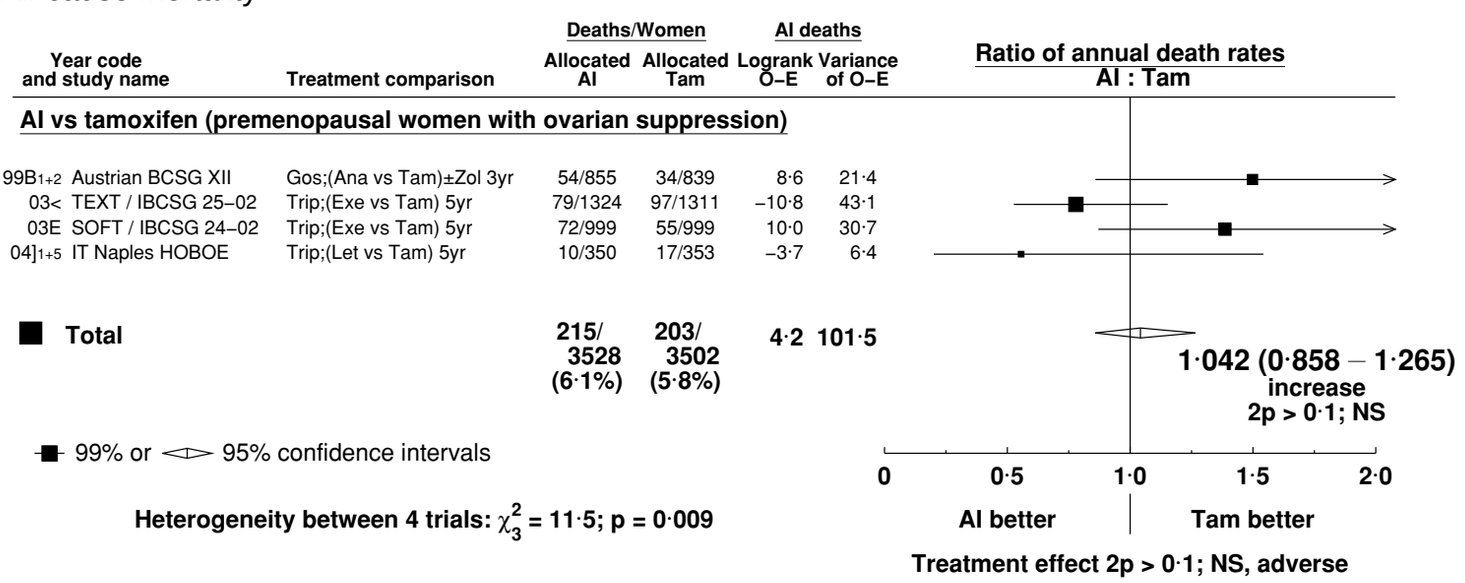
Breast cancer mortality



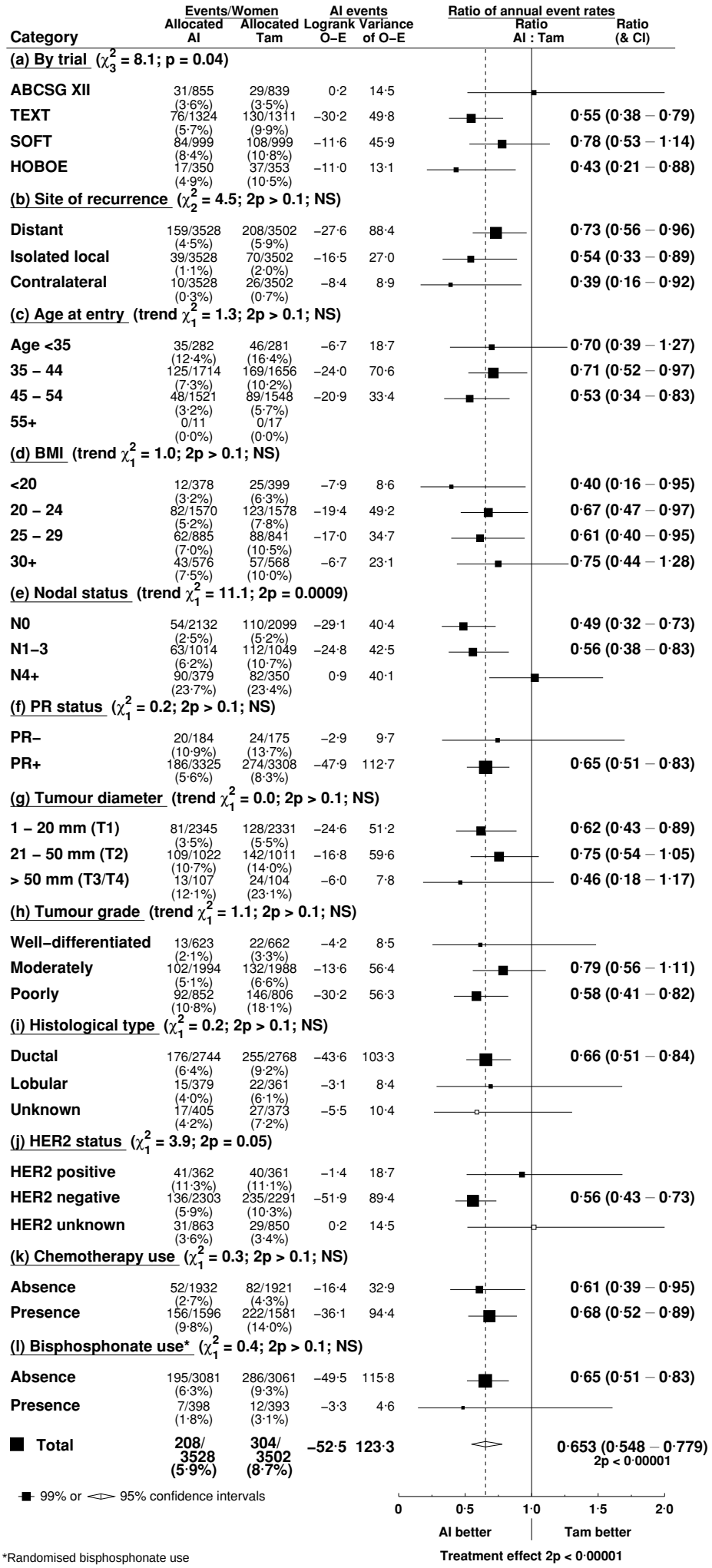
Death without recurrence



All-cause mortality

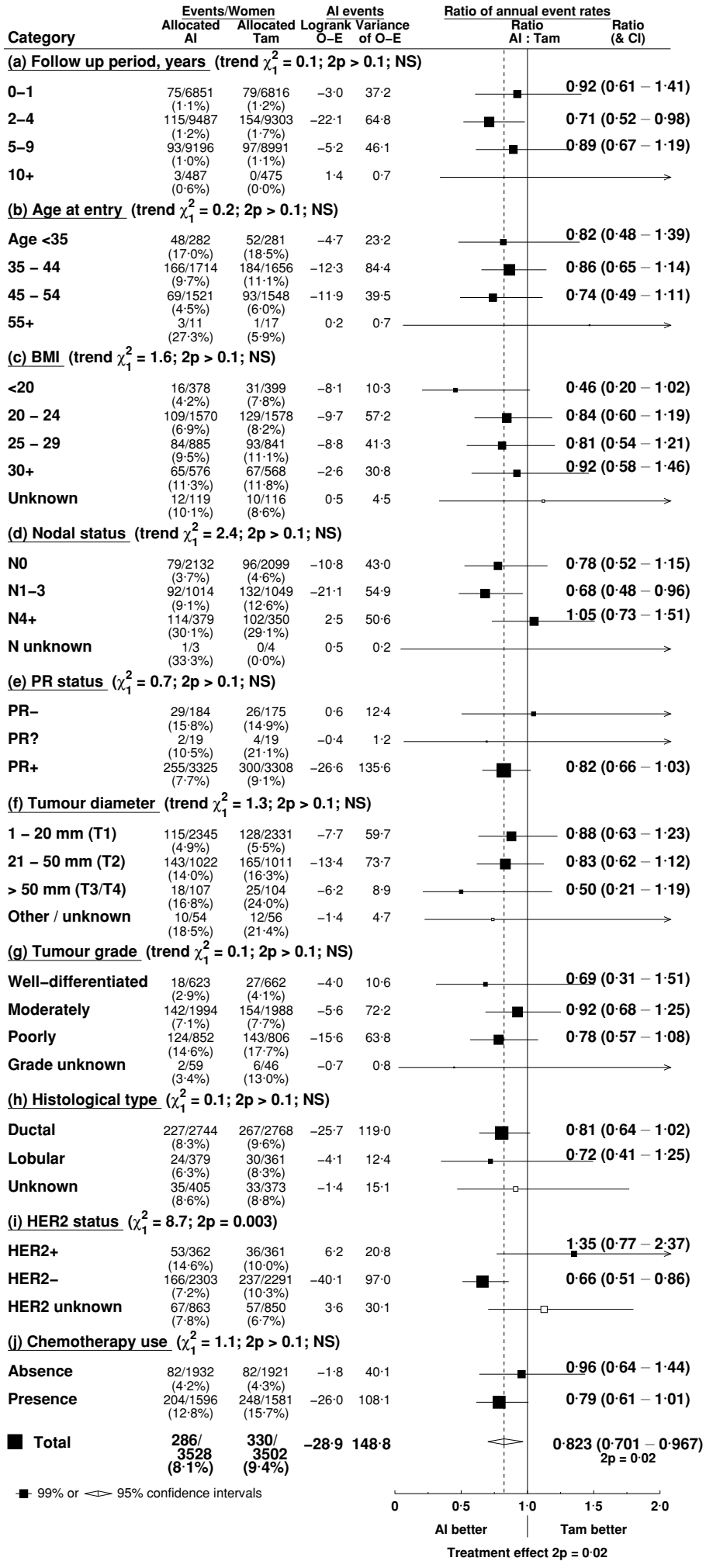


P7: Subgroup analyses of AI vs tamoxifen in premenopausal women; any recurrence only during the time period when treatment differ

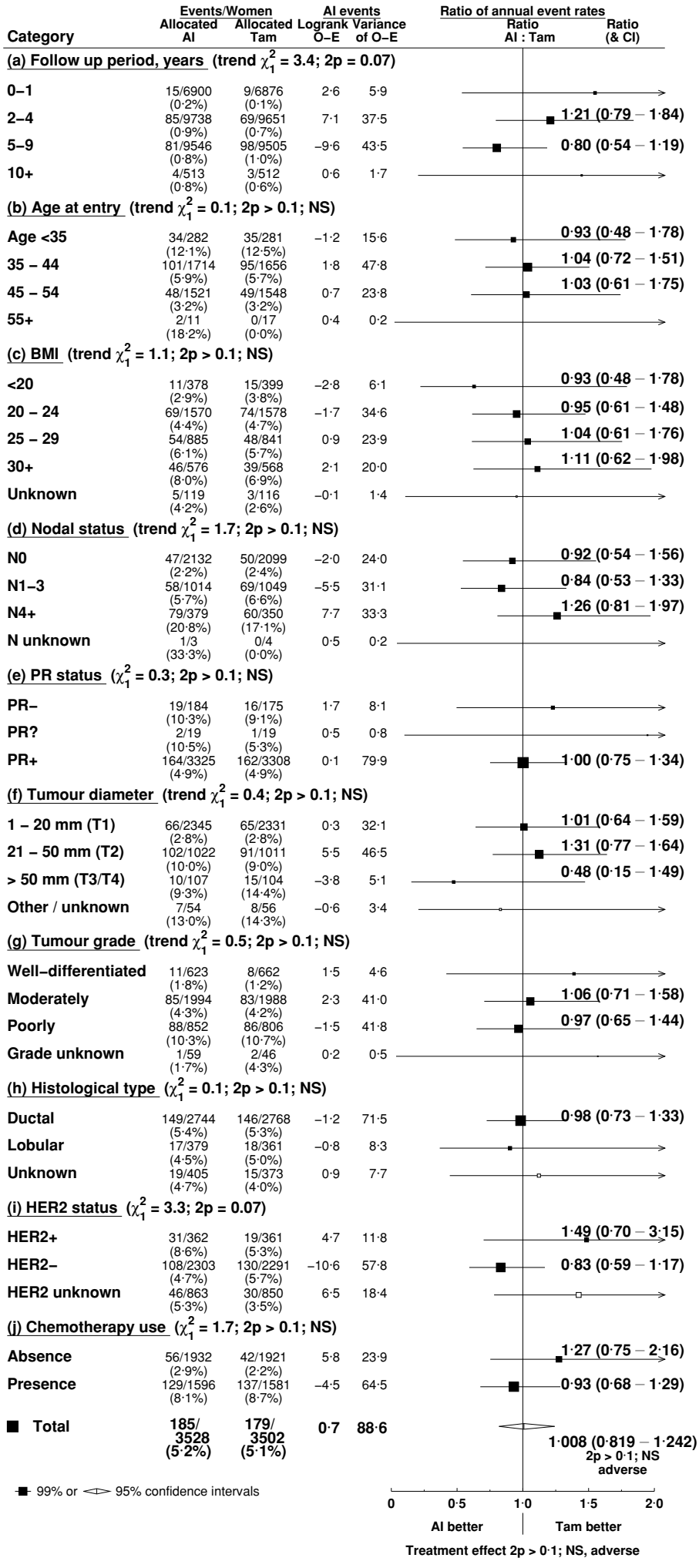


*Randomised bisphosphonate use

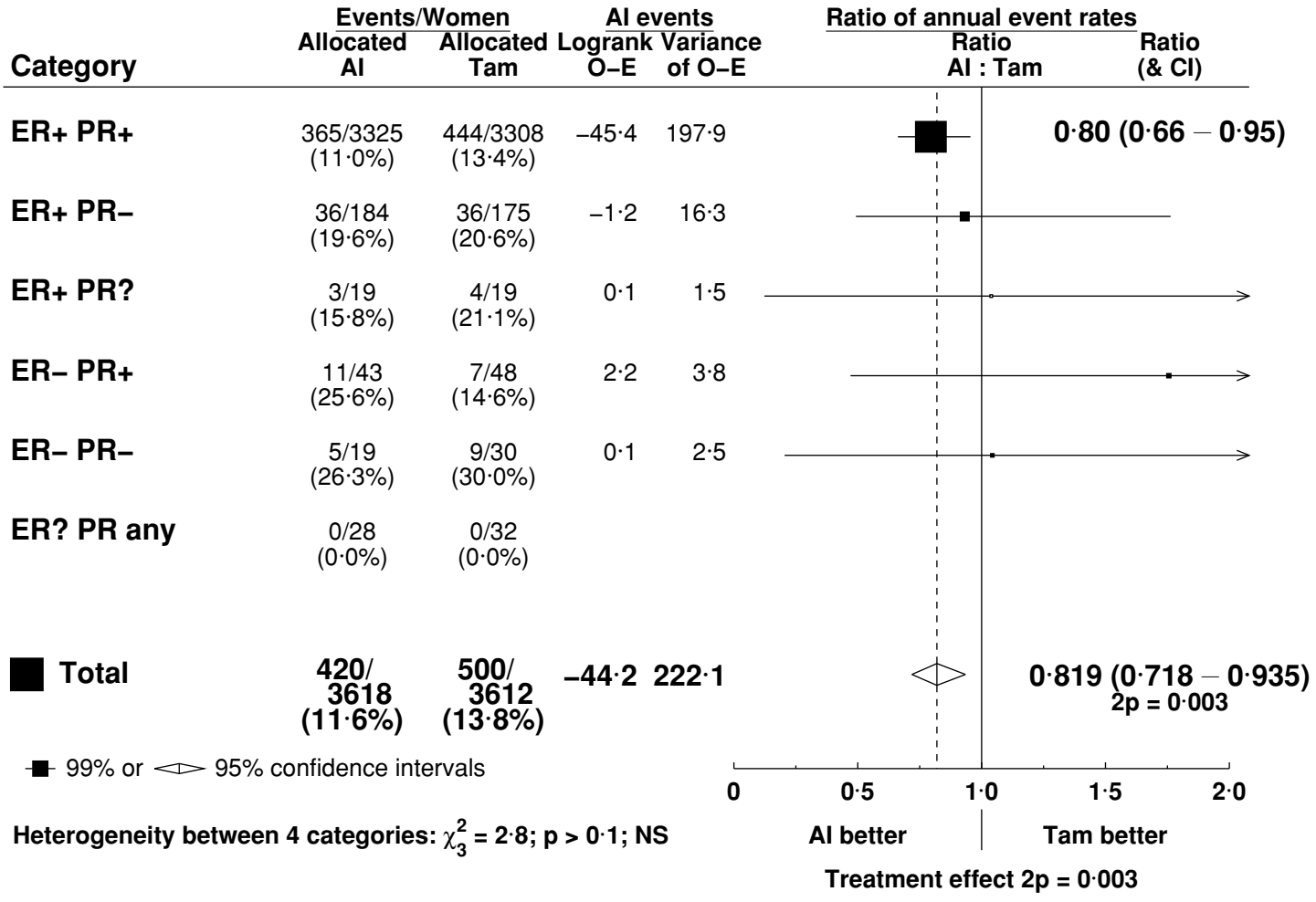
P8: Subgroup analyses of AI vs tamoxifen in premenopausal women; distant recurrence at any time



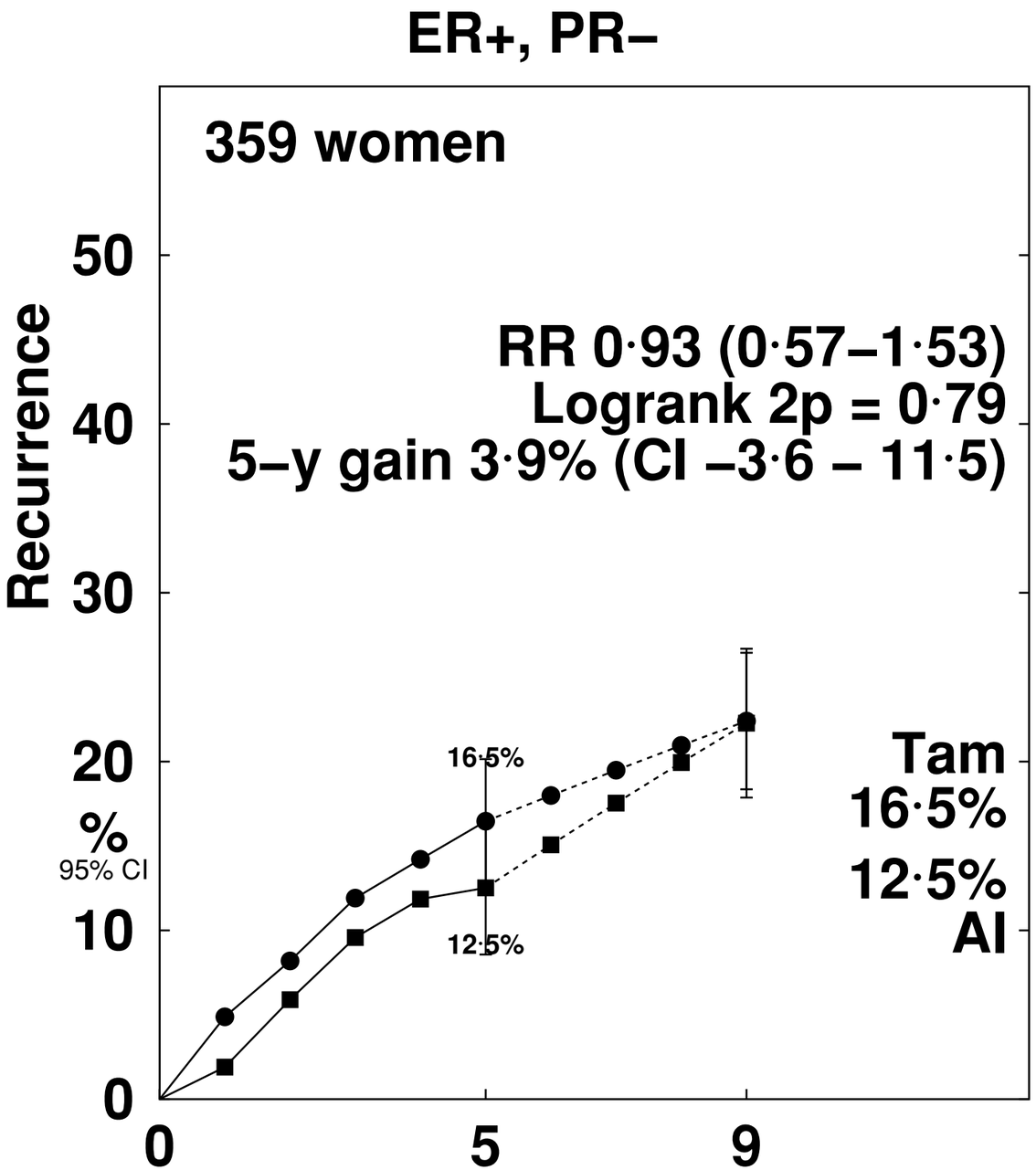
P9: Subgroup analyses of AI vs tamoxifen in premenopausal women; breast cancer mortality



P10: Subgroup analyses of AI vs tamoxifen in premenopausal women; any recurrence by ER and PR

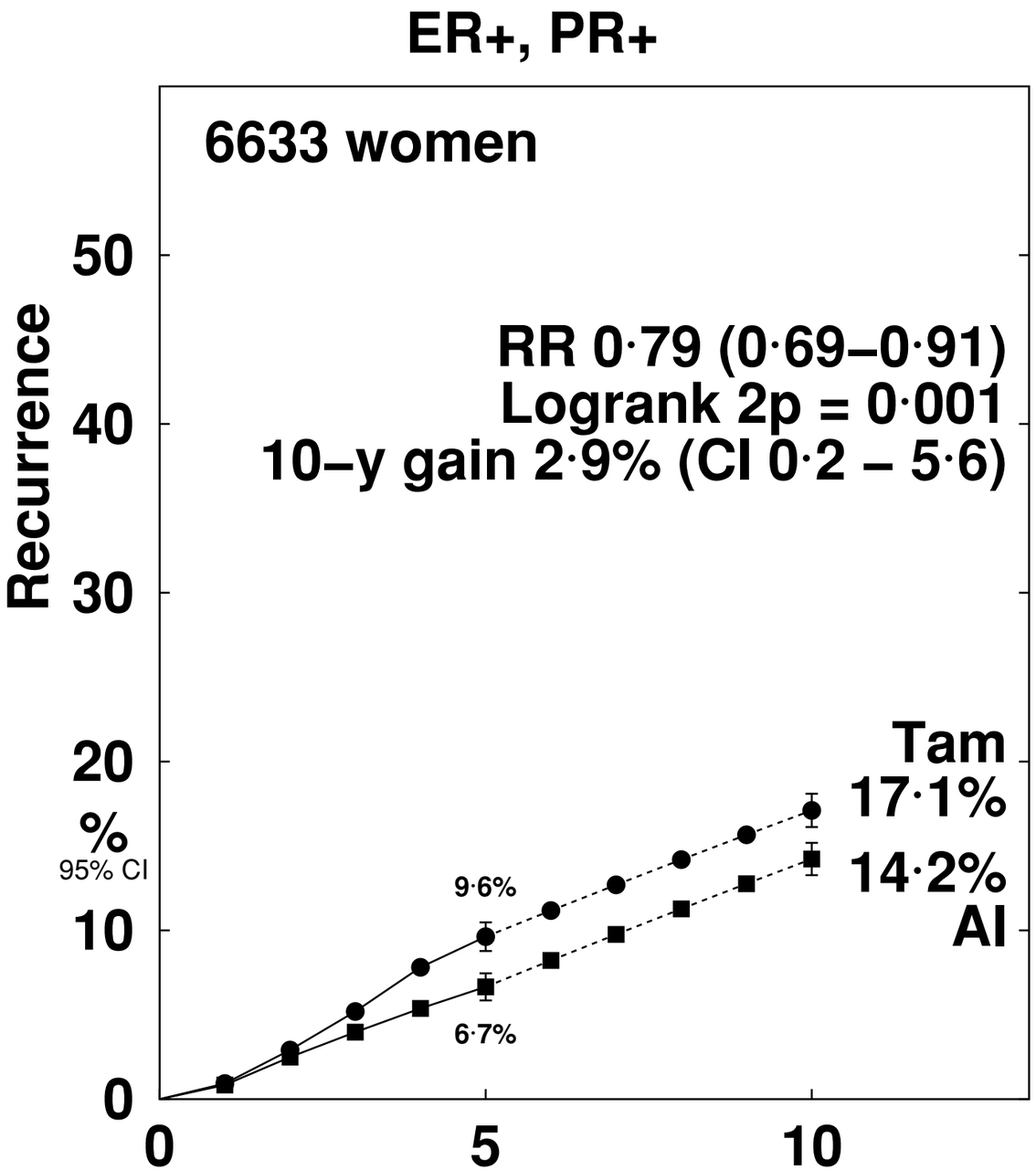


P11: 10 year cumulative risk recurrence by PR status with smoothing from year 5



Recurrence rates (% / year) and logrank analyses

Allocation	Years 0 – 4	Year 5+
AI	2·80 (23 / 821)	2·91 (13 / 446)
Tam	3·57 (28 / 785)	1·86 (8 / 431)
Rate ratio, from	0·75 CI 0·42 – 1·34	1·65 CI 0·64 – 4·20
(O–E) / V	–3·3 / 11·5	2·2 / 4·4

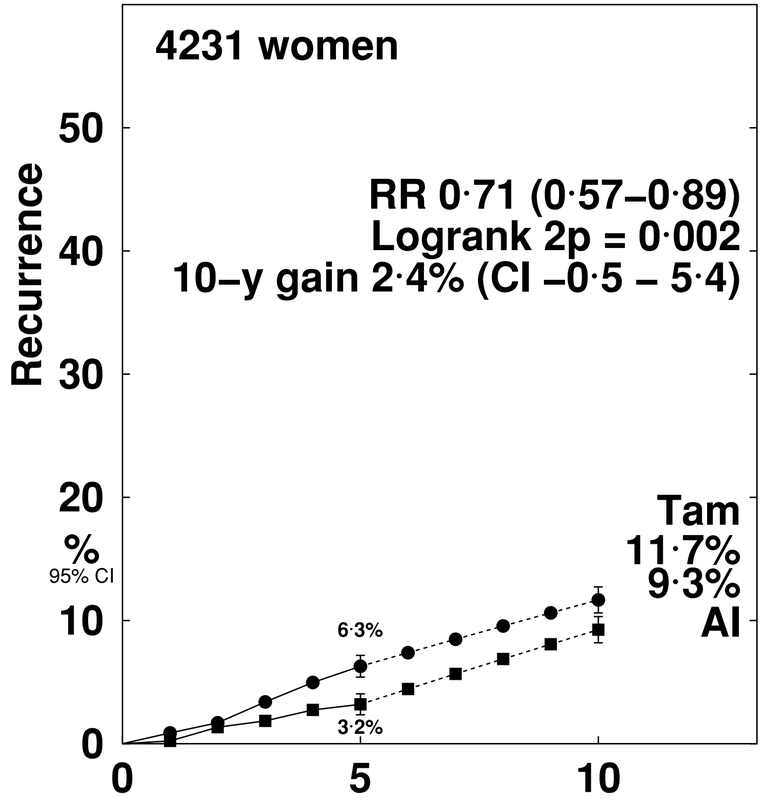


Recurrence rates (% / year) and logrank analyses

Allocation	Years 0 – 4	Years 5 – 9	Year 10+
AI	1·38 (212 / 15368)	1·70 (145 / 8524)	1·74 (8 / 461)
Tam	2·00 (302 / 15096)	1·72 (141 / 8197)	0·23 (1 / 438)
Rate ratio, from	0·69 CI 0·57 – 0·82	0·98 CI 0·77 – 1·24	4·44 CI 1·19 – 16·59
(O–E) / V	–47·2 / 124·9	–1·5 / 69·0	3·3 / 2·2

P12: 10 year cumulative risk recurrence by nodal status smoothed from year 5

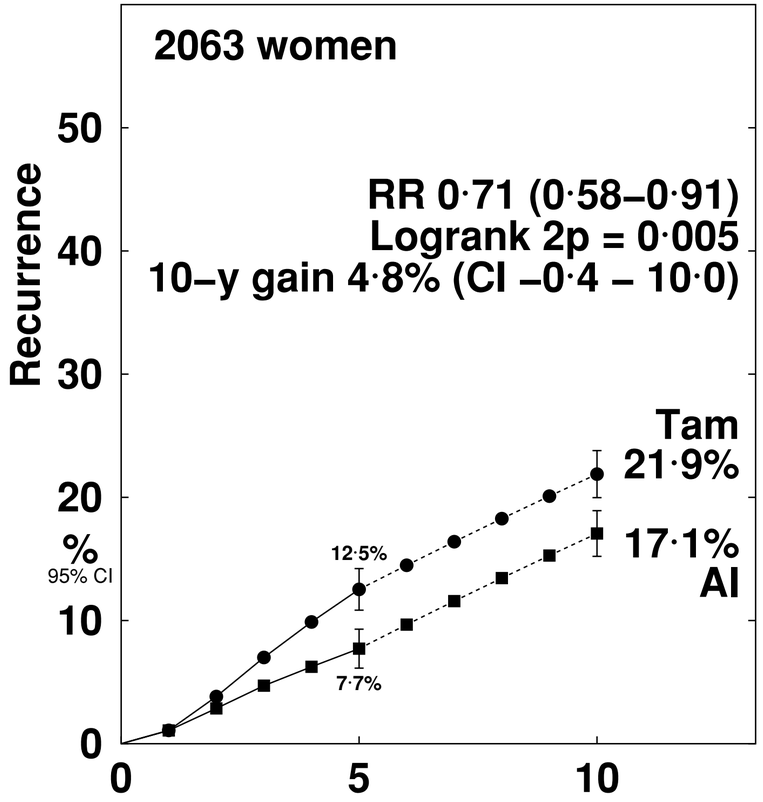
ER+, N0



Recurrence rates (% / year) and logrank analyses

Allocation	Years 0 – 4	Years 5 – 9	Year 10+
AI	0·66 (66 / 9994)	1·33 (74 / 5575)	1·48 (4 / 270)
Tam	1·28 (125 / 9741)	1·14 (61 / 5328)	0·71 (2 / 280)
Rate ratio, from (O–E) / V	0·52 CI 0·39 – 0·69 –31·2 / 47·0	1·09 CI 0·78 – 1·54 2·9 / 32·8	1·78 CI 0·35 – 8·98 0·8 / 1·5

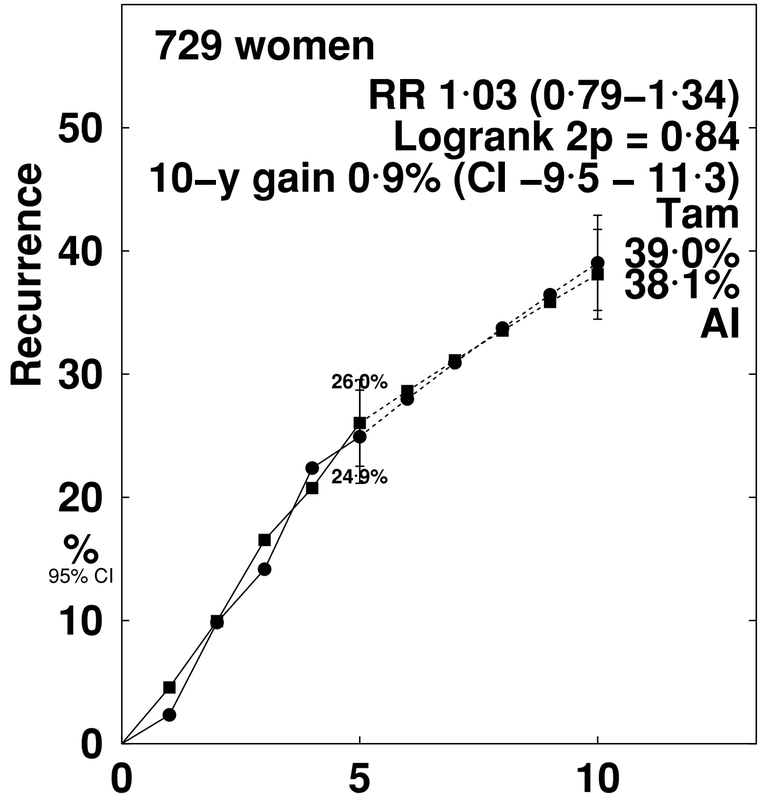
ER+, N1–3



Recurrence rates (% / year) and logrank analyses

Allocation	Years 0 – 4	Years 5 – 9	Year 10+
AI	1·62 (76 / 4697)	2·14 (57 / 2661)	1·21 (2 / 165)
Tam	2·62 (124 / 4729)	2·26 (58 / 2567)	0·0(0/130)
Rate ratio, from (O–E) / V	0·61 CI 0·46 – 0·81 –24·2 / 48·6	0·94 CI 0·65 – 1·37 –1·6 / 27·8	6·69 CI 0·42 – 107·56 0·9 / 0·5

ER+, N4+

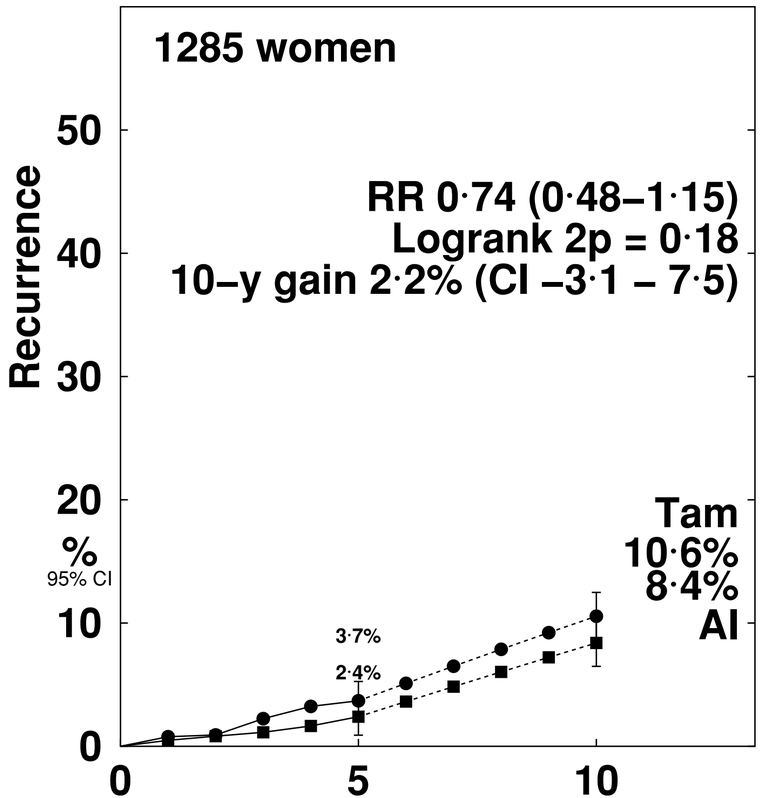


Recurrence rates (% / year) and logrank analyses

Allocation	Years 0 – 4	Years 5 – 9	Year 10+
AI	5·98 (94 / 1572)	3·55 (27 / 761)	7·50 (3 / 40)
Tam	5·62 (83 / 1477)	4·19 (31 / 740)	0·0(0/47)
Rate ratio, from (O–E) / V	1·06 CI 0·78 – 1·44 2·3 / 41·2	0·85 CI 0·50 – 1·46 –2·1 / 13·5	7·62 CI 0·70 – 83·24 1·4 / 0·7

P13: 10 year cumulative risk recurrence by tumour grade smoothed from year 5

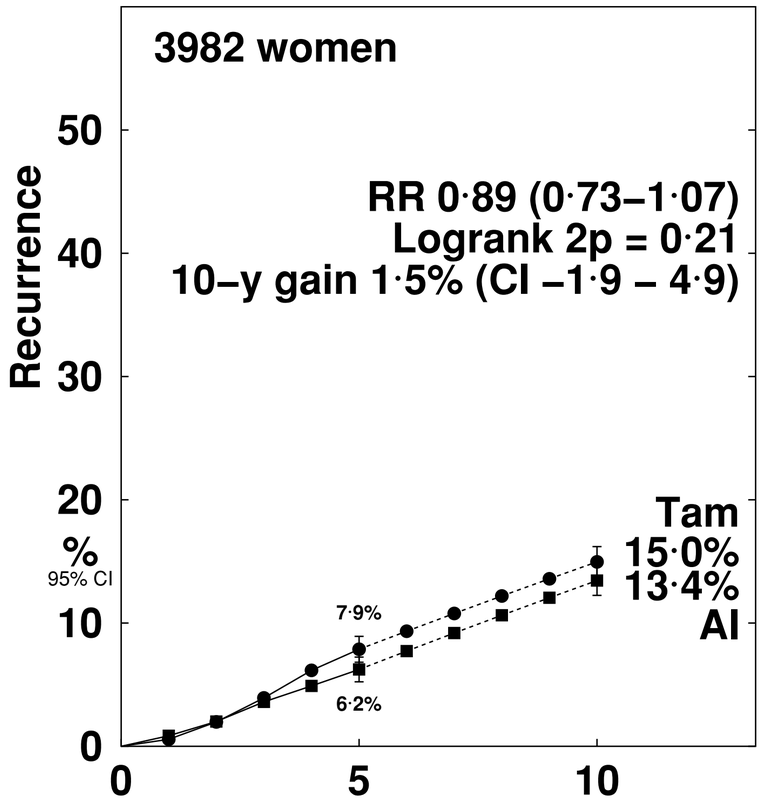
ER+, Well–differentiated



Recurrence rates (% / year) and logrank analyses

Allocation	Years 0 – 4	Years 5 – 9	Year 10+
AI	0·48 (14 / 2935)	1·30 (22 / 1695)	0·0(0/116)
Tam	0·75 (23 / 3070)	1·44 (26 / 1804)	0·77 (1 / 130)
Rate ratio, from	0·65 CI 0·34 – 1·24	0·86 CI 0·47 – 1·56	0·21 CI 0·00 – 13·02
(O–E) / V	–3·9 / 9·0	–1·6 / 10·6	–0·4 / 0·2

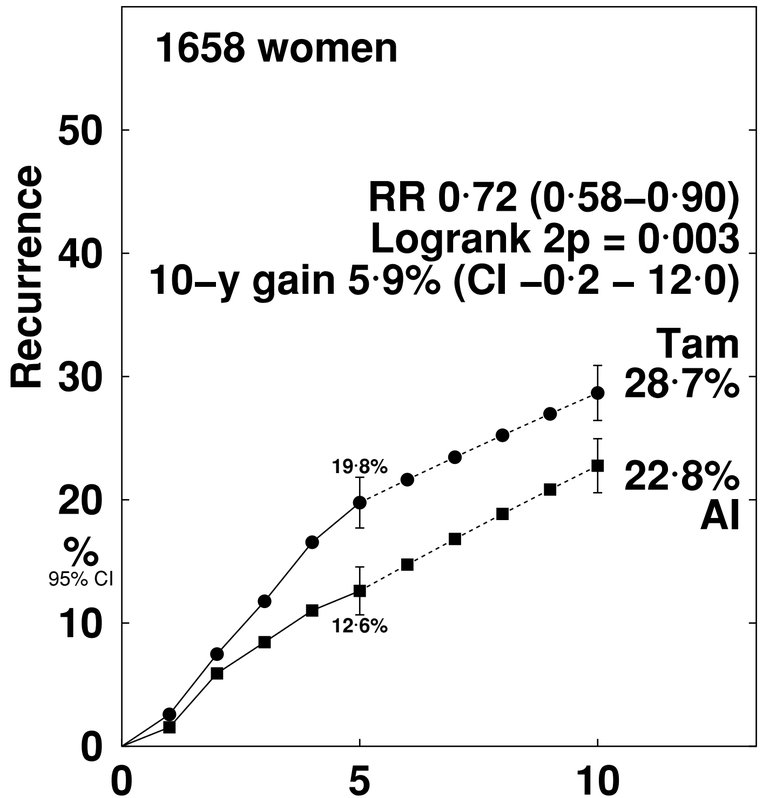
ER+, Moderately



Recurrence rates (% / year) and logrank analyses

Allocation	Years 0 – 4	Years 5 – 9	Year 10+
AI	1·26 (117 / 9257)	1·59 (81 / 5097)	1·95 (5 / 256)
Tam	1·63 (150 / 9183)	1·61 (80 / 4955)	0·0(0/241)
Rate ratio, from	0·79 CI 0·62 – 1·01	1·00 CI 0·73 – 1·37	7·04 CI 1·18 – 41·87
(O–E) / V	–15·1 / 64·5	–0·0 / 39·1	2·4 / 1·2

ER+, Poorly

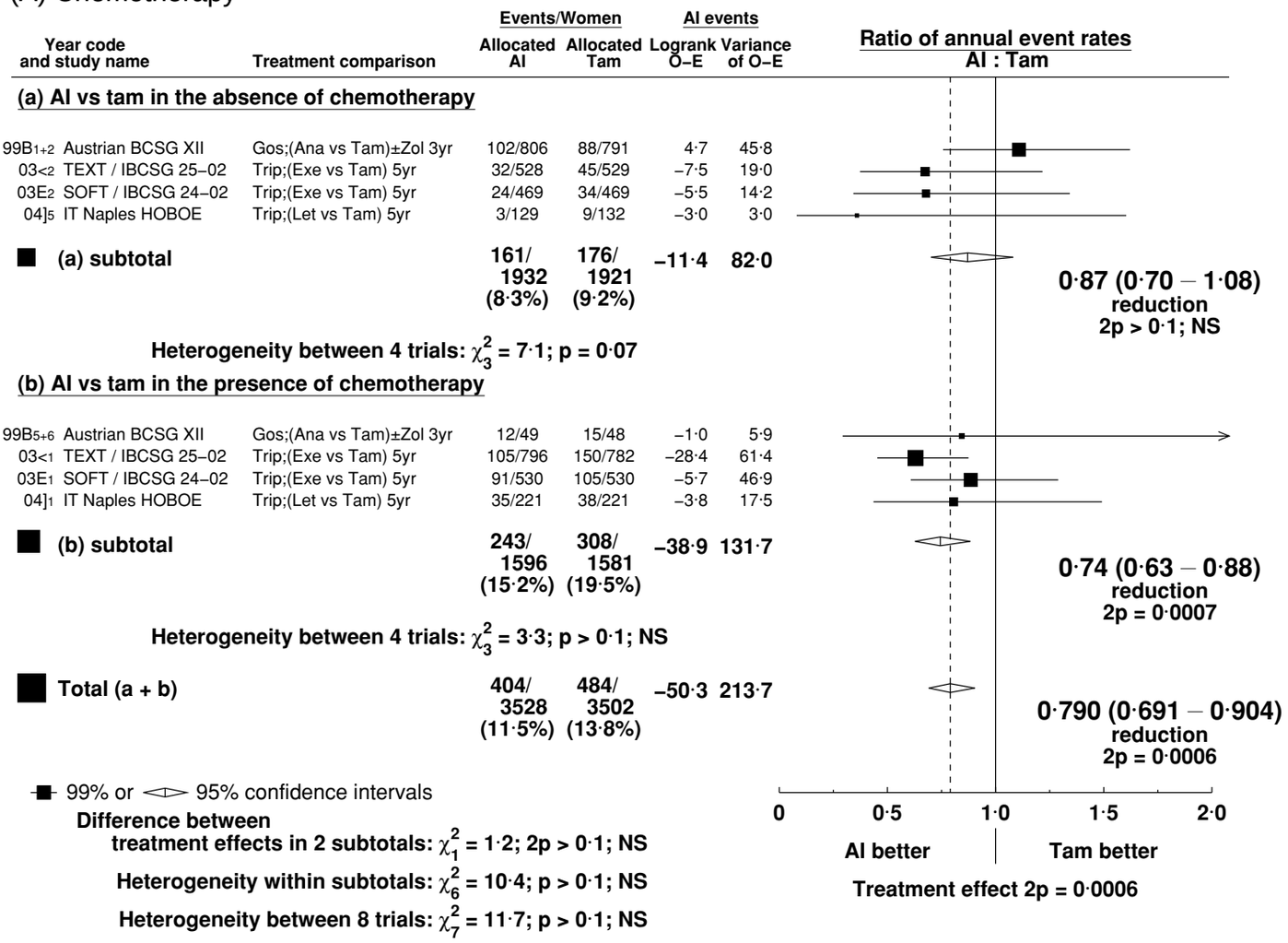


Recurrence rates (% / year) and logrank analyses

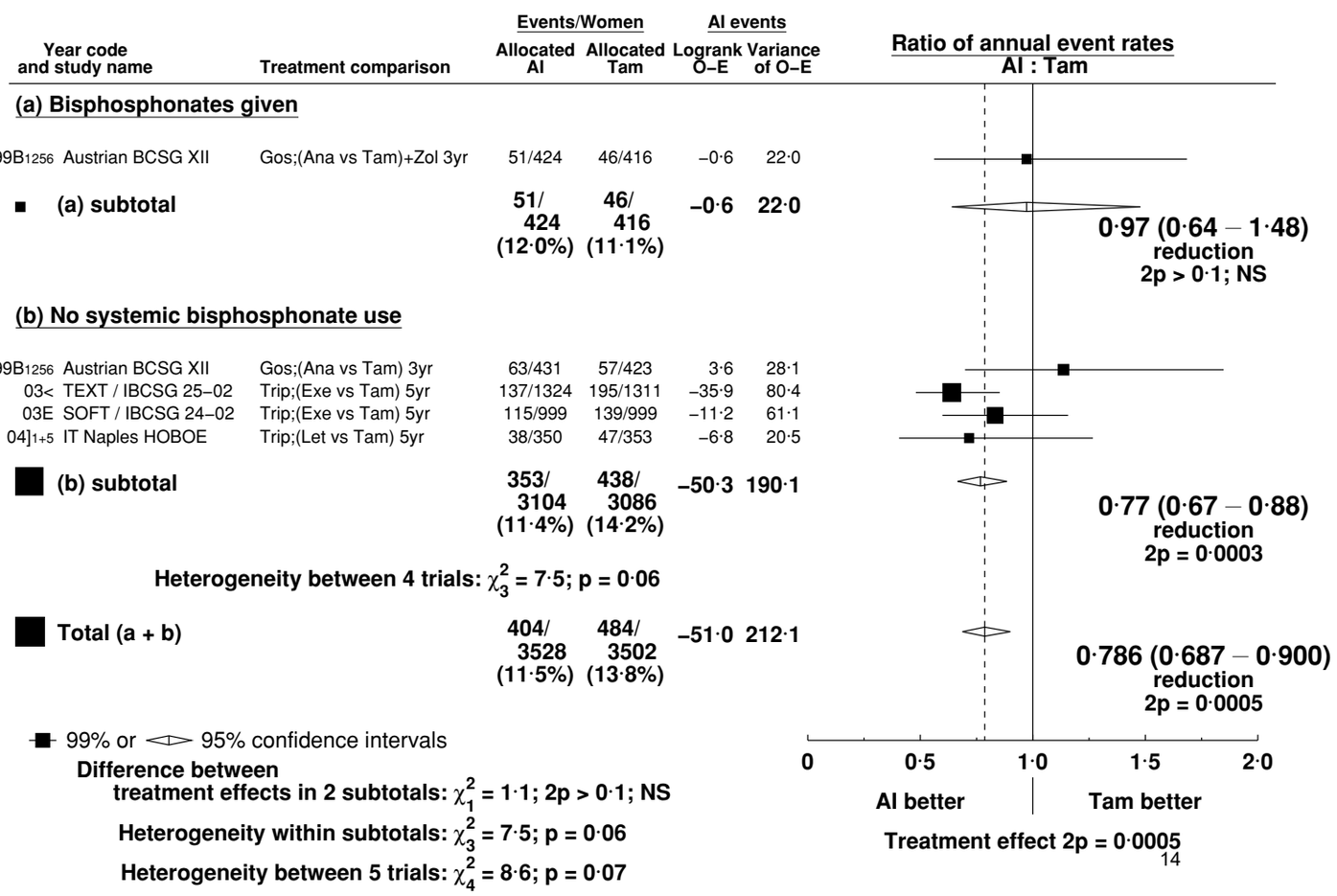
Allocation	Years 0 – 4	Years 5 – 9	Year 10+
AI	2·74 (105 / 3828)	2·53 (53 / 2091)	4·04 (4 / 99)
Tam	4·41 (155 / 3516)	2·27 (41 / 1807)	1·16 (1 / 86)
Rate ratio, from	0·62 CI 0·48 – 0·80	1·05 CI 0·69 – 1·60	2·46 CI 0·40 – 15·11
(O–E) / V	–29·5 / 61·5	1·1 / 21·9	1·0 / 1·2

P14: Any recurrence in AI versus tamoxifen in premenopausal women: (A) in the presence and absence absence chemotherapy, (B) in the presence and absence of bisphosphonate

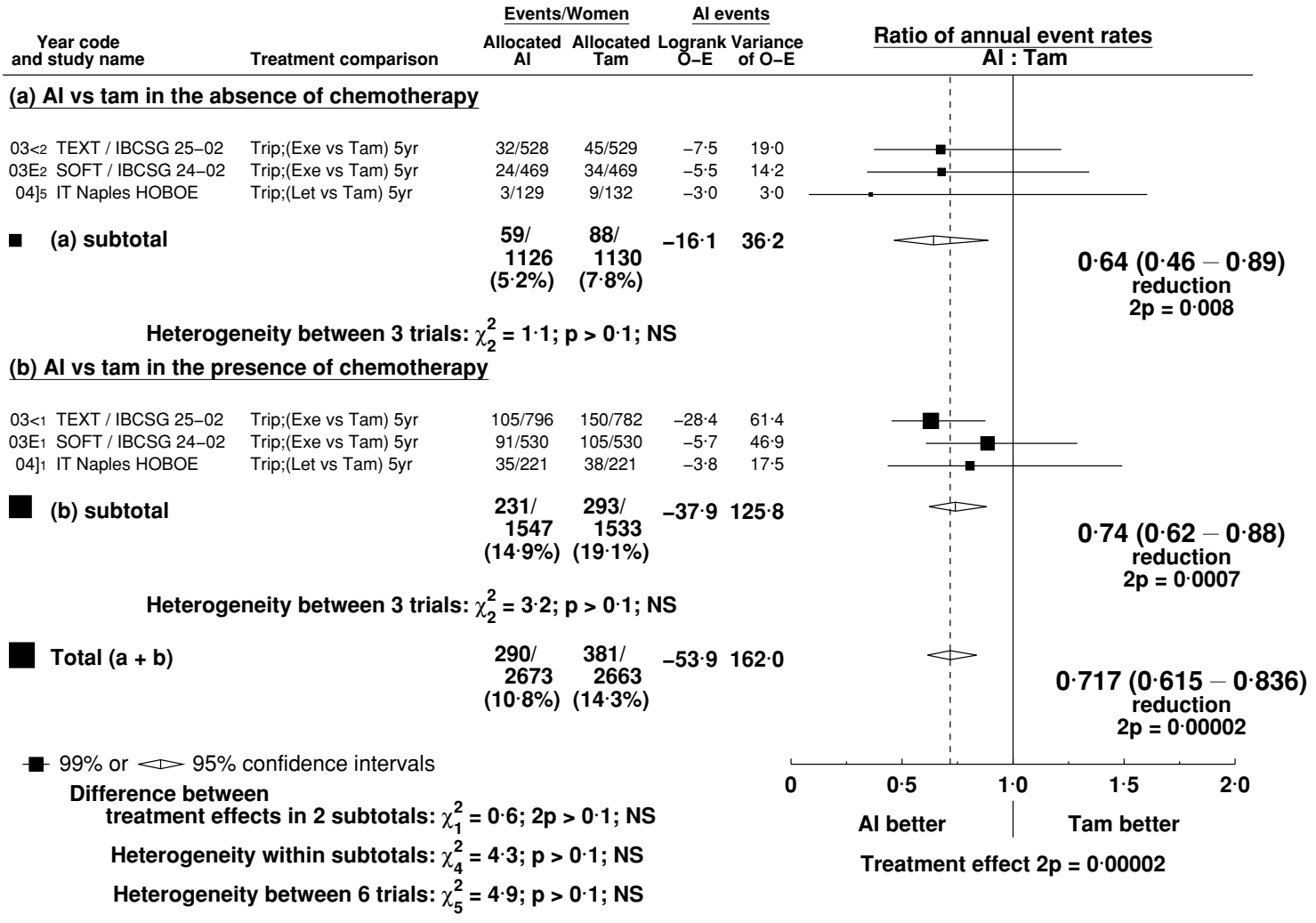
(A) Chemotherapy



(B) Randomised bisphosphonate use



P15: Any recurrence in AI versus tamoxifen in premenopausal women: in the presence and absence of chemotherapy (excluding ABCSG XII)

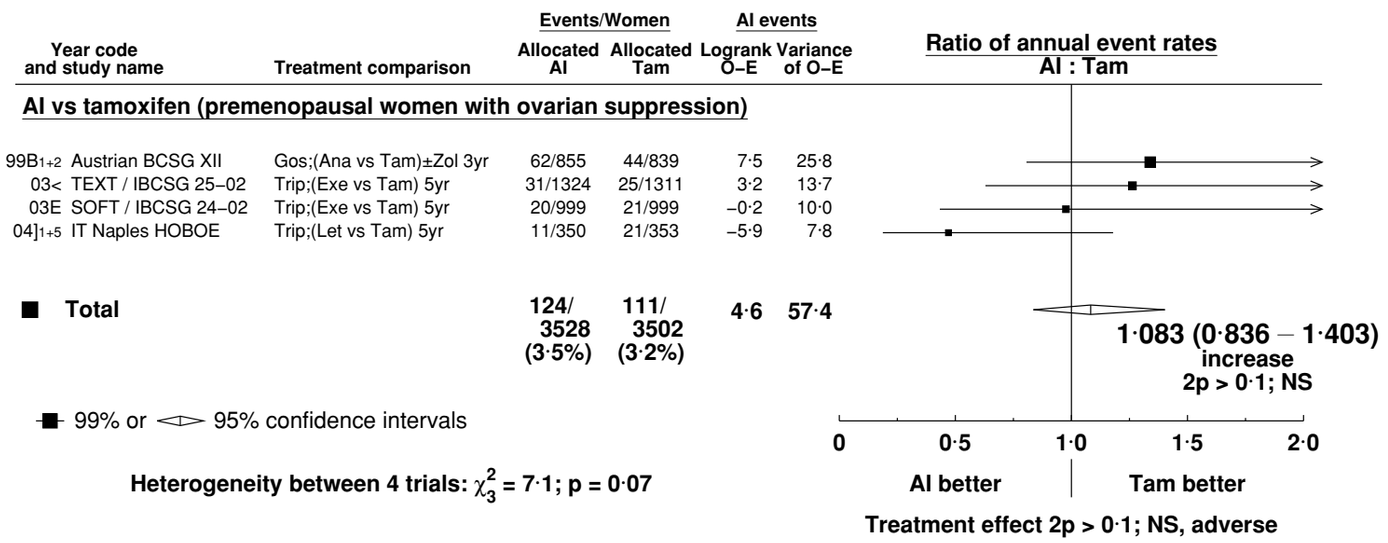


Cause of death	AI	Tam	(O-E)	Variance	RR	95% CI	p
Cancers other than breast	**22**	**10**	**6.6**	**7.4**	**2.44**	**(1.19 – 5.01)**	**0.02**
Haematological	3	1	1.0	1.0	2.7	(0.38 – 19.2)	0.32
Lung	6	4	1.0	2.5	1.49	(0.43 – 5.15)	0.53
Ovarian	4	1	1.5	1.3	3.17	(0.57 – 17.7)	0.19
Pancreatic	4	1	1.5	1.3	3.17	(0.57 – 17.7)	0.19
Other non-breast*	5	3	1.6	1.3	3.42	(0.61 – 19.1)	0.16
Vascular disease	**2**	**1**	**0.4**	**0.7**	**1.77**	**(0.17 – 18.4)**	**0.63**
Other specified disease**	**3**	**9**	**-3.0**	**3.1**	**0.38**	**(0.13 – 1.16)**	**0.09**
Unknown	**3**	**4**	**-0.6**	**1.7**	**0.70**	**(0.16 – 3.16)**	**0.65**
Total	**30**	**24**	**3.4**	**12.9**	**1.30**	**(0.75 – 2.25)**	**0.34**

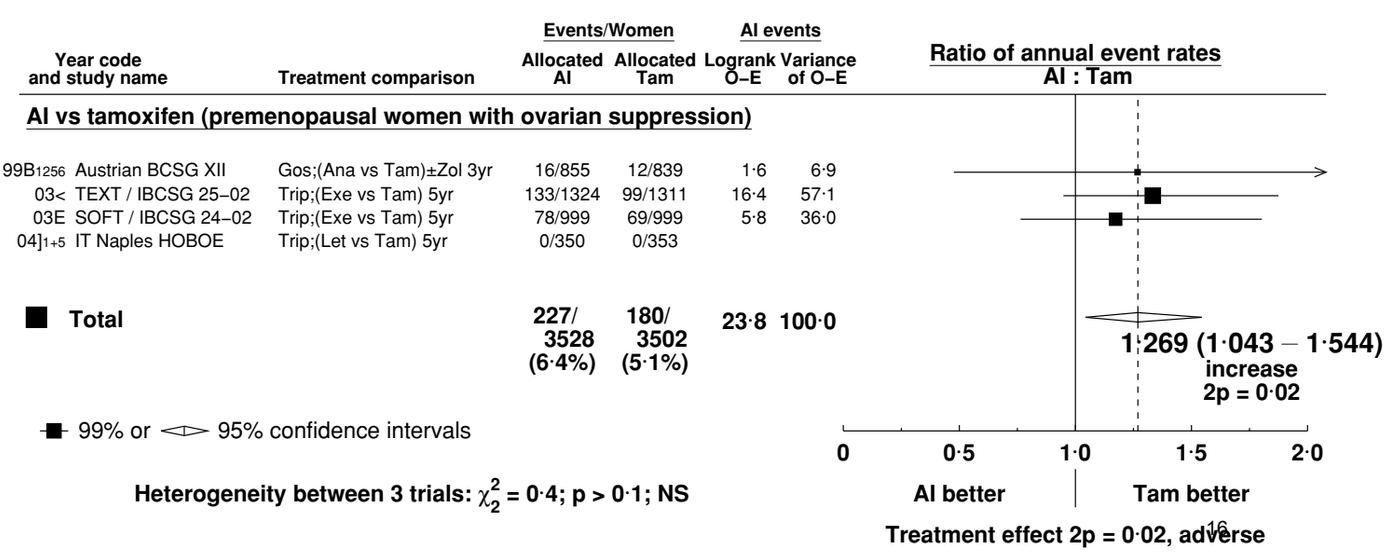
*Other non-breast cancer deaths in AI group: colorectal, gastric, head & neck, skin (2). In Tam group: brain, primary liver, uterine.

**Other specified disease in AI group: hepatic disease, other specified not breast cancer (2). In Tam group: diabetes, mixed drug intoxication, suicide (2), accident, infectious/parasitic, other specified not breast cancer (3).

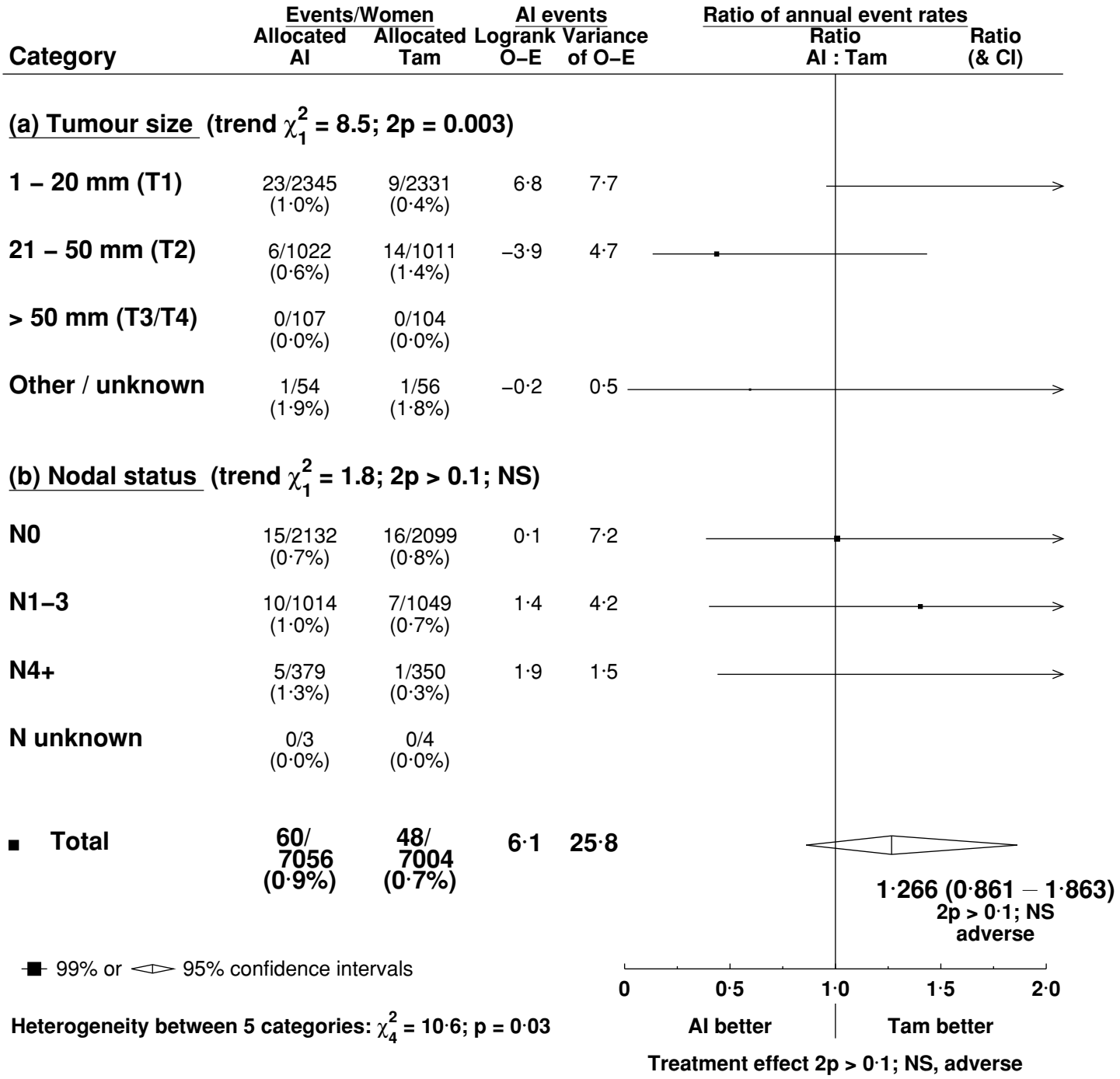
SECOND PRIMARY CANCER INCIDENCE in trials of Aromatase inhibitor versus tamoxifen, ER+



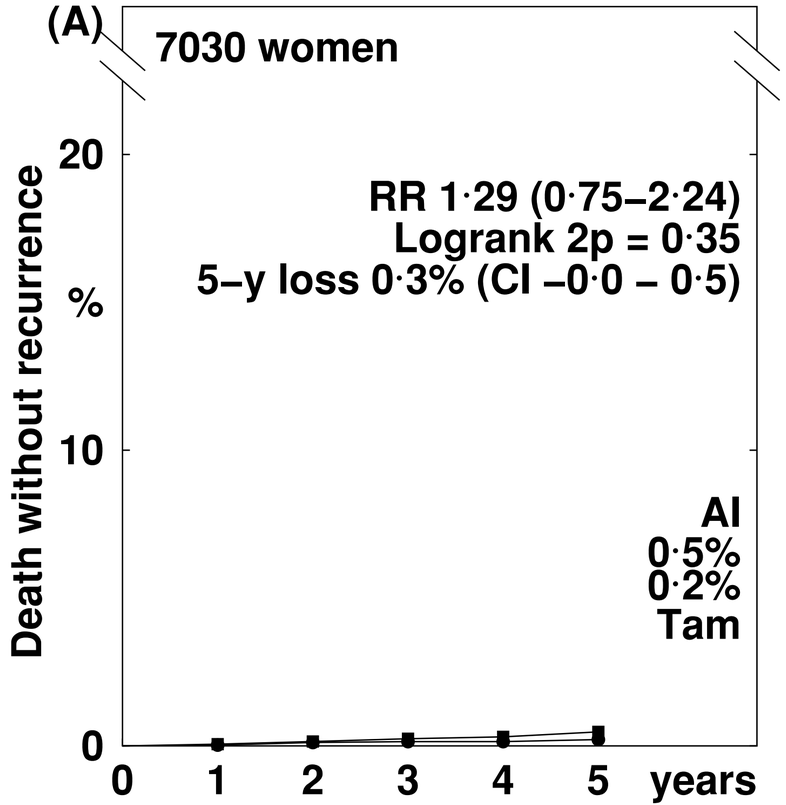
BONE FRACTURE INCIDENCE in trials of Aromatase inhibitor versus tamoxifen, ER+



P17: Death without recurrence by tumour size and nodal status

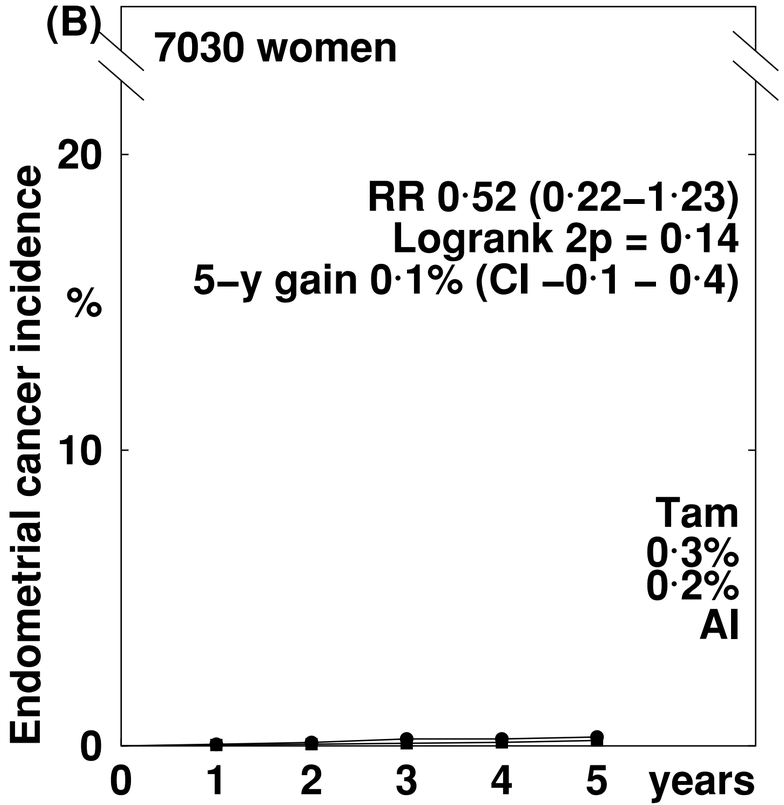


P18: 5 year cumulative risk recurrence by (A) death without recurrence, (B) endometrial cancer incidence, (C) bone fracture incidence



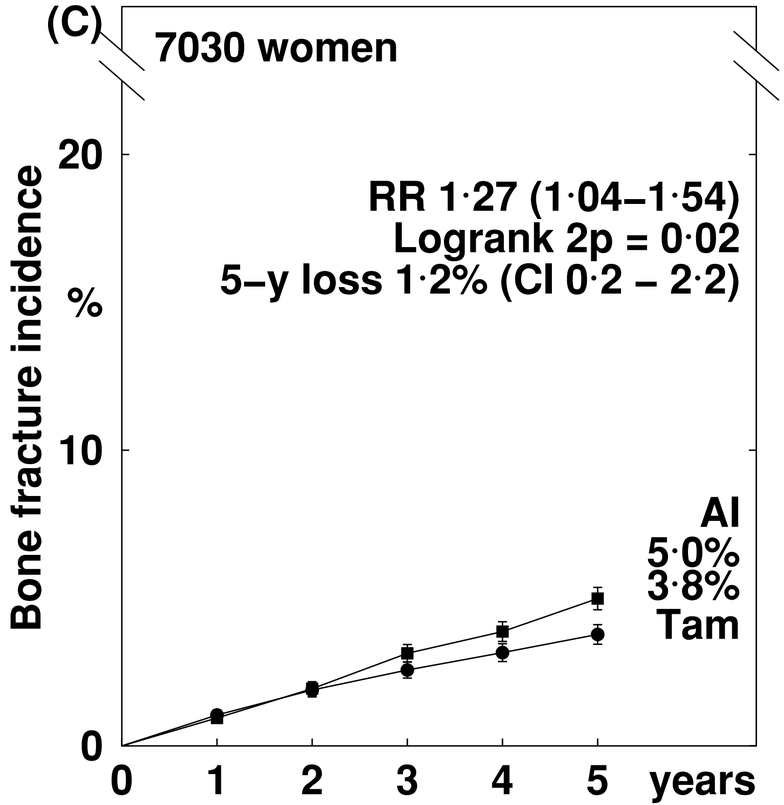
Death–without–recurrence rates (% / year) and logrank analyses

Allocation	Years 0 – 1	Years 2 – 4	Year 5+
AI	0·07 (5 / 6854)	0·11 (10 / 9421)	0·16 (15 / 9476)
Tam	0·06 (4 / 6805)	0·03 (3 / 9152)	0·19 (17 / 9098)
Rate ratio, from	1·29 CI 0·35 – 4·78	2·94 CI 0·99 – 8·73	0·90 CI 0·44 – 1·86
(O–E) / V	0·6 / 2·2	3·5 / 3·2	–0·7 / 7·4



Event rates (% / year) and logrank analyses

Allocation	Years 0 – 1	Years 2 – 4	Year 5+
AI	0·03 (2 / 6898)	0·04 (4 / 9669)	0·01 (1 / 9882)
Tam	0·06 (4 / 6868)	0·06 (6 / 9600)	0·05 (5 / 9827)
Rate ratio, from	0·51 CI 0·10 – 2·56	0·67 CI 0·19 – 2·32	0·32 CI 0·05 – 1·84
(O–E) / V	–1·0 / 1·5	–1·0 / 2·5	–1·4 / 1·2



Bone fracture incidence rates (% / year) and logrank analyses

Allocation	Years 0 – 1	Years 2 – 4	Year 5+
AI	0·98 (67 / 6853)	1·05 (99 / 9425)	0·65 (61 / 9356)
Tam	0·95 (65 / 6816)	0·65 (61 / 9416)	0·57 (54 / 9471)
Rate ratio, from	1·02 CI 0·73 – 1·44	1·62 CI 1·18 – 2·21	1·16 CI 0·80 – 1·68
(O–E) / V	0·8 / 32·4	18·9 / 39·3	4·2 / 28·4

Pages 19-20: Published adverse events in trials of AI verse tamoxifen in pre-menopausal women (% of women, unless otherwise stated)

Year code	Trial name	N	Bone fracture	Non-fatal cardiac events	Thrombosis or embolism	Stroke	Uterine	Compliance	Follow-up (median)
99B	ABCSG XII	1803	Ana vs Ana+Z vs Tam vs Tam+Z: * • Fracture: 0.2 vs 0 vs 0.2 vs 0.2; p=0.91 • Serious Fracture: 0.9 vs 1.6 vs 1.3 vs 0.9; p=0.75	Ana vs Ana+Z vs Tam vs Tam+Z: * • Tachycardia: 1.1 vs 2.2 vs 0.4 vs 2.0; p=0.07 • Serious tachycardia: 0.2 vs 0.2 vs 0.2 vs 0.0; p=1.00	Ana vs Ana+Z vs Tam vs Tam+Z: * • Thrombosis: 0.0 vs 0.2 vs 0.0 vs 0.0; p=0.50 • Serious thrombosis: 0.0 vs 0.0 vs 0.7 vs 1.1; p=0.01	Not reported	Ana vs Ana+Z vs Tam vs Tam+Z: * • Uterine polyp: 0.2 vs 0.2 vs 0.0 vs 1.1; p=0.07 • Serious uterine polyp: 1.5 vs 1.1 vs 8.9 vs 11.4; p<0.001	Number of patients who 'did not comply with treatment': Ana vs Ana+Z vs Tam vs Tam+Z: 5 v 2 v 0 v 2 Patient compliance with endocrine therapy was confirmed during the 3-year treatment phase: • Ana alone: in 4362 (89%) of 4884 visits • Ana+Z: in 4769 (98%) of 4886 visits • Tam alone: in 4492 (93%) of 4823 visits • Tam + Z: in 4450 (93%) of 4778 visits	47.8 months
03E 03<	SOFT TEXT	3066 2672	Ex+OS vs Tam vs Tam+OS: ¶ • any: 7.7 vs 5.3 vs 6.0 • grade 3 or 4: 1.6 vs 0.8 vs 1.0	Ex+OS vs Tam vs Tam+OS: ¶ cardiac ischemia or infarction: • any: 0.7 vs 0.5 vs 0.4 • grade 3 or 4: 0.3 vs 0.4 vs 0.3 (One patients in the Tam group in SOFT had a grade 5 event)	Ex+OS vs Tam vs Tam+OS: ¶ • any event: 1.2 vs 2.2 vs 2.3 • grade 3 or 4: 0.9 vs 1.7 vs 2.0	Ex+OS vs Tam vs Tam+OS: ¶ CNS cerebrovascular ischemia: • any event: 0.3 vs 0.6 vs 0.4 • grade 3 or 4: 0.2 vs 0.4 vs 0.3 CNS haemorrhage: • any event: 0.8 vs 1.5 vs 1.1 • grade 3 or 4: <0.1 vs 0.0 vs 0.1	‡ Ex+OS vs Tam+OS Gynaecological cancer: 0.3 (7/2346) vs 0.4 (9/2344), including endometrial cancer: 0.1 (2/2346) vs 0.2 (5/2344)	Exe+OS vs Tam vs Tam+OS: ¶ Discontinuation of endocrine therapy (with/without alternative therapy): 23.7 vs 22.5 vs 19.3	9 years
04]	HOBOE	483	No bone fracture.	Let vs Tam: • Arrhythmia: - any: 3.3 vs 3.1 - severe: 0 vs 0 • Cardiac ischemia: - any 0.3 vs 0; - severe 0.3 vs 0. • Cardiovascular – other: - any: 2.2 vs 3.1 - severe: 0 vs 0	Let vs Tam: • any: 0 vs 0.3 • severe: 0 vs 0	Let vs Tam: Central nervous system cerebrovascular ischemia: • any: 0 vs 0.6 • severe: 0 vs 0.6	Not reported	Let vs Tam: • Treatment stopped because of toxicity: 6.7 vs 5.6 • Treatment stopped because of refusal: 0.6 vs 1.7 • Treatment completed per protocol: 59.3 vs 49.2	64 months

Abbreviations: Exe=exemestane. N=number of women randomised (unless otherwise specified). Let=Letrozole. OS= ovarian suppression. Tam=tamoxifen. Z=zoledronic acid.

Data in the table is from most recent publication unless '‡' indicated then from a previous publication

* P values are for a four-group comparison according to Fisher's exact test. Serious adverse events: any adverse events that were lethal or life-threatening, resulted in permanent damage, required inpatient hospitalization or extension of inpatient treatment, or placed the patient at risk and necessitated medical or surgical intervention.

† Data based on the 4643 patients in the safety population in SOFT and TEXT who received a protocol-assigned treatment.

¶ Data based on safety populations in SOFT and TEXT who initiated a protocol-assigned treatment, including 1005 patients who were randomly assigned to receive tamoxifen in SOFT and 4643 patients who were randomly assigned to receive tamoxifen plus ovarian suppression (2326 patients) or exemestane plus ovarian suppression (2317 patients) in SOFT or TEXT

Reference

Year code	Trial name	Reference
99B	ABCSG XII	Gnant M, Mlineritsch B, Schippinger W, Luschin-Ebengreuth G, Pöstlberger S, Menzel C, Jakesz R, Seifert M, Hubalek M, Bjelic-Radisic V, Samonigg H. Endocrine therapy plus zoledronic acid in premenopausal breast cancer. *N Engl J Med*. 2009;360(7):679-691. Gnant M, Mlineritsch B, Stoeger H, et al. Adjuvant endocrine therapy plus zoledronic acid in premenopausal women with early-stage breast cancer: 62-month follow-up from the ABCSG-12 randomised trial. *Lancet Oncol* 2011;12(7):631-41.
03E 03<	SOFT TEXT	Pagani O, Regan MM, Walley BA, Fleming GF, Colleoni M, Láng I, Gomez HL, Tondini C, Burstein HJ, Perez EA, Ciruelos E. Adjuvant exemestane with ovarian suppression in premenopausal breast cancer. *N Engl J Med*. 2014;371(2):107-118. Francis PA, et al. Tailoring adjuvant endocrine therapy for premenopausal breast cancer. *N Engl J Med*. 2018;379(2):122-137.
04]	HOBOE	Nuzzo F, et al. Bone effect of adjuvant tamoxifen, letrozole or letrozole plus zoledronic acid in early-stage breast cancer: the randomized phase 3 HOBOE study. *Ann. Oncol*. 2012;23(8):2027-2033. Perrone F, et al. Adjuvant zoledronic acid and letrozole plus ovarian function suppression in premenopausal breast cancer: HOBOE phase 3 randomised trial. *Eur J Cancer*. 2019;118:178-186.

Pages 21-29: Statistical Analysis Plan

DRAFT Statistical Analysis Plan for a Meta-Analysis of trials of Aromatase inhibitors versus tamoxifen in premenopausal ER+ women with ovarian suppression
November 11th 2021, Version 1.2

Objective

To collate individual patient data to characterise the benefits and risks of trials comparing aromatase inhibitors (AI) with tamoxifen in premenopausal ER+ women with ovarian suppression.

Analyses

Primary analyses will be by Intention-to-Treat (ITT), including all randomised patients irrespective of treatment compliance.

Sensitivity analyses will also be undertaken to investigate the potential impact of non-compliance on treatment efficacy.

The primary analyses will be of the following comparisons (trials included in each section are listed in appendix 1):

- Trials of AI versus tamoxifen in premenopausal ER+ women with operable breast cancer who have undergone ovarian ablation or ovarian suppression

Primary outcomes

The main endpoint definitions and analyses methods are those used in previous EBCTCG reports, with amendments reflecting the potential impact of endocrine therapy.

1. **Time to recurrence:** includes distant recurrence, local recurrence and new second primary breast cancer (ipsilateral or contralateral), and the definitions of these will be as in each trial.
2. **Breast cancer mortality:** Information about mortality rates without recurrence will be subtracted from information about overall mortality rates (as in previous EBCTCG reports). The same statistical methods will be used to construct Kaplan-Meier graphs that estimate breast cancer mortality (i.e. the pattern of mortality that would have been seen if it had been possible to avoid all deaths before or after recurrence from causes other than breast cancer).
3. **All-cause mortality**

Exploratory endpoints

1. **Time to first distant recurrence:** includes distant recurrence and ignores any prior loco-regional or contralateral recurrences. Information about the site(s) of first distant recurrence will be collected to allow additional exploratory analyses to be performed. Patients can have more than one site recorded as their first distant recurrence. Categorisation of sites will include, but is not limited to: liver, lung, bone, node, soft tissue, brain.
2. **Time to loco-regional recurrence as first event**: includes ipsilateral breast, chest wall and locoregional lymph nodes (axilla and SCF).
3. **Time to new contralateral breast cancer**
4. **Death without recurrence:** i.e. without the secretariat having any record of recurrence. Because the quality of the recurrence data in the AI trials is likely to be reasonably good, deaths from a wholly unknown cause without record of recurrence will be treated as deaths from an unknown cause that was not breast cancer.
5. **New primary endometrial cancer rates:** Endocrine therapy with tamoxifen is associated with an increased risk of endometrial cancers though aromatase inhibitors may reduce the risk of endometrial cancer.
6. **Fracture rates:** Information on fractures, including date of first fracture and site where available will be collected.

7. **Safety –** Data on other non-fatal adverse events in individual trials (e.g. ischemic heart disease, stroke, pulmonary embolus, DVT) will be sought and analysed.
8. **Incidence and site of second cancers** – Quality assurance checks will be undertaken to compare the incidence of second cancers (overall and by site) with that of distant recurrence by age group to characterise whether these are true second cancers or miscoded recurrences. Where there is confidence that second cancers are correctly diagnosed, the incidence of second cancers before any recurrence of breast cancer will be provided by treatment group for each cohort. In addition, second cancers will be categorised by site.

Subgroup analyses:

Exploratory subgroup analyses will be undertaken but, given the well-known hazards of subgroup analysis, will be interpreted appropriately cautiously. Investigation of potential interactions between tumour or patient characteristics and treatment efficacy will be undertaken with breast cancer recurrence as primary outcome. However, if almost all of the benefits of AI over tamoxifen are seen in the period when the treatments differed, as they were in the EBCTCG meta-analysis of AI vs tamoxifen in postmenopausal women (Lancet 2015), then subgroup analyses of recurrence just in this period will also be undertaken to enhance the statistical power to investigate any variability in treatment efficacy by patient or tumour characteristics.

Subgroup analyses of Recurrence Forest plots for subgroup analyses by:

- Site of recurrence (distant metastasis, local recurrence or contralateral breast cancer)
- Age (<45, 45-54, 55-69, ≥70, unknown),
- PR status (PR-, PR+, PR unknown)
- HER2:CEP17 ratio (HER2-, HER2+, HER2 unknown)
- Nodal status (N0, N1-3, N4+, N unknown)
- Tumour stage (T1, T2, T3/T4, T stage unknown);
- Histological grade (1, 2, 3, unknown),
- Proliferation index (%Ki-67: 0-9, 10-19, 20+, unknown)
- Tumour histology (ductal, lobular, other, unknown)
- Presence/absence of chemotherapy
- Period of follow-up years 0-1, 2-4, 5-9, and 10+ after randomisation

Tests of heterogeneity and of trend

First calculate the log-rank statistic (o-e) and its variance v in each separate stratum, and add these up to get the overall logrank (O-E) and it variance V (i.e. the sum of the separate variances). Delete any uninformative strata (i.e. those for which v is zero), and number the remaining strata from 1 to n. A χ^2 test (on n-1 degrees of freedom) for heterogeneity between the treatment effects in different strata can be obtained by subtracting $(O-E)^2/V$ from the sum of the separate values, one per stratum, of $(o-e)^2/v$.

Alternatively, a χ^2 test for trend (i.e. for whether the treatment effect changes progressively from one stratum to the next) will be calculated as follows: if the stratum numbered s has logrank statistics (o-e) and v then define m, the mean stratum number, to be the sum, one term per stratum, of sv/V and define T to be the sum, one term per stratum, of (s-m)(o-e). The variance of T, var(T), is then the sum, one term per stratum, of $(s-m)^2v$, and the χ^2 test (on 1 degree of freedom) for trend is $T^2/var(T)$. If there are only two strata then the tests for trend and heterogeneity are identical.

Appendix 1. Trial list

Trial (Yr) Code	Trial name	Comparison	Size	Data status
AI vs tamoxifen in premenopausal women				
2618/23 (99B)	ABCSG XII, Austria	Goserelin; (Anastrozole vs tamoxifen) ± Zoledronic Acid 3yr	1803	Received
7629/36 (03<)	TEXT/IBCSG 25-02, Switzerland	Triptorelin; (Exemestane vs tamoxifen) 5yr	2639	Received
7627 (03E)	SOFT/IBCSG 24-02, Switzerland	Triptorelin; (Exemestane vs tamoxifen) 5yr	3066	Received
21204 (04])	IT Naples HOBOE, Italy	Triptorelin; (Letrozole vs tamoxifen) 5yr	1065	Received

Appendix 2.

EBCTCG seventh cycle variables and data format

Either using the codes we suggest below or using your own codes, please extract from your dataset the variables that correspond most closely to the items listed below and send them to us. Please provide one record for each person ever randomised (including any person who was randomised and then was later categorised as ineligible, withdrawn, unevaluable, lost or "protocol deviant"—but, please tell us in question 9 which patients your group's preferred analyses would exclude, and why).

For trials where a dataset has previously been sent to the EBCTCG it is probably easiest and most reliable to update by re-sending all variables. If, however, this would cause difficulties then you can send only the additional variables; let us know if you want a file of the data you previously sent and we will provide it.

If any variable is not available or not applicable, please omit it and send only the remaining variables. If you have any of the requested variables in your records in a form that would require substantial additional work to supply (e.g. computerisation, or manual coding), feel free to omit them for now, but in your cover document please tell us of their existence. Please send the following:

- Your data in a separate Excel spreadsheet for each separate trial, if possible.
- A cover document giving all your coding conventions (including your format for dates).
- Send data to: bc.overview@ndph.ox.ac.uk with your research group's name (and/or the EBCTCG number for your research group) and your group's name for the trial in the subject line.

If you have any questions about this data request, please contact the EBCTCG secretariat on bc.overview@ndph.ox.ac.uk (Telephone: +44-1865-743852). All data supplied to the secretariat will be held securely and treated confidentially, in accordance with the EBCTCG Data Policy (available at: https://www.ctsu.ox.ac.uk/research/ebctcg).

For additional notes on supplying data to the overview, please see the EBCTCG Collaborators' Space web pages, or email the EBCTCG secretariat

CORE VARIABLES—BASELINE (Q1–27)

A) RANDOMISATION AND PATIENT CHARACTERISTICS (Q1–9)

1. **Your patient identifier** *(preferably specifying uniquely which trial as well as which patient)*
2. **Date of randomisation** *(specify your format for dates [in your covering document])*
3. **Allocated treatment** *(specify your codes)*
4. **Age** at randomisation *(years)* **NB Here & everywhere else, leaving an item Blank means Not Known**
5. **Height** at randomisation *(m)*
6. **Weight** at randomisation *(kg)*
7. **Menopausal status** at randomisation *(1=pre-, 2=peri-, 3=postmenopausal with intact ovaries & uterus, 4=ovarian ablation, 5=hysterectomy, 6=both [ie, 4 and 5]), 7=artificial, 8=male patient)*
8. **Did chemotherapy cause apparently permanent cessation of menses?** *(1=no/not applicable, 2=yes)*
9. **Would your group's preferred analyses exclude this patient?** NB A few trial patients may be randomised in error, otherwise ineligible, lost with no follow-up, unevaluable or withdraw consent.
   (1=no known reason for exclusion, 2=yes [specify main reason(s) for preferring exclusion, if known])

B) SURGICAL DETAILS (Q10–11; OR, DEFINE AND USE YOUR OWN CODES)

10. **Breast surgery** *(1=none, 2=only lumpectomy or wide local excision, 3=quadrantectomy or sector resection, 4=partial mastectomy, 5=simple or total mastectomy, 6=radical mastectomy, 7=modified radical mastectomy)*
11. **Axillary surgery** *(1=none, 2=sentinel node biopsy only, 3=axillary sampling, 4=surgical clearance of less than levels I & II, 5=full clearance of axillary levels I & II, 6=clearance of more than levels I & II, 7=axillary clearance, but levels cleared unspecified)*

C) PATHOLOGICAL NODAL STATUS (Q12–13; OR, USE YOUR OWN CODES [EG, TNM])

Note: In patients receiving neo-adjuvant treatment (or in trials where neo-adjuvant or axillary treatment differs between groups) give nodal status prior to neo-adjuvant (or axillary) therapy in section K

12. **Sentinel node biopsy** *(1=not done; 2=done and negative for cancer; 3=no greater involvement than isolated tumour cells [≤0.2 mm and/or < 200 cells]; 4=no greater involvement than micrometastases [>0.2mm or >200cells but ≤2mm]; 5=macroscopic nodal deposits [>2 mm]; 6=positive but size unknown)*
13. **Axillary status** *(specify codes, or: 1=pN- histologically; 2=N- other/unknown method; 3=1-3 positive nodes ; 4=4-9 [or 4+] positive; 5=10+ positive; 6=N+ histologically, unknown number; 7=N+ other/unknown method)*

D) TUMOUR CHARACTERISTICS (Q14–18; OR, USE YOUR OWN CODES [EG, TNM])

14. **Method first detected** *(1=mammographic screening, 2=incidental, 3=symptomatic, 4=other)*
15. **Laterality** *(1=left, 2=right, 3=bilateral)*
16. **Pathological grade prior to any neo-adjuvant therapy** *(1=well differentiated, 2=moderately,3=poorly)*
17. **Histological type** (if not locally determined please state)*(1=invasive, not otherwise specified, 2=ductal, 3=lobular, 4=other invasive, 5=mixed, 6=carcinoma in situ (CIS) only)*
18. **Tumour diameter:** largest diameter of excised primary *(mm)*

E) RECEPTOR STATUS (Q19–27; OR, USE YOUR OWN CODES)

Note: In trials with some neo-adjuvant treatment give receptor status prior to any neo-adjuvant therapy

19. **Summary of Estrogen Receptor (ER) status of primary tumour** *(1=ER-poor, 2=ER+, 3=ER++*
    [define in cover document, unless ER-poor is <10 fMol/mg and ER++ is ER definitely ≥100 fMol/mg])
20. **Quantitative ER measurement** *(measured in central/reference lab if possible, otherwise best available)*
21. **Units for ER** *(1=fMol/mg, 2=% +ve by IHC, 3=Allred score [category score], 4=H-score, 9=other [specify]*)

22. **Summary of Progesterone Receptor (PR) status of primary tumour** *(1=PR-poor, 2=PR+, 3=PR++*
    [define in cover document, unless PR-poor is <10 fMol/mg and PR++ is PR definitely ≥100 fMol/mg])
23. **Quantitative PR measurement** *(done in central/reference lab if possible, otherwise best available)*
24. **Units for PR** *(coded as Q21)*

25. **Summary of HER2 status of primary** *(1=negative/normal, 2=positive/over-expressing)*
26. **Quantitative HER2 measurement** *(done in central/reference lab if possible, otherwise best available)*
27. **Units for HER2** *(1=IHC [% staining], 2=IHC score [0, 1+, 2+, 3+], 3=FISH [# copies], 4=FISH [HER2:CEP17 ratio], 5=CISH [# copies], 6=CISH [HER2:CEP17], 9=other [please specify])*

F) NON-COMPLIANCE BEFORE ANY RECURRENCE (Q28–29; OR, USE YOUR OWN CODES)

28. **Any substantial deviation from trial treatment allocation (before any breast cancer recurrence)?**
(1=no, 2=never started, 3=discontinued, 4=switched to opposite trial group, 5=other [specify])

29. **Date of first such deviation from allocated treatment** *(ignore deviations after recurrence)*

G) CANCER RECURRENCE AND SECOND CANCERS (Q30–40; OR, USE YOUR OWN CODES)

30. **Any recurrence of invasive breast cancer (ie, locoregional, contralateral or distant)?**
NB Includes any occurrence of new ipsilateral or contralateral breast cancer *(1=no, 2=yes)*

31. **<u>If no:</u> Date patient last known to be free of such recurrence; <u>If yes:</u> Date of first such recurrence**

32. **Site of first distant recurrence (ie, possibly distant; not just locoregional/contralateral)**
(1=no distant recurrence, 2=recurrence-unknown if distant, 3=distant recurrence-unknown site(s), 4=only in distant soft tissue, 5=only in distant nodes, 6=only in bone, 7=only visceral, 8=only in CNS, 9=multiple sites including bone but not CNS/brain, 10=multiple sites not including bone or CNS/brain, 11=multiple sites including CNS/brain but not bone, 12=multiple sites including CNS/brain and bone)

33. **Date of first distant recurrence** NB Locoregional recurrence can precede first distant recurrence

34. **Site of first locoregional recurrence** *(1=no locoregional recurrence recorded, 2=multiple or unspecified locoregional sites 3=only in breast [new or recurrent invasive cancer] or chest wall, 4=only in axilla, 5=only in other locoregional nodes [eg, infraclavicular fossa], 6= multiple locoregional sites, 7=only in internal mammary nodes, 8=only in supraclavicular nodes, 9=tumour bed, 10=breast but known not tumour bed, 11=CIS only (if index CIS or if subsequent invasive cancer not collected))*

35. **Date of first locoregional recurrence**

36. **Contralateral breast cancer?** *(1=no, 2=yes: new invasive cancer thought to have arisen during follow-up in the contralateral breast, 3=CIS only (if index CIS))*

37. **Date of first contralateral breast cancer**
NB If patient had more than one second malignancy during follow-up, <u>repeat</u> variables 37-39 for each.

38. **Site of any second malignancy *[except breast cancer (including contralateral)]* during follow-up** *(Describe ALL sites. Use and specify your own codes; if you use ICD codes specify revision, eg ICD-9 or ICD-10)*

39. **Date of this second malignancy**

40. **MIGHT this have been a breast cancer metastasis?** *(1=no, 2=possibly/not yet certain [eg, possible lung, liver, bone or brain metastasis: please do not report definite breast metastases as second cancers])*

H) SURVIVAL (Q41–43)

41. **Is patient known to have died?** *(1=no, 2=yes)*

42. **<u>If NO:</u> Date patient last known to be alive; <u>If yes:</u> Date of death**

43. **<u>If YES:</u> Cause of death** *(use and specify your own codes; if you use ICD codes specify which version, eg ICD-9 or ICD-10)*

ADDITIONAL VARIABLES (Q44–59)

I) ADDITIONAL TUMOUR MARKER DATA (Q44–52; OR, USE YOUR OWN CODES)

Note: If tests of gene expression or special tests of IHC quantitation were done on the excised primary then please send a separate file in your own format with the fully detailed set of results on each individual.

44. **Summary of gene-expression status of primary tumour** *(1=low risk, 2=intermediate risk, 3=high risk): NB Please also provide the fully detailed gene expression results for each patient as a separate dataset.*

45. **Quantitative gene-expression prognostic score** *(best available single numerical measure)*

46. **Prognostic score used to quantify gene expression profile** *(use own code, or: 1=OncotypeDx prognostic score, 2=Mammaprint prognostic score, 3=EndoPredict, 4=Prosigna,9=other [please specify])*

47. **Summary of Topo-isomerase II alpha (TOPO2A) status of primary tumour** *(1= normal [ie, no gene over-expression or deletion], 2=positive/over-expressing, 3=deleted)*

48. **Quantitative TOPO2A measurement** *(done in central/reference laboratory if possible)*

49. **Units for TOPO2A** *(1=IHC [% staining], 2=IHC score [0, 1+, 2+, 3+], 3=FISH [number of copies], 4=FISH [TOPO:CEP17 ratio], 5=CISH [# copies], 6=CISH [TOPO:CEP17], 9=other [please specify])*

50. **Summary of Proliferation Index of primary tumour** *(1=low, 2=intermediate, 3=high)*

51. **Quantitative Proliferation Measure** *(best available numerical measure, in central/ ref lab if possible)*

52. **Factor measured for Proliferation Index** *(1=S-phase fraction [%], 2=thymidine labelling index [%], 3=Ki-67 by IHC [% staining], 9=other [please specify]*)

J) NON-FATAL ADVERSE EVENTS (Q53–54; OMIT IF NOT SOUGHT)

Note: Some treatments may cause or prevent bone fractures, cardiovascular events, lymphoedema, or lung fibrosis. Please describe all such events (eg, hip fracture, spinal fracture, myocardial infarction, stroke, pulmonary embolus, episode of cardiac failure) if, but only if, such events were sought and recorded systematically for all arms of the trial.

If more than one such event was recorded, repeat variables 52-53 for each.

53. **Nature and severity of event** *(use your own codes; if you use ICD codes, specify which version, eg ICD-9 or ICD-10, and if you use CTC Adverse Event codes, please specify version number, eg CTCAE-3 or CTCAE-4)*

54. **Date of event**

K) TRIALS WITH SOME NEO-ADJUVANT SYSTEMIC THERAPY OR WHERE AXILLARY TREATMENT DIFFERS BETWEEN GROUPS (Q55–59; OR, USE OWN CODES)

55. **Apparent axillary nodal status (clinical, radiological or other) before neo-adjuvant** *(1=N-, 2=N+)*

56. **Apparent tumour diameter (clinical or radiological) before neo-adjuvant***:* largest diameter *(mm)*

57. **Operability before any neo-adjuvant therapy** *(define your own codes, or: 1=Breast-conserving surgery feasible, 2=Mastectomy but not BCS feasible, 3=inoperable, 4=uncertain operability)*

58. **Breast tumour response after completion of neo-adjuvant** *(define your own codes, or: 1=clinically complete response [cCR] & negative pathology (for invasive disease and DCIS), 2=cCR with DCIS, 3=cCR with invasive cancer remaining pathologically, 4=cCR with no pathological information, 5=partial response, 6=stable disease, 7=progression [define 5–7])*

59. **Axillary response after neo-adjuvant** *(coded as Q58)*

ADDITIONAL VARIABLES FOR SPECIFIC META-ANALYSES (Q60–87)

For some meta-analyses we may need additional information on surgery, radiotherapy, adjuvant treatments received, or additional tumour markers but this section does not need to be completed for most meta-analyses so omit unless specifically requested.

L) TRIALS OF EXTENDED ENDOCRINE THERAPY (Q60–63)

Note: This information is needed only for trials of longer versus shorter endocrine therapy

60. **Endocrine therapy given prior to randomisation** *(1=Tamoxifen, 2=Aromatase Inhibitor, 3=Tamoxifen then Aromatase Inhibitor, 4=Aromatase Inhibitor then Tamoxifen, 5=Other (specify))*
61. **Date initial endocrine therapy started** *(Approximate date, or date of surgery if unknown)*
62. **Date of first switch from tamoxifen to AI, or vice versa** *(if applicable)*
63. **Date initial endocrine therapy completed** *(Approximate date)*

M) DETAILS OF ADDITIONAL TREATMENTS FOR TRIALS OF LOCAL, BIOLOGICAL, OR ENDOCRINE THERAPY (Q64–68)

64. **Neoadjuvant chemotherapy received** *(1=no, 2= non-anthracycline, non-taxane, 3=anthracycline, non-taxane, 4=taxane+anthracycline, 5=other taxane-containing, 9=Yes – type unknown)*
65. **Adjuvant chemotherapy received** *(as for Q64)*
66. **Endocrine therapy received** *(1=no, 2=tamoxifen, 3=aromatase inhibitor, 4=sequential tamoxifen and aromatase inhibitor, 5=ovarian ablation/suppression alone, 6=ovarian ablation+tamoxifen, 7=ovarian ablation + aromatase inhibitor, 9=Yes – type unknown))*
67. **HER2 directed therapy received** *(1=no, 2=Yes)*
68. **Radiotherapy received (irrespective of site)** *(1=No, 2=Yes)*

N) ADDITIONAL THERAPY DETAILS FOR TRIALS OF LOCAL THERAPY (Q69–79)

69. **Date of first breast surgery** *(dd/mm/yyyy)*
70. **Site of tumour in breast quadrant** *(1=lateral, 2=medial, 3=central, 4=medial or central, 5=not specified)*
71. **Lymphovascular invasion** *(1=no, 2=yes)*
72. **Date of first axillary surgery** *(dd/mm/yyyy)*
73. **Total number of sentinel lymph nodes excised and examined pathologically** *(-1=none, 1=one, 2=two, etc)*
74. **Number of sentinel lymph nodes excised and were:**
  a. **isolated tumour cells [≤0.2 mm and or ≤200 cells]** *(-1=none, 1=one, 2=two, etc)*
  b. **micrometastasis (>0.2mm or >200cells but ≤2mm)** *(-1=none, 1=one, 2=two, etc)*
  c. **macroscopic nodal deposits [>2 mm]** *(-1=none, 1=one, 2=two, etc)*
  d. **positive (unknown whether ITC, micro or macrometastasis)** *(-1=none, 1=one, 2=two, etc)*
  e. **negative** *(-1=none, 1=one, 2=two, etc)*

75. **Radiotherapy to whole breast/chest wall** *(1=No, 2=Yes)*
76. **Radiotherapy to partial breast** *(1=No, 2=Yes)*
77. **Radiotherapy to supraclavicular fossa** *(1=No, 2=Yes)*
78. **Radiotherapy to axilla** *(1=No, 2=Yes)*
79. **Radiotherapy to internal mammary chain** *(1=No, 2=Yes)*

O) ADDITIONAL DETAILS FOR TRIALS OF DCIS (Q80–84)

80. **Closest relevant excision margin** *(please give in mm where available; 99=no excision performed)*
81. **Diagnosis at entry** *(1=DCIS only, 2=DCIS+LCIS, 3 = DCIS (+/- LCIS) with microinvasion, 4=unknown)*
82. **Comedo** *(1=Present, marked or severe; 2=Present, moderate; 3=Present, slight; 4=Present, NOS; 5=Absent; 6=Unknown)*
83. **Architecture** *(1=Cribriform; 2=Micropapillary; 3=Papillary; 4=Solid; 5=Other (please specify), 6=Unknown)*
84. **Focality** *(1=Unifocal;2=Multifocal/multicentric;3=Unknown)*

P) ADDITIONAL QUESTIONS FOR TRIALS OF DURATION OF BIOLOGICAL THERAPY *(Q85–88)*

Note: This information is needed only for trials of longer versus shorter biological therapy (e.g. trastuzumab duration)

85. **Biological therapy given prior to randomisation** *(1=None, 2=Yes)*
86. **Date initial biological therapy started** *(Approximate date, or date of surgery if unknown)*
87. **Date initial biological therapy completed** *(Approximate date)*

Q) FURTHER TUMOUR MARKERS *(Q88–91)*

88. **Summary of Tumour Infiltrating Lymphocytes (TILS) status** *(1=low, 2=intermediate, 3=high)*
89. **Quantitative TILS Measure** *(% staining positive, in central/ ref lab if possible)*
90. **Central histological type** *(coded as for Q17)*
91. **E-Cadherin data** *(1=negative, 2=positive,3=intermediate/unclear)*