

## Appendix Four. Clinical effectiveness

### Contents

1. Treatment regimens and available trial data.....	1
Treatment regimens in randomised controlled trials.....	1
Trial outcomes as reported.....	10
2. Increasing uncertainty around sequence and durations .....	13
Non-publication of data for 12 month sequential treatment from N9831 .....	13
Initial HERA results overstate effects on disease free survival .....	16
Impact of N9831 Arm B and the updated HERA results on disease free survival benefits of 12 month sequential treatment .....	17
The impact of sequencing on the effectiveness of 12 month treatment .....	19
The impact of sequencing on the cardiotoxicity of 12 month treatment.....	21
Durability of response.....	21
3. Concurrent treatment pre-anthracyclines (FinHer and E2198).....	26
FinHer .....	26
Reliability and validity of the FinHer trial and nine-week concurrent treatment.....	27
Patient numbers.....	28
Overall survival.....	30
Data from other clinical trials (E2198) .....	31
4. The SOLD study.....	33
References and endnotes .....	36

### 1. Treatment regimens and available trial data

Clinical effectiveness is one of PHARMAC’s decision criteria (<http://www.pharmac.govt.nz/pdf/231205.pdf> ‘The clinical benefits and risks of pharmaceuticals’).

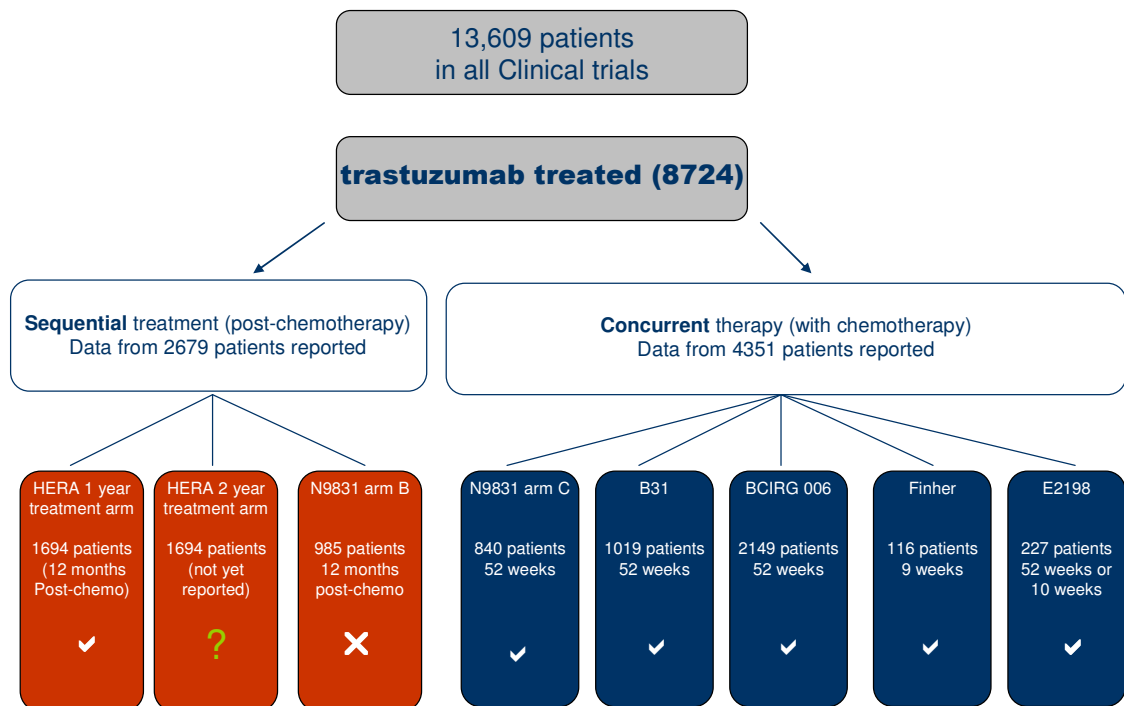
#### *Treatment regimens in randomised controlled trials*

To date there have been five open-label randomised controlled trials (RCTs) reporting outcomes for adjuvant trastuzumab given in addition to standard chemotherapy treatment against standard chemotherapy treatment alone in HER2-positive early breast cancer<sup>1</sup>—HERA<sup>2 3 4</sup>, NASBP B31<sup>5 6</sup>, NCCTG N9831<sup>5 7</sup>, BCIRG 006<sup>8</sup>, and FinHer.<sup>9</sup> (A sixth study, ECOG E2198<sup>10</sup>, which compared 12 months with 10 weeks trastuzumab given concurrently with paclitaxel, was presented as a poster at the San Antonio Breast Cancer Symposium in 2006; however this study has not reported outcomes against standard chemotherapy treatment.)

In these studies, trastuzumab was investigated in one of two broad treatment regimens (Figure 1):

1. ‘sequential treatment’—trastuzumab for 12 months following completion of chemotherapy (anthracycline +/- taxane): HERA and NCCTG N9831 Arm B;
2. ‘concurrent treatment’—trastuzumab for 9-10 weeks, or 12 months, started in combination with taxane chemotherapy (either preceding or following completion of anthracycline chemotherapy): NASBP B31, NCCTG N9831 Arm C, BCIRG 006, and FinHer.

Figure 1. Clinical trials of trastuzumab



**Key to graphic:**

✓ statistically significant improvement in disease-free survival (DFS)

✗ no statistically significant improvement in DFS

? results awaited

**Notes to graphic:**

HERA patient numbers derive from the first (12 month median f/u) interim analysis (Piccart-Gebhart 2005), where n=1694 patients in the 1 year trastuzumab arm and n=1693 in the standard treatment arm; this compares with n=1703 and n=1698 for those respective arms reported in the second (23 month median f/u) interim analysis (Smith 2007).

BCIRG 006 includes both Arm AC-TH (post anthracycline treatment, n=1074) and Arm TCH (concurrent trastuzumab with docetaxel and carboplatin, no anthracycline, n=1075). Includes study ECOG E2198.

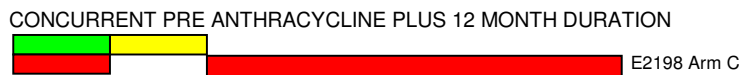
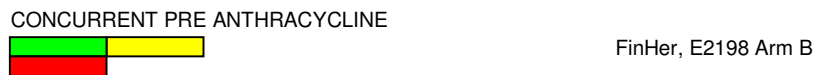
Trastuzumab treatment regimens in these studies varied in their association with taxanes (sequential to, concurrent with), duration of trastuzumab (between 9 weeks and 2 years), sequence relative to anthracycline chemotherapy and weekly or every three-weekly dosing (Figure 2).

Figure 2. Anthracycline-containing treatment regimens in trials of trastuzumab in early stage HER2-positive breast cancer

## 1) Sequential



## 2) Concurrent



Key:

- taxane chemotherapy
- anthracycline-containing chemotherapy
- trastuzumab

Broad schematic only. Other regimens, not illustrated, include concurrent use of trastuzumab with docetaxel and carboplatin chemotherapies in BCIRG 006 (no anthracycline chemotherapy), and 2 years' sequential trastuzumab (following anthracycline and then taxane chemotherapies) in HERA (no data available to date). HERA permitted the use of a variety of different combinations of chemotherapy, including non-anthracycline regimens using CMF (<http://content.nejm.org/cgi/data/353/16/1659/DC1/1>).

Specific treatment regimens studied in the trials were as follows:

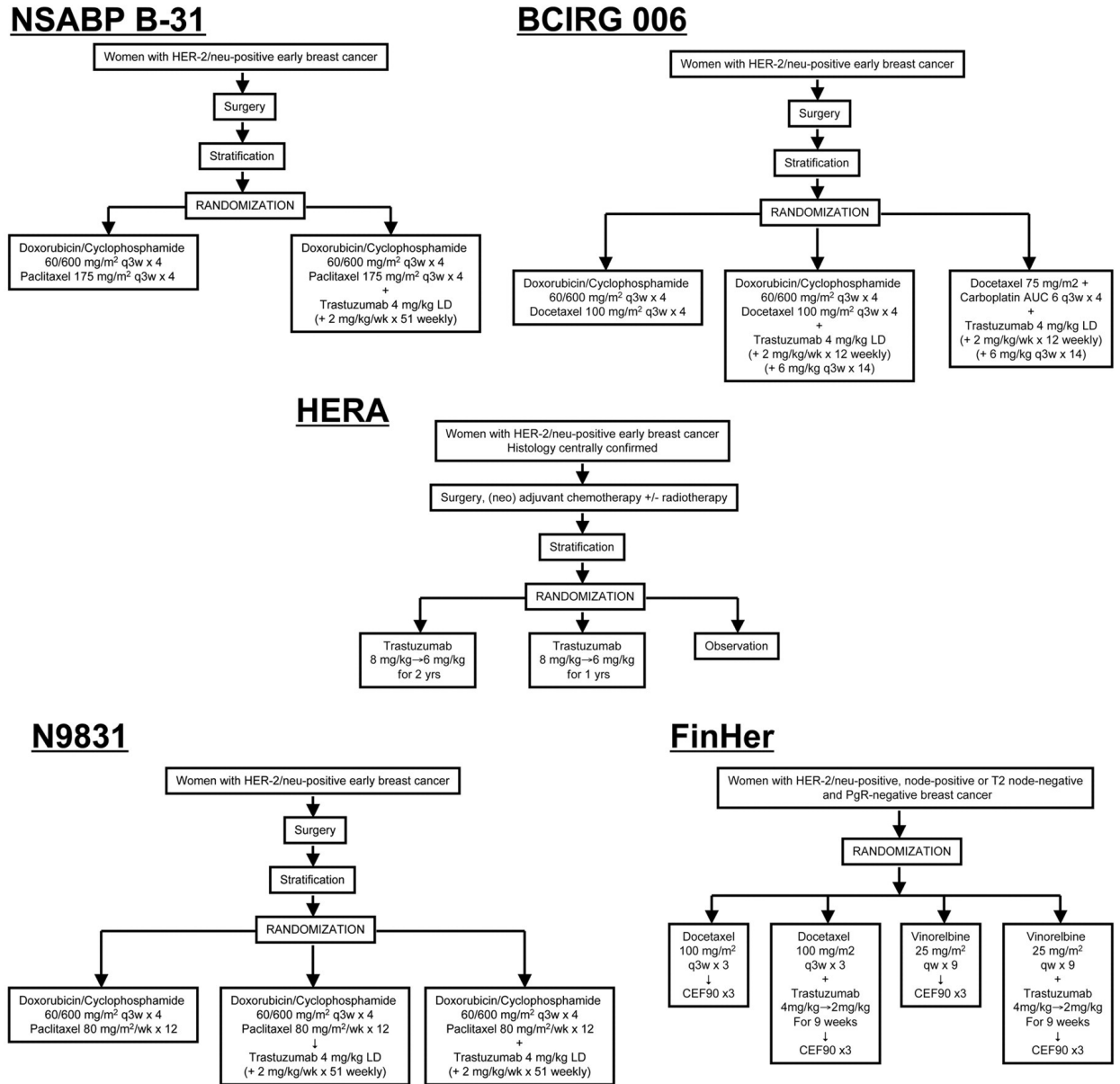
- HERA—a three-arm open-label RCT comparing 12 months and 24 months sequential with standard chemotherapy:
  - **control (n=1693-1698<sup>11</sup>):** *Observation alone following completion of standard neoadjuvant or adjuvant chemotherapy*
  - **1 Year arm (n=1694-1703<sup>11</sup>):** *12 months sequential trastuzumab treatment following completion of standard neoadjuvant or adjuvant chemotherapy*
  - **2 Year arm (n=1694-1701<sup>11</sup>):** *24 months sequential trastuzumab treatment following completion of standard neoadjuvant or adjuvant chemotherapy*
- NCCTG N9831—a three-arm open-label RCT comparing sequential and concurrent trastuzumab treatment with standard chemotherapy:
  - **Arm A (control) (n= 979):** *4 cycles of AC treatment followed by 12 weeks paclitaxel*
  - **Arm B (12 months sequential trastuzumab treatment) (n=985):** *4 cycles of AC treatment followed by 12 weeks paclitaxel and then 52 weeks trastuzumab*
  - **Arm C (12 months concurrent trastuzumab treatment) (n=840):** *4 cycles AC followed by 12 weeks paclitaxel; trastuzumab started on day 1 of paclitaxel and continued for 12 months.*
- NASBP B31—a two-arm open-label RCT comparing concurrent trastuzumab treatment with standard chemotherapy:
  - **Control (n=1024):** *4 cycles of AC treatment followed by 12 weeks paclitaxel*
  - **12 months concurrent trastuzumab treatment (n=1019):** *4 cycles AC followed by 12 weeks paclitaxel; trastuzumab started on day 1 of paclitaxel and continued for 12 months*
- BCIRG 006—a three-arm open-label RCT comparing concurrent trastuzumab treatment (with 2 different chemotherapy regimens) with standard chemotherapy:
  - **Arm AC-T (control) (n=1073):** *4 cycles of AC treatment followed by 4 cycles docetaxel*
  - **Arm AC-TH (12 months concurrent trastuzumab) (n=1074):** *4 cycles of AC treatment followed by 4 cycles docetaxel; trastuzumab started on day 1 of docetaxel and continued for 12 months*
  - **Arm TCH (12 months concurrent trastuzumab) (n=1075):** *6 cycles docetaxel and carboplatin; trastuzumab started on day 1 of docetaxel/carboplatin and continued for 12 months*

- FinHer—an open-label RCT comparing docetaxel with vinorelbine for the adjuvant treatment of early stage breast cancer (n=1010). For the subset of women with HER2-positive cancers, patients were further randomised to four arms comparing concurrent trastuzumab (with 2 different chemotherapy regimens—docetaxel or vinorelbine) with docetaxel or vinorelbine chemotherapy alone:
  - **docetaxel + FEC, no trastuzumab (n=58):** 3 cycles of docetaxel followed by 3 cycles of AC
  - **docetaxel + FEC, trastuzumab (n=54):** 3 cycles of docetaxel followed by 3 cycles of AC; trastuzumab started on day 1 of docetaxel and continued for 9 weeks
  - **vinorelbine + FEC, no trastuzumab (n=58):** 3 cycles of vinorelbine followed by 3 cycles of AC
  - **vinorelbine + FEC, trastuzumab (n=62):** 3 cycles of vinorelbine followed by 3 cycles of AC; trastuzumab started on day 1 of vinorelbine and continued for 9 weeks

Nearly all regimens have involved anthracycline chemotherapy, apart from BCIRG 006 which also included an arm that assessed concurrent trastuzumab with docetaxel and carboplatin (not anthracycline—Arm ‘TCH’, 1075 patients), and in HERA (which permitted a variety of chemotherapy regimens) where 6% of primary treatments contained no anthracyclines. In the context of the rest of this appendix, discussions about trastuzumab regimens are restricted to those containing anthracycline, which comprise most of the available evidence and are currently most widely used and applicable to New Zealand as standard care chemotherapy. Results from BCIRG 006 discussed in this appendix are therefore largely restricted to the anthracycline-containing arms AC-T and AC-TH (control and added trastuzumab). Study designs and their treatment regimens are summarised further in Figure 3 and at the end of this Appendix.

Figure 3. Schemas of published trials using adjuvant trastuzumab

Abbreviations: CEF, Cyclophosphamide, epirubicin, 5FU; LD, loading dose; PgR, progesterone receptor; qw, weekly; q3w, every 3 weeks



Source: Gonzalez-Angulo AM, Hortobagyi GN, Esteva FJ. Adjuvant therapy with trastuzumab for HER-2/neu-positive breast cancer. *Oncologist* 2006 Sep;11(8):857-67.

The results from these trials is summarised on the bpac<sup>nz</sup> website at <http://www.bpac.org.nz/magazine/2007/april/herceptin.asp> and Adjuvant! online (<http://www.adjuvantonline.com/breasthelp0306/Trastuzumab.html>). The study designs and the efficacy and cardiotoxicity results are summarised further in Tables 1 to 3 and Figures 4 to 5B below.

Table 1. Disease recurrence in sequential and concurrent trastuzumab regimens in RCTs of trastuzumab vs. standard care in the adjuvant treatment for HER2-positive early breast cancer

Trial	Sequential (trastuzumab post taxane) [n tmt/N tmt vs. n obs/N obs], hazard ratio (95% CI), % absolute improvement <sup>1</sup> (minimal from 95% CI <sup>1</sup> )	Concurrent (trastuzumab with taxane) [n tmt/N tmt vs. n obs/N obs], hazard ratio (95% CI), % absolute improvement <sup>1</sup> (minimal from 95% CI <sup>1</sup> )
(HERA one-year treatment arm, 12 mth f/up–interim result)	[127/1694 vs. 220/1693], 0.54 (0.43-0.67), 5.5%	
HERA one-year treatment arm, 23 mth median f/up	[218/1703 vs. 321/1698], 0.64 (0.54-0.76), 6.1% (3.9%)	
NSABP B31		[83/864 vs. 171/872], 0.45 (CI not reported; 2P=10 <sup>-9</sup> ), 10.0% (7.3%)
NCCTG N9831	[103/985 vs. 117/979], 0.87 (0.67-1.13), 1.5% (-1.5%)	[50/808 vs. 90/807], 0.55 (CI not reported; 2P=0.0004), 5.0% (2.5%)
joint analysis of N9831/B31		[134/1672 vs. 261/1679], 0.48 (0.39-0.59), 7.5% (5.8%)
BCIRG 006 Arm AC-TH (36 mth year f/u)		[128/1074 vs. 192/1073], 0.61 (0.48-0.76), 6.0% (3.2%)
<b>Overall post-anthracycline treatment</b>	[321/2688 vs. 438/2677], <b>0.70 (0.61-0.81), 4.4%<sup>2</sup> (2.7%)</b>	[262/2291 vs. 453/2235], <b>0.53 (0.46-0.62), 6.9%<sup>2</sup> (5.5%)</b>
FinHer (pre-anthracycline treatment) 36 mth f/u–interim		[12/115 vs. 27/116], 0.42 (0.21-0.83), 12.8% (3.7%)
non-anthracycline, concurrent (BCIRG 006 Arm TCH), 36 mth		[142/1075 vs. 192/1073], 0.67 (0.54-0.83), 4.7% (1.8%)

Notes:

1 Absolute improvements in DFS are the differences between rates of disease recurrence in the observation and treatment arms. Minimum absolute improvements derive from upper limits of 95% confidence intervals for calculated DFS relative risks.

2 Overall results for sequential treatment comprise the N9831 sequential (B) and HERA one year trastuzumab arms compared with respective control groups (HERA 2-year median follow-up results).

3 Overall results for concurrent treatment comprise N9831 concurrent (C), NSABP B31 trastuzumab and BCIRG 006 AC-TH arms compared with respective control groups. These results are post-anthracyclines, and hence do not include concurrent treatment given in FinHer (pre-anthracyclines), nor the non-anthracycline (carboplatin + docetaxel + trastuzumab) Arm TCH of BCIRG 006.

Table 2. Trials reporting outcomes for adjuvant trastuzumab against vs. standard treatment in HER2-positive early breast cancer

	Sequential treatment trials, long duration (12 month) <sup>i</sup> regimens		Concurrent treatment trials, long duration (12 month) regimens			Concurrent treatment trials with short duration regimens	
	HERA Trial	N9831 (Arm B)	N9831 (Arm C)	BCIRG 006 <sup>ii</sup>	B31	FinHer Trial	E2198 <sup>iii</sup>
<b>Patient Numbers</b>	Observation: 1,693 Trastuzumab (1 yr): 1,694 Trastuzumab (2 yr): <sup>i</sup> 1,694	Observation: 979  Trastuzumab: 985	Observation: 979  Trastuzumab: 840	Observation: 1,073  Trastuzumab: 1,074	Observation: 1,024  Trastuzumab: 1,019	Observation: (116) vinorelbine – 58 docetaxel – 58  Trastuzumab: (116) Docetaxel – 54 Vinorelbine – 62	Short duration: 115 patients  Long duration: 112 patients
<b>Intervention</b>	1 loading dose (8mg/kg) trastuzumab, then 6mg/kg every 3 weeks for one year or two years (17 or 35 infusions, respectively).	1 loading dose (4mg/kg) trastuzumab, then 2mg/kg every week for 52 weeks	1 loading dose (4mg/kg) trastuzumab, then 2mg/kg every week for 52 weeks	1 loading dose (4mg/kg) trastuzumab, then 2mg/kg every week for 52 weeks	1 loading dose (4mg/kg) trastuzumab, then 2mg/kg every week for 52 weeks	9 Trastuzumab infusions at 1 week intervals. First dose 4mg/kg (90min infusion), remaining doses 2mg/kg (30 min infusion)	Short duration: 10 Trastuzumab infusions at 1 week intervals. Trastuzumab given in combination with paclitaxel– Loading dose 4mg/kg followed by 9 weeks 2mg/kg  Long duration: As above but with further 52 weeks of Trastuzumab at 2mg/kg per week.
<b>Timing of treatment</b>	Sequential (after completion of all chemotherapy – anthracycline chemotherapy <sup>iv</sup> and taxane treatment <sup>v</sup> )		Concurrent with taxane (paclitaxel), after completion of anthracycline chemotherapy			Concurrent with taxane (docetaxel) treatment, before anthracycline chemotherapy	Short duration: Concurrent with taxane treatment, before AC anthracycline chemotherapy <sup>2</sup>  Long duration: concurrent and sequential
<b>Disease free survival Hazard Ratio (95% confidence interval)</b>	12-mth median f/up: 0.54 (0.43-0.67)  23-mth median f/up: 0.64 (0.54-0.76)	0.87 (0.67-1.13)	0.55 (CI not reported <sup>vi</sup> )	23-mth median f/up: 0.49 (0.37-0.79)  36-mth median f/up: 0.61 (0.48-0.76)	0.45 (CI not reported <sup>vi</sup> )	36-mth median f/up: 0.42 (0.21-0.83)	DFS at 5 yrs: 76% short duration 75% long duration  Note that E2198 is a pilot study, not designed to report efficacy <sup>iii</sup>
<b>Overall DFS HR</b>	0.70 (0.61-0.81) (2-yr HERA f/u)		0.53 (0.46-0.62) (3-yr BCIRG 006 f/u)			0.42 (0.21-0.83)	N/A – placebo group data not reported
<b>Overall survival (95% CI)</b>	12-mth median f/up: 0.76 (0.47-1.23)  23-mth median f/up: 0.66 (0.47-0.91)	0.85 (0.55-1.33)	Not reported  (joint analysis with B31 = 0.67 (0.48-0.93))	36-mth median f/up: 0.59 (0.42-0.85)	Not reported  (joint analysis with N9831 arm C = 0.67 (0.48-0.93))	36-mth median f/up: 0.41 (0.16-1.08)	OS at 5 yrs: 89% short duration 83% long duration  Note that E2198 is a pilot study, not designed to report efficacy <sup>iii</sup>
<b>Overall OS HR (95% CI)</b>	0.72 (0.55-0.94)		0.63 (0.50-0.80)			0.41 (0.16-1.08)	N/A – placebo group data not reported

<sup>i</sup> No evidence is available on the outcomes of the 2 year trastuzumab treatment arm in the HERA trial.

<sup>ii</sup> Note that there was also an arm to BCIRG 006 ('arm TCH') that consisted of 6 cycles of docetaxel and carboplatin with concurrent trastuzumab (i.e. no anthracycline chemotherapy). However, because this regimen is not comparable to the other regimens, these results are not presented in this table. For further information regarding BCIRG 006 see Appendix One: Minutes of the relevant clinical advisory committee meetings. Anthracycline-containing regimens comprise most of the available evidence and are currently most widely used and applicable to New Zealand as standard care chemotherapy.

<sup>iii</sup> The E2198 study (Sledge et al, poster presentation at SABCs 2006) was not designed to test efficacy, and was not powered to determine equivalence, and results comparing the treatment arms to the control arm have yet to be reported. However, the results supported the efficacy of short duration concurrent trastuzumab therapy when administered before anthracycline containing chemotherapy, as demonstrated in the FinHer study (Appendix One: Minutes of the relevant clinical advisory committee meetings).

<sup>iv</sup> Anthracycline containing chemotherapy regimens (FEC or FAC).

<sup>v</sup> The HERA trial allowed several different chemotherapy regimens.

<sup>vi</sup> Note that N9831 Arm C and trial B31 data were only published as a joint analysis (Romond, 2005) without stating the hazard ratios' confidence intervals for the individual trials. Confidence limits for the disease recurrence HR for N9831 Arm B were stated in the 2005 conference presentation on the ASCO website.

Table 3. Event rates, absolute improvements and numbers needed to treat (NNTs) in adjuvant trastuzumab trials in HER2-positive early breast cancer, for disease free survival (DFS) and overall survival (OS)

Disease free survival																								
	median f/u (years)	totals		treatment group			observation group			hazard ratio			relative risk			ARR/improvement in DFS				disease recurrence free survival (DFS)			number needed to treat (NNT)	
		n	N pts	n	N pts	rate	n	N pts	rate	HR	-95% CL	+95% CL	RR	-95% CL	+95% CL	from HR	from RR	-95% CL, from RR	+95% CL, from RR	obs	tmt (from RR)	-95% CL, from RR	from RR	-95% CL, from RR
sequential post-anthracyclines																								
HERA 23-mth FU data	1.9	539	3,401	218	1703	13%	321	1698	19%	<b>0.64</b>	0.54	0.76	<b>0.68</b>	0.58	0.79	6.8%	6.1%	3.9%	8.0%	81.1%	87.2%	85.0%	16	26
N9831 sequential arm	1.5	220	1,964	103	985	10%	117	979	12%	<b>0.87</b>	0.67	1.13	<b>0.87</b>	0.68	1.12	1.6%	1.5%	-1.5%	3.8%	88.0%	89.5%	86.6%	67	-68
overall HR (95% CI)	1.8	759	5,365	321	2688	12%	438	2677	16%	<b>0.70</b>	0.61	0.81	<b>0.73</b>	0.64	0.83	4.9%	4.4%	2.7%	5.9%	83.6%	88.1%	86.3%	23	37
concurrent post-anthracyclines																								
B31 & N9831 arm C	2.0	395	3,351	134	1672	8%	261	1679	16%	<b>0.48</b>	0.39	0.59	<b>0.52</b>	0.42	0.63	8.1%	7.5%	5.8%	9.0%	84.5%	92.0%	90.2%	13	17
BCIRG 006 arm AC-TH	3.0	320	2,147	128	1074	12%	192	1073	18%	<b>0.61</b>	0.48	0.76	<b>0.67</b>	0.54	0.82	7.0%	6.0%	3.2%	8.2%	82.1%	88.1%	85.3%	17	31
overall HR (95% CI)	2.4	715	5,498	262	2746	10%	453	2752	16%	<b>0.53</b>	0.46	0.62	<b>0.58</b>	0.50	0.67	7.7%	6.9%	5.5%	8.2%	83.5%	90.5%	89.0%	14	18
concurrent pre-anthracyclines																								
FinHer	3.0	39	231	12	115	10%	27	116	23%	<b>0.42</b>	0.21	0.83	<b>0.45</b>	0.24	0.84	13.5%	12.8%	3.7%	17.7%	76.7%	89.6%	80.4%	8	27
overall HR (95% CI)	3.0	39	231	12	115	10%	27	116	23%	<b>0.42</b>	0.21	0.83	<b>0.45</b>	0.24	0.84	13.5%	12.8%	3.7%	17.7%	76.7%	89.6%	80.4%	8	27
concurrent non anthracycline																								
BCIRG 006 arm TCH	3.0	334	2,148	142	1075	13%	192	1073	18%	<b>0.67</b>	0.54	0.83	<b>0.74</b>	0.60	0.90	5.9%	4.7%	1.8%	7.1%	82.1%	86.8%	83.9%	21	57
overall HR (95% CI)	3.0	334	2,148	142	1075	13%	192	1073	18%	<b>0.67</b>	0.54	0.83	<b>0.74</b>	0.60	0.90	5.9%	4.7%	1.8%	7.1%	82.1%	86.8%	83.9%	21	57
all regimens																								
overall HR (95% CI)	2.2	1,847	13,242	737	6624	11%	1110	6618	17%	<b>0.62</b>	0.57	0.68	<b>0.66</b>	0.61	0.72	6.3%	5.6%	4.6%	6.6%	83.2%	88.9%	87.9%	18	22
Overall survival																								
	median f/u (years)	totals		treatment group			observation group			hazard ratio			relative risk			ARR/improvement in OS				overall survival (OS)			number needed to treat (NNT)	
		n	N pts	n	N pts	rate	n	N pts	rate	HR	-95% CL	+95% CL	RR	-95% CL	+95% CL	from HR	from RR	-95% CL, from RR	+95% CL, from RR	obs	tmt (from RR)	-95% CL, from RR	from RR	-95% CL, from RR
sequential post-anthracyclines																								
HERA 23-mth FU data	1.9	149	3,401	59	1703	3%	90	1698	5%	<b>0.66</b>	0.47	0.91	<b>0.65</b>	0.47	0.90	1.8%	1.8%	0.5%	2.8%	94.7%	96.5%	95.2%	54	191
N9831 sequential arm	1.5	44	1,964	1	985	0%	43	979	4%	<b>0.85</b>	0.55	1.33	<b>0.02</b>	0.00	0.17	0.7%	4.3%	3.7%	4.4%	95.6%	99.9%	99.3%	23	27
overall HR (95% CI)	1.9	193	5,365	60	2688	2%	133	2677	5%	<b>0.72</b>	0.55	0.94	<b>0.45</b>	0.33	0.61	1.4%	2.7%	2.0%	3.3%	95.0%	97.8%	97.0%	37	51
concurrent post-anthracyclines																								
B31 & N9831 arm C	2.0	154	3,351	62	1672	4%	92	1679	5%	<b>0.67</b>	0.48	0.93	<b>0.68</b>	0.49	0.93	1.8%	1.8%	0.4%	2.8%	94.5%	96.3%	94.9%	56	251
BCIRG 006 arm AC-TH	3.0	129	2,147	49	1074	5%	80	1073	7%	<b>0.59</b>	0.42	0.85	<b>0.61</b>	0.43	0.86	3.1%	2.9%	1.0%	4.2%	92.5%	95.4%	93.6%	35	99
overall HR (95% CI)	2.5	283	5,498	111	2746	4%	172	2752	6%	<b>0.63</b>	0.50	0.80	<b>0.65</b>	0.51	0.82	2.3%	2.2%	1.1%	3.0%	93.8%	96.0%	94.9%	45	87
concurrent pre-anthracyclines																								
FinHer	3.0	20	231	6	115	5%	14	116	12%	<b>0.41</b>	0.16	1.08	<b>0.43</b>	0.17	1.09	7.1%	6.9%	-1.0%	10.0%	87.9%	94.8%	86.9%	15	-97
overall HR (95% CI)	3.0	20	231	6	115	5%	14	116	12%	<b>0.41</b>	0.16	1.08	<b>0.43</b>	0.17	1.09	7.1%	6.9%	-1.0%	10.0%	87.9%	94.8%	86.9%	15	-97
all concurrent non anthracycline																								
BCIRG 006 arm TCH	3.0	136	2,148	56	1075	5%	80	1073	7%	<b>0.66</b>	0.47	0.93	<b>0.70</b>	0.50	0.97	2.5%	2.2%	0.2%	3.7%	92.5%	94.8%	92.7%	45	491
overall HR (95% CI)	3.0	136	2,148	56	1075	5%	80	1073	7%	<b>0.66</b>	0.47	0.93	<b>0.70</b>	0.50	0.97	2.5%	2.2%	0.2%	3.7%	92.5%	94.8%	92.7%	45	491
all regimens																								
overall HR (95% CI)	2.6	632	13,242	233	6624	4%	399	6618	6%	<b>0.66</b>	0.56	0.77	<b>0.58</b>	0.50	0.68	2.0%	2.5%	1.9%	3.0%	94.0%	96.5%	95.9%	40	52

Figure 4. Hazard ratios for disease recurrence in sequential and concurrent trastuzumab regimens in RCTs of trastuzumab vs. standard care in the adjuvant treatment for HER2-positive early breast cancer

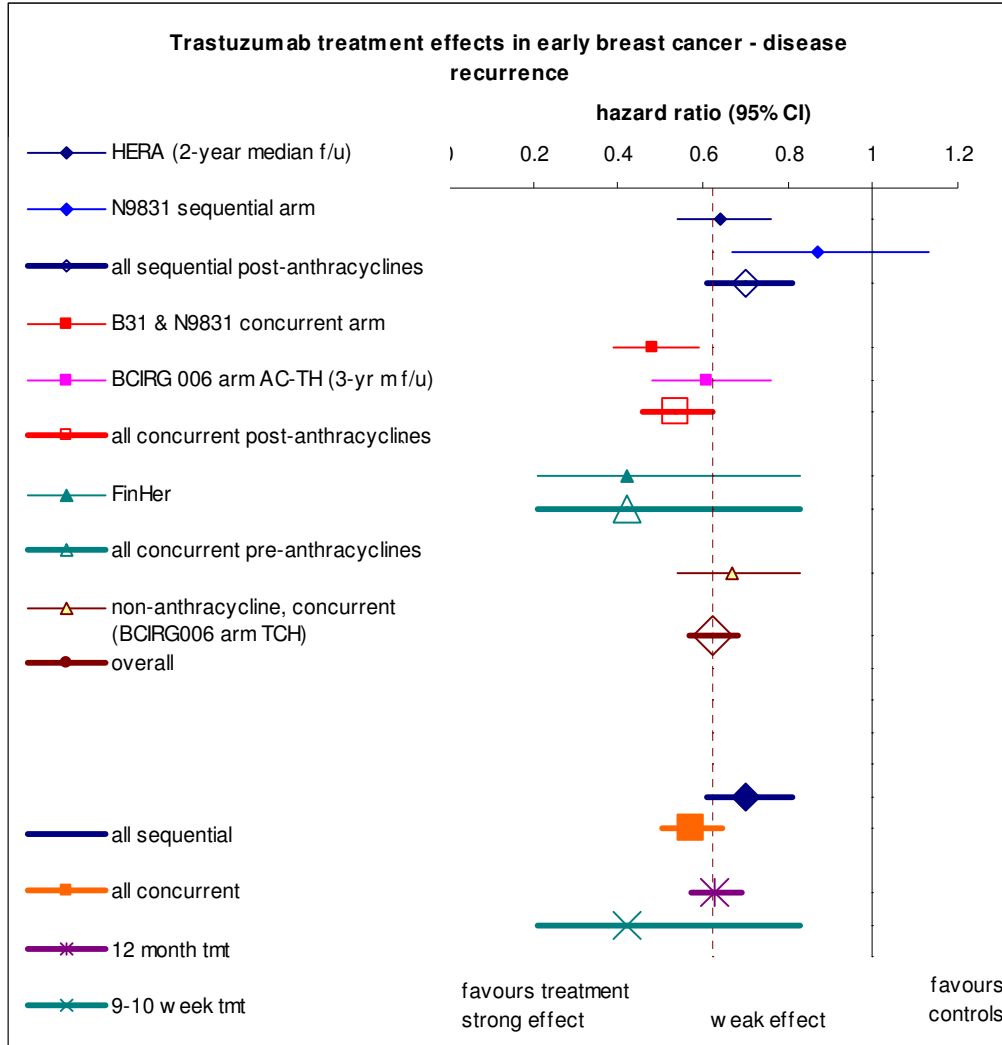
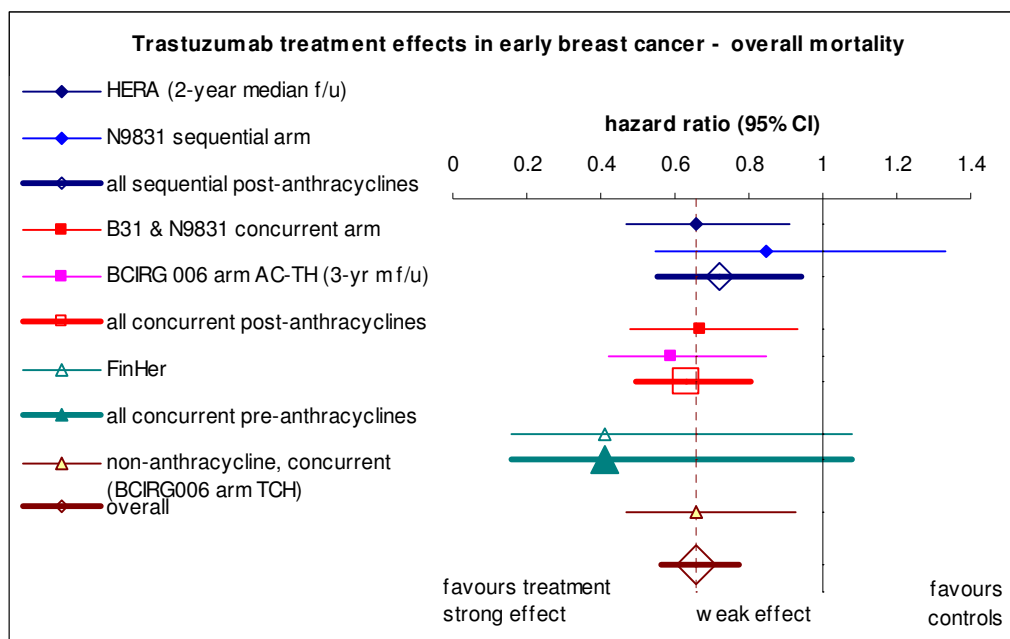


Figure 4A. Hazard ratios for overall mortality in sequential and concurrent trastuzumab regimens in RCTs of trastuzumab vs. standard care in the adjuvant treatment for HER2-positive early breast cancer



### ***Trial outcomes as reported***

Tables 1 to 3 and figures 4 to 5B (below) indicate that overall trastuzumab, when given in addition to standard chemotherapy, reduces the risk of both disease recurrence and overall deaths by  $\sim 1/3^{\text{rd}}$  relative to standard chemotherapy along, with overall (pooled) hazard ratios (HRs) for the five RCTs of 0.62 (0.57-0.68) for disease recurrence and 0.66 (0.56-0.77) for overall mortality. This translates to overall absolute improvements in DFS of 5.6% (on top of 83% untreated free of disease recurrence/death over 2.2 years average median follow-up (f/u), number needed to treat (NNT) 18); and a 2.5% absolute improvement in overall survival (on top of 94% untreated survival by 2.2 years, NNT 40). All RCTs show statistically significant DFS gains except the sequential arm (Arm B) of N9831, and significant overall survival gains are demonstrated in the N9831 Arm C/B31, BCIRG 006 and HERA trials.

The tables and figures also show the breadth of disparities in efficacy between the studies, i.e. differences in HRs between studies and regimens and differences in the ranges and breadths of confidence limits for those HRs:

- For trials, hazard ratios for DFS vary between 0.42 (FinHer) to 0.87 (sequential arm of trial N9831), and confidence limits (precision) vary threefold in range (HERA 95% CI 0.54-0.76 vs. FinHer 0.21-0.83). These translate to ranges in DFS absolute improvements of 1.5% to 12.8% (NNT 8 by median 3 years for FinHer, 67 by median 1.5 years for N9831 sequential arm). These also translate to a range of minimum absolute improvements (from the upper 95% confidence limits of calculated relative risks (RR) applied to untreated disease recurrence rates) of -1.5% for the N9831 sequential arm (NNT -68 i.e. would harm every 68<sup>th</sup> patient by a median of 1.5 years) to 5.8% (B31 and N9831 concurrent arm, NNT 17 by median 2 years).
- For broad treatment regimens, overall (pooled) HRs for disease recurrence vary between 0.42 (0.21-0.83) for concurrent treatment pre-anthracycline, 0.53 (0.46-0.62) for concurrent treatment post-anthracycline and 0.70 (0.61-0.81) for sequential treatment post-anthracycline.

These values translate to absolute DFS improvements of 12.8%, 6.9% and 4.4% for concurrent pre-anthracycline (AC), concurrent post-AC and sequential post-AC respectively (NNTs 8 by median 3.0 years, 14 by 2.4 years and 23 by 1.8 years respectively). Minimum confident absolute DFS improvements for the broad regimens (from the RR upper 95% confidence limits) vary between 3.7% for concurrent treatment pre-anthracycline (NNT 27), 5.5% (18) for concurrent treatment post-anthracycline and 2.7% (37) for sequential treatment post-anthracycline.

Figure 5. DFS in adjuvant trastuzumab trials in HER2-positive early breast cancer for untreated and treated groups

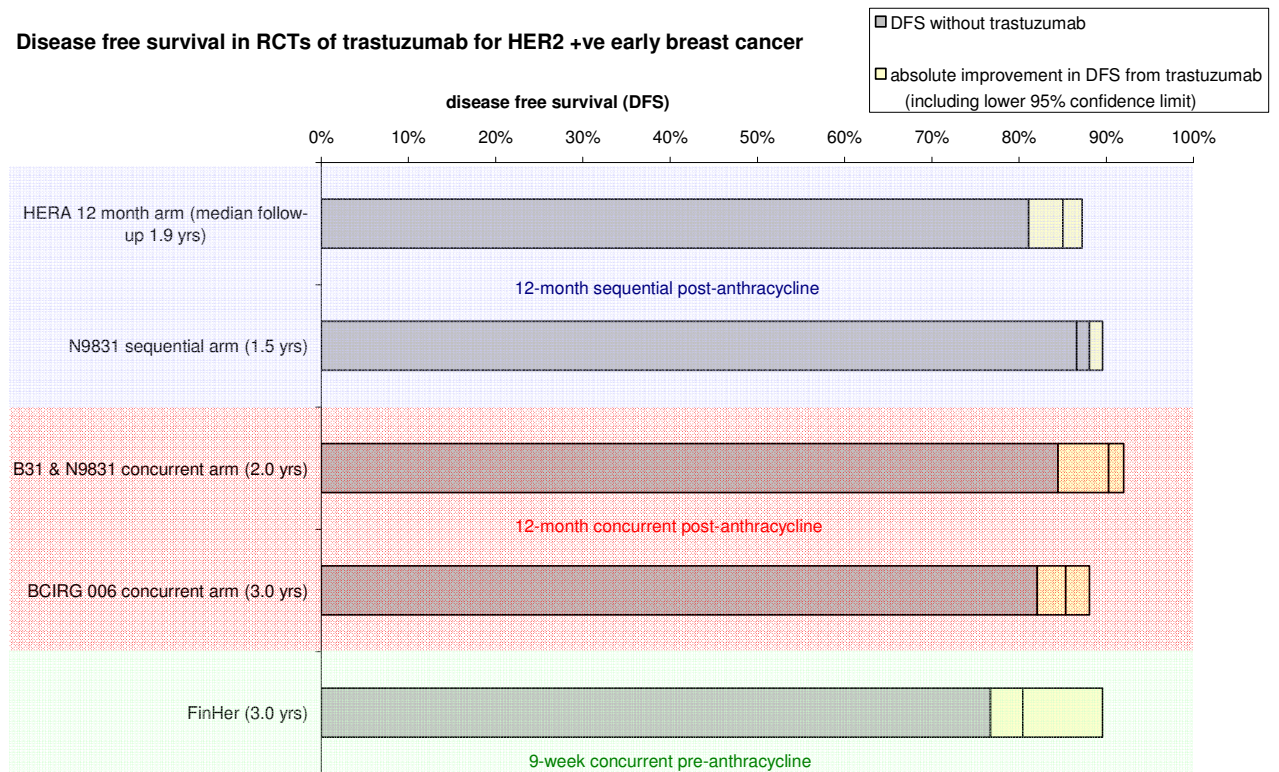


Figure 5A. DFS in adjuvant trastuzumab regimens for HER2-positive early breast cancer for untreated and treated groups

**Disease free survival in trastuzumab regimens for HER2 +ve early breast cancer**

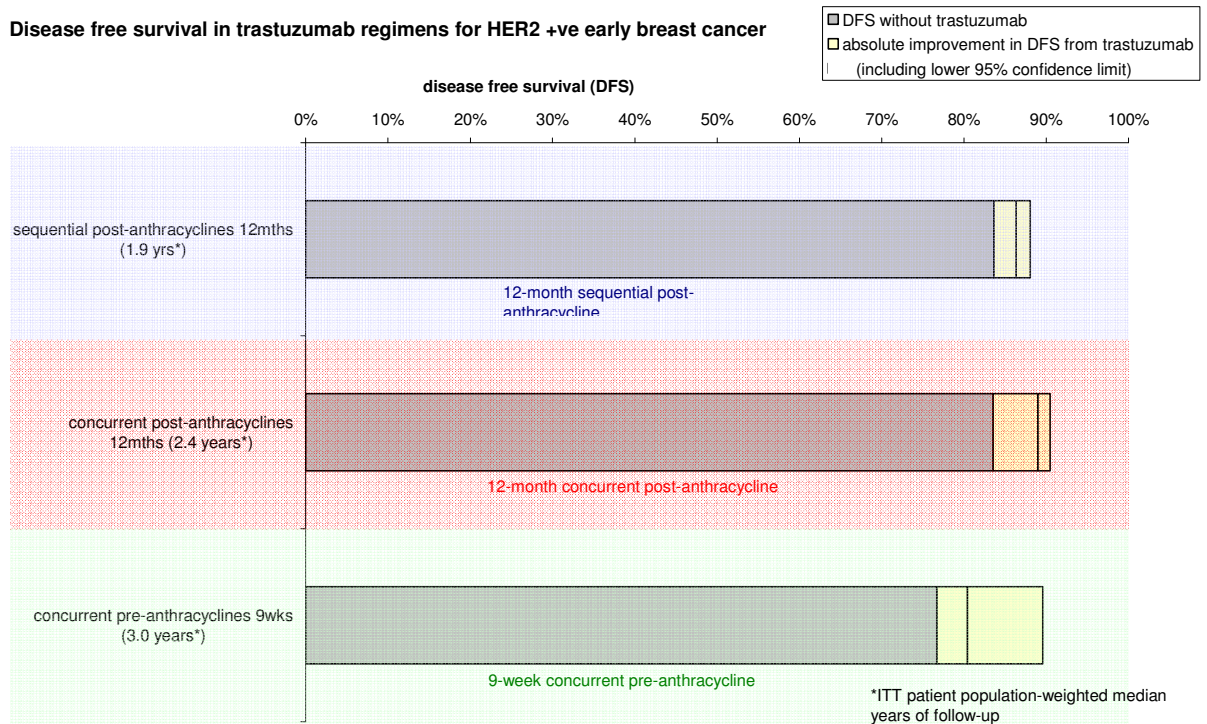
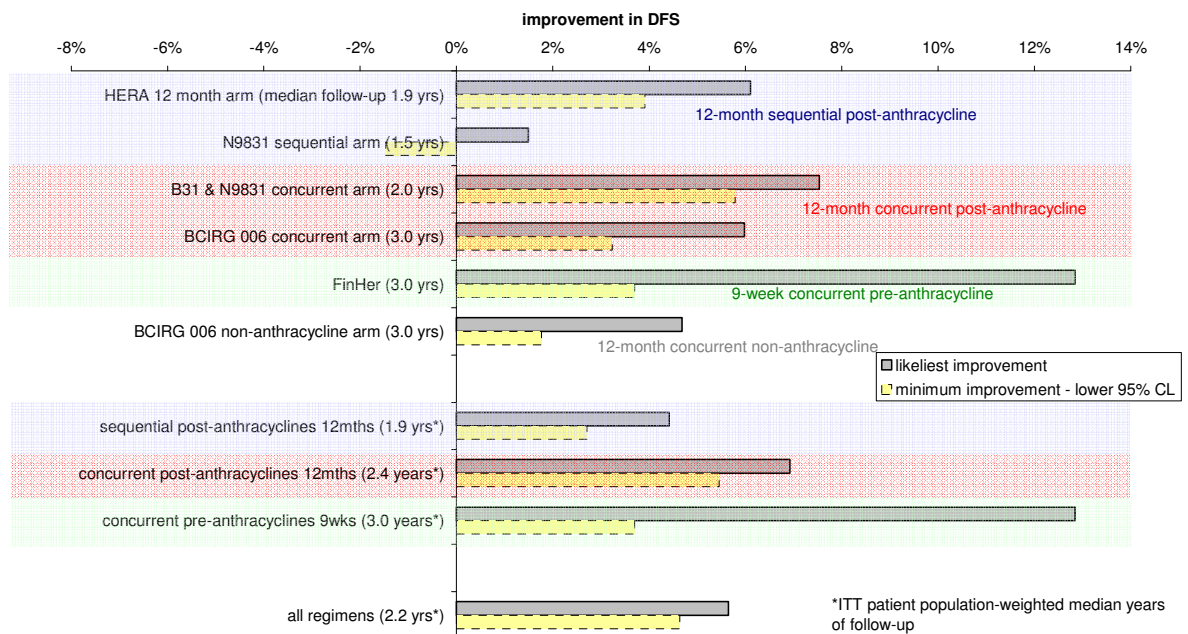


Figure 5B. Absolute improvements in DFS in adjuvant trastuzumab trials in HER2-positive early breast cancer, including minimum confident absolute improvements (from upper 95% confidence limits of hazard ratios)

**Improvements in disease free survival in RCTs of trastuzumab for HER2 +ve early breast cancer**



## 2. Increasing uncertainty around sequence and durations

With data from the five RCTs, questions remain over the extent that sequential treatment prevents recurrence in the short and long term and the duration of this benefit.

Firstly, the trials have a number of methodological issues that may affect their validity:

- results from all five trials reported to date have been preliminary (interim)—all continue to follow-up patients, and none have met their preset target event accruals (e.g. HERA requires 951 primary endpoint events for final analysis<sup>4</sup>);
- it is difficult to assess the quality of two RCTs (B31 and N9831) because reporting has been limited to either a published joint analysis for B31 and the N9831 concurrent arm (Romond 2005<sup>5</sup>)—with little disaggregation into the separate studies<sup>6</sup> or description of key validity aspects of the separate studies—or a conference slideshow presentation for the N9831 sequential arm;<sup>7</sup>
- results for BCIRG 006 too have been limited to conference slideshow presentations<sup>8</sup>, limiting assessibility;
- all trials are open-label in trial design (unblinded);
- allocation concealment methodology is not adequately reported (except FinHer)—where inadequate or unclear allocation concealment has been associated with 30-40% larger estimates of treatment effects<sup>12</sup>; and
- reporting of compliance, contamination and co-intervention has been variable.

Further details on the quality of the trials are available in GATE<sup>13</sup> appraisals undertaken by EPIQ at [http://www.health.auckland.ac.nz/population-health/epidemiology-biostats/epiq/critical\\_appraisal\\_library/Herceptin](http://www.health.auckland.ac.nz/population-health/epidemiology-biostats/epiq/critical_appraisal_library/Herceptin).<sup>14 15</sup>

Secondly, the optimal schedule and duration of treatment for trastuzumab in the adjuvant treatment of early stage HER2-positive breast cancer cannot be determined from the current evidence.<sup>16</sup> This is not unusual for a new treatment with emerging evidence. Particular questions are raised about the durability of benefit, and the scheduling in relation to chemotherapy.

Combinations of trastuzumab with taxane chemotherapy (i.e. concurrent treatment) seem to have a synergistic effect in both adjuvant treatment and treatment of metastatic disease.<sup>17</sup>

However, evidence is available that raises questions about the place of 12-month sequential therapy as the standard-of-care recommended by the supplier in its datasheet<sup>18</sup>. In New Zealand, the Medsafe-approved indication is currently restricted to the HERA regimen<sup>2</sup> (12 months sequential), unlike Australia<sup>19</sup>, where sequential and concurrent (12 months or 9 weeks) is approved and the US<sup>20</sup> where 12 month concurrent treatment is approved.

### ***Non-publication of data for 12 month sequential treatment from N9831***

PHARMAC (and other international bodies such as NICE in the UK) have been asked by the supplier to make decisions on funding trastuzumab on the basis of 12-month median follow-up data from the HERA study<sup>2</sup> (i.e. the sequential treatment regimen). This was supported with longer-term follow-up data from BCIRG 006 and combined results of the NSABP B31 and NCCCTG N9831 studies (all using concurrent treatment).

However, trial N9831, which had three arms, is the only study to have investigated both sequential and concurrent treatment with trastuzumab:

**Arm A (control):** 4 cycles of AC treatment followed by 12 weeks paclitaxel

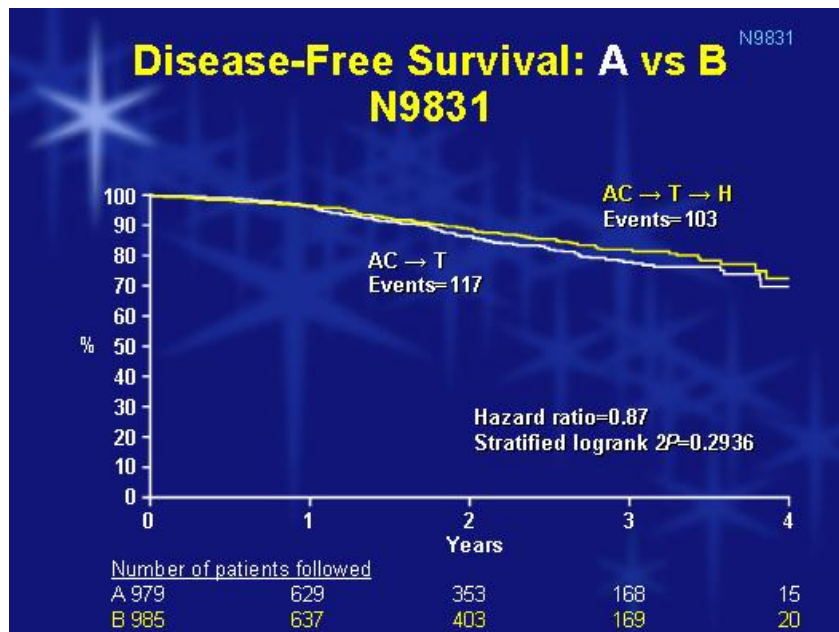
**Arm B (post-anthracycline and sequential with trastuzumab treatment):** 4 cycles of AC treatment followed by 12 weeks paclitaxel and then 52 weeks trastuzumab

**Arm C (post-anthracycline and concurrent with trastuzumab treatment):** 4 cycles AC followed by 12 weeks paclitaxel with trastuzumab, and a further 52 (I think) weeks trastuzumab.

Although not initially provided to PHARMAC and others, separate data from each of the arms of the N9831 from an unplanned interim analysis were presented at the American Society of Clinical Oncology (ASCO) conference in 2005 (at [http://www.asco.org/ac/1.1003\\_12-002511-00\\_18-0034-00\\_19-005815-00\\_21-001.00.asp](http://www.asco.org/ac/1.1003_12-002511-00_18-0034-00_19-005815-00_21-001.00.asp)<sup>7</sup>). The separation of the data for sequential treatment from the concurrent treatment, albeit not yet formally published, is important as it raises questions about optimal treatment scheduling in relation to chemotherapy.

At 1.5 years' median follow-up, N9831 Arm C, one year of trastuzumab administered concurrently with taxane chemotherapy, demonstrated a large and significant benefit in disease-free survival (DFS) over usual care (control Arm A) (HR 0.48, 95% CI 0.39-0.59). However, N9831 Arm B (one year sequential trastuzumab therapy) did not show a statistically significant improvement in DFS compared with usual care (control Arm A) (HR 0.87, 95% CI 0.67-1.13) (see Figure 6A).

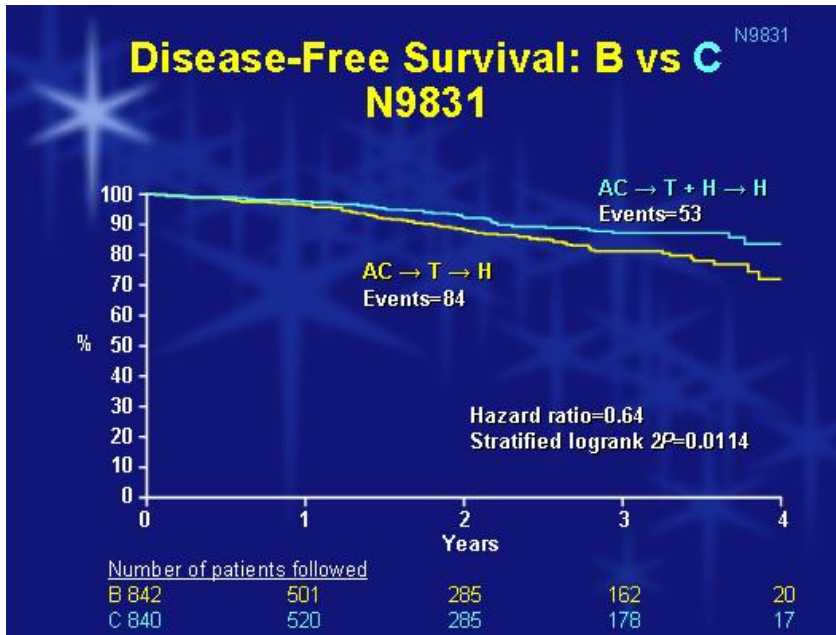
Figure 6A. Disease free survival in trial N9831, sequential trastuzumab vs. control arms



Perez EA. Further Analysis of NCCTG-N9831. Slide presentation at ASCO annual meeting 2005, available online at [http://www.asco.org/ac/1.1003\\_12-002511-00\\_18-0034-00\\_19-005815-00\\_21-001.00.asp](http://www.asco.org/ac/1.1003_12-002511-00_18-0034-00_19-005815-00_21-001.00.asp), accessed January 2007.

Importantly, Arm C (12 months' concurrent trastuzumab) showed a statistically significant improvement in DFS over Arm B (12 months' sequential trastuzumab) standard chemotherapy (HR 0.64, 95% CI 0.46- 0.91)) (see Figure 6B).

Figure 6B. Disease free survival in trial N9831, concurrent vs. sequential trastuzumab arms



Perez EA. Further Analysis of NCCTG-N9831. Slide presentation at ASCO annual meeting 2005, available online at [http://www.asco.org/ac/1.1003\\_12-002511-00\\_18-0034-00\\_19-005815-00\\_21-001.00.asp](http://www.asco.org/ac/1.1003_12-002511-00_18-0034-00_19-005815-00_21-001.00.asp), accessed January 2007.

The sequential regimen in Arm B of N9831 is essentially the same 12-month regimen as examined in the HERA study (i.e. the regimen indicated in New Zealand (Medsafe-approved NZ datasheet<sup>21</sup>)). The conflicting results from these two studies raise questions about the efficacy of 12 months sequential trastuzumab.

The evidence available for N9831 Arm B is confined to a conference presentation of the interim analysis (median f/up 1.5years) and has not been formally published. The results for the sequential arm were not statistically significant, and may have therefore been subjected to publication bias, although there may be other reasons for non-publication.<sup>vii</sup>

At its May 2006 meeting, the Pharmacology and Therapeutics Advisory Committee (PTAC, PHARMAC's independent clinical advisory body <http://www.pharmac.govt.nz/ptac.asp>) considered the above unpublished disease-free survival data for N9831, in the form of MS PowerPoint slides of the conference presentation. The PTAC minute states that members noted that sequential trastuzumab treatment (Arm B) was not statistically superior to non-trastuzumab treatment (Arm A), but that concurrent trastuzumab treatment (Arm C) resulted in a significant improvement in disease-free survival compared with Arm B. Members considered that although these data were preliminary, they raised concerns about the optimal dosing schedule of trastuzumab treatment. PTAC requested that the supplier provide full data from the N9831 trial.

Non-publication of what are said to be interim data reduces confidence in the results; however, earlier 1.5-year data for Arm C (also interim) were used in the high-profile published Romond paper<sup>5</sup> (NEJM 2005) when they were pooled with trial B31 data, and in fact were published individually in the on-line appendix to the Romond paper (<sup>6</sup> Figure 1). The de facto publication of the Arm C results lends legitimacy to the Arm B results, particularly when the key reason stated in that paper for excluding Arm B was simply because trastuzumab was not given concurrently with paclitaxel<sup>5</sup> (and hence was not comparable with trial B31 in joint analysis), when in turn the

<sup>vii</sup> Publication bias emerges when published trials do not represent all trials undertaken, usually because statistically significant results tend to be submitted and published more frequently than indeterminate results.

differences between concurrent treatments in N9831 (Arm C) and B31 were appreciable. [The N9831 Arm C and B31 trials in the Romond interim analysis differed in patient eligibility (high risk negative node status), methods of randomisation allocation, taxane regimens, anthracyclines, sequencing with radiotherapy, sequencing with hormonal therapy, and aromatase inhibitor types and when they started to be used in the trials.<sup>22</sup>]

Another key reason stated for treating the Arm C results differently, that it was planned to be jointly analysed with B31 as approved by the FDA, is unclear, given that the joint analysis may not have been pre-planned.<sup>23</sup>

In general terms, it has been observed that failing to publish means, at best, that ineffective treatments are widely used in patients and, at worst, can lead to unnecessary illness and even death if the reported risks of harms are underestimated.<sup>24 25 26</sup> The authors of the joint N9831-C/B31 trials considered that further follow-up of groups B and C in trial N9831 would be necessary for an adequate evaluation of the efficacy of concurrent as compared with sequential administration of trastuzumab (Romond et al 2005). Professor Ian Smith and colleagues in the publication of the HERA 23-month median follow-up data made similar observations, stating:

“The first [major question regarding adjuvant trastuzumab] is whether trastuzumab started concurrently with taxane chemotherapy (as in the USA trials) is better than trastuzumab starting sequentially after completion of chemotherapy (as here). The NCCTG N9831 trial addresses this issue—this trial includes a third group given sequential trastuzumab. Preliminary data suggest that sequential treatment might be less effective than concurrent treatment [Perez slideshow ASCO 2005] but this was an unplanned comparison with low statistical power, and longer follow-up is needed for confirmation.” (Smith et al. Lancet 2007).

### ***Initial HERA results overstate effects on disease free survival***

Because the HERA study of sequential treatment came to a different result than the sequential arm of N9831, the conflicting information from these studies raises questions about the extent of effect and duration of response to sequential trastuzumab. This becomes more relevant in light of the 23 month median follow-up data from HERA; these reported improvements in overall survival to now be statistically significant<sup>27</sup>, but at the same time showed reduced benefit in disease-free survival over time.

Specifically, in HERA by 23 months median time the risk reduction in DFS for patients who had been treated with trastuzumab had reduced from what had been reported in the 12 month median f/u HERA publication; the HERA disease-free survival 12 month median f/u hazard ratio was 0.54 (95% CI 0.43-0.67) (Piccart-Gebhart et al 2005<sup>2</sup>); 23 month median f/u HR 0.64 (0.54-0.76) (Smith et al 2007<sup>4</sup>). These translated to an 8.3% absolute difference in DFS at the 12 month median follow-up, reducing to 6.1% at 23 months, this reduction being statistically significant<sup>28</sup> (detailed later).

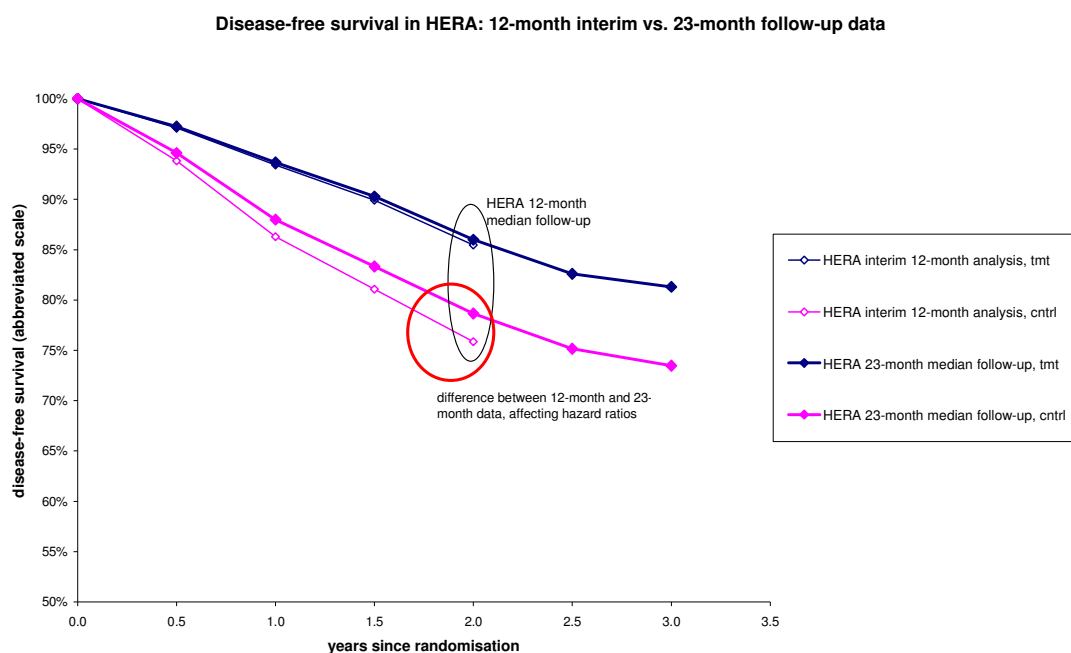
Note that this apparent decrease in the effect on DFS over time in the HERA trial does not appear to be due to the appreciable contamination that occurred in the study (where patients in the comparator arm were offered trastuzumab treatment once the 12-month DFS results emerged, causing a substantial loss of patients in the control arm in a non-randomised fashion). The opportunity to cross-over occurred relatively late, and the hazard ratios for both intention-to-treat analysis (ITTA) of all patients and a censored analysis (where cross-over patient data were removed) were virtually identical for DFS (0.64 ITTA, 0.63 censored analysis).<sup>4</sup> This

concordance of hazard ratios suggests a genuine effect (rather than artefact due to comparator group patients latterly receiving trastuzumab treatment).

Importantly, this cross-over has limited the ability of the HERA study to examine longer-term outcomes for 12 months' sequential trastuzumab compared to standard chemotherapy, because the standard chemotherapy arm has largely been lost, possibly to a point where longer-term comparison questions can never be resolved.

The impact of timing of analysis in HERA, with the apparent waning of treatment effect over time, can be seen in the following graph (Figure 7) of disease free survival over time (Figure 7; note the use of abbreviated scale for DFS).

Figure 7. Comparison 12-month and 23-month median follow-up DFS in HERA



***Impact of N9831 Arm B and the updated HERA results on disease free survival benefits of 12 month sequential treatment***

The non-inclusion of the Arm B N9831 data and use of the interim HERA results (limited to 12 months median follow-up) means that initial estimates of the effectiveness of trastuzumab in the sequential setting may have been overstated. Including the available interim N9831 Arm B data and then updating the HERA results for the 23-month median data means that the effects of sequential treatment on disease progression reduce by one third—the overall hazard ratio increases from 0.54 (95% CI 0.43-0.67) for the HERA 12-month median follow-up data to 0.70 (0.61-0.81) for the pooled HERA median 23-month follow-up and N9831 Arm B data.

The potential impact on the overall effectiveness (HR) for 12 month sequential treatment of the missing N9831 Arm B vs. A comparison may be even larger, considering the potential added events accruing since the time of its initial analysis. The N9831 Arm B vs. A events were analysed at around April 2005, compared with May 2006<sup>4</sup> for the 23-month median follow-up for HERA trial. Had the N9831 B vs. A results been re-analysed and then presented at the same time

as were HERA 23-month median follow-up data at the 2006 ASCO conference, the additional numbers of events would influence the weight given to N9831 Arm B overall—similar to the impact of time on event numbers in the HERA follow-ups, and as suggested in the following graphs (Figures 8 and 9) that model possible disease events (numerical) and disease free survival (comparative) over time.

Figure 8. Timecourse of disease recurrence events in HERA and N9831 Arm B vs. A

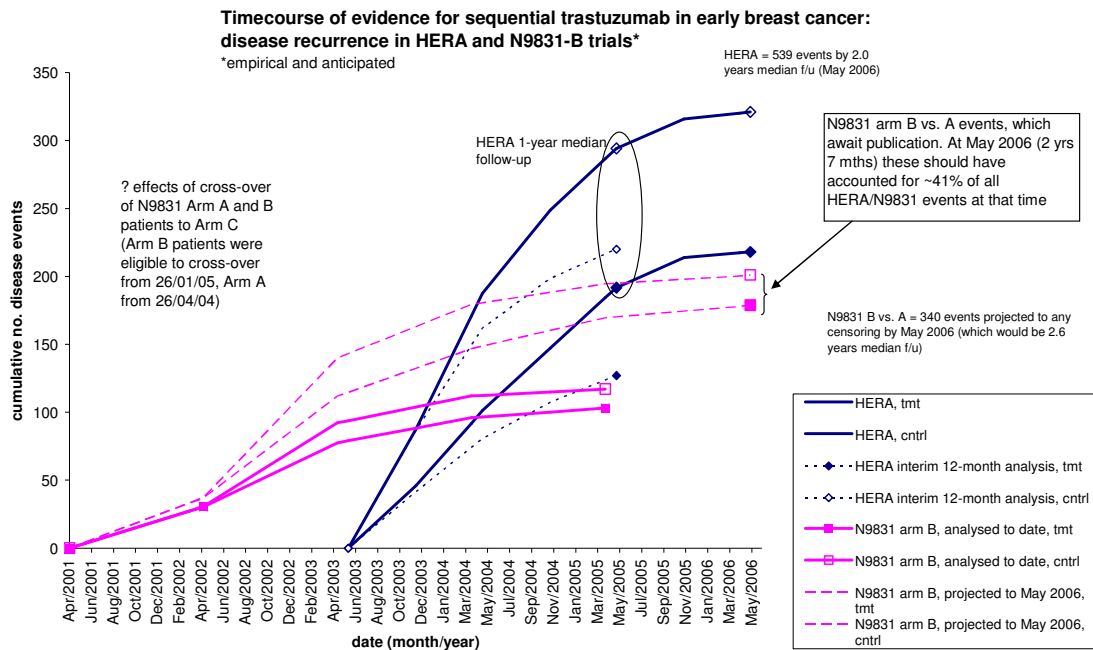
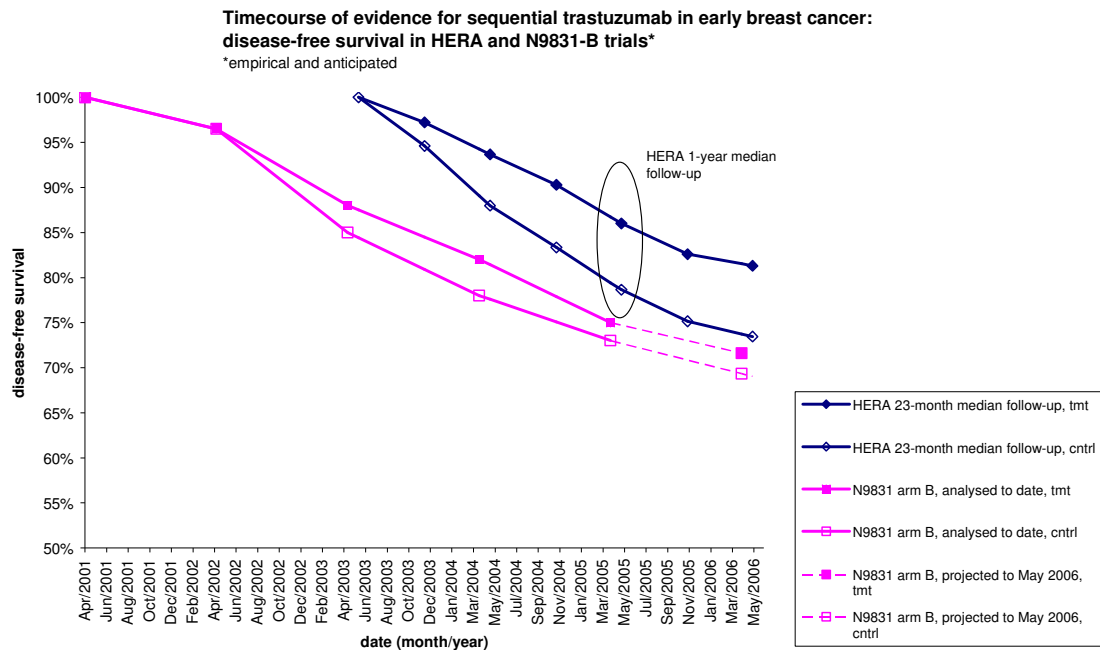


Figure 9. Timecourse of DFS in HERA and N9831 Arm B vs. A



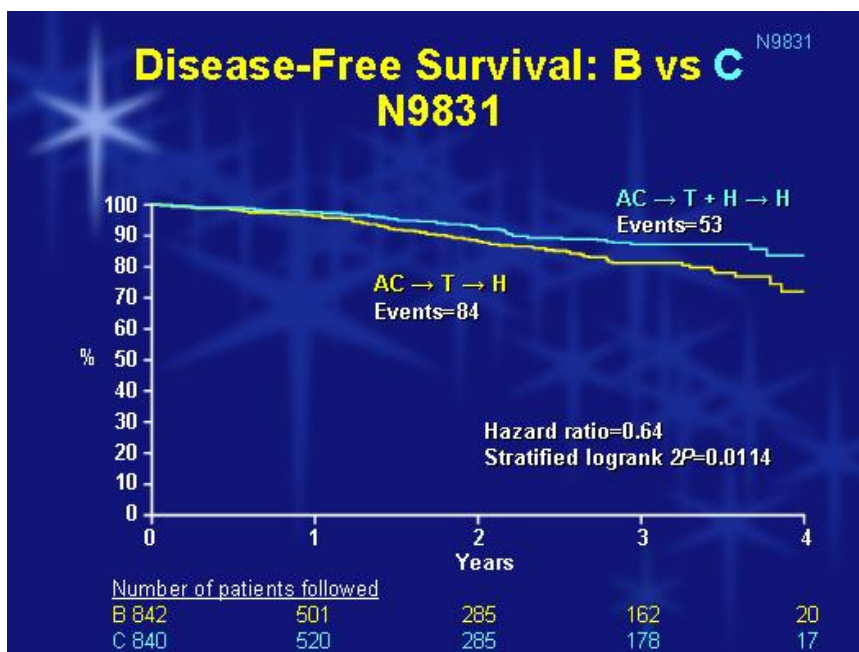
However, the results for N9831 Arm B were an unplanned interim analysis with low statistical power, and longer follow-up is needed for confirmation.<sup>5</sup> PTAC (February 2007) considered that there was now likely to be longer-term follow-up of outcomes (disease free survival and mortality) in this study, and asked for all the updated data from all three arms of the trial be made available. Such data should be available, the authors of the joint B31/N9831 Arm C analysis (Romond et al.) commented that the group C vs. B comparison requires longer follow-up and that such further follow-up is necessary.<sup>5</sup>

***The impact of sequencing on the effectiveness of 12 month treatment***

Trial N9831 was also designed to directly address the efficacy of trastuzumab initiated concurrently with paclitaxel versus sequentially after completion of paclitaxel (Arm C vs. Arm B). Despite an overall small number of events, the N9831 trial showed a statistically significant benefit of concurrent therapy over sequential therapy for disease progression (HR 0.64, 2P logrank 0.0114), although this finding is qualified<sup>viii</sup> (Figure 6B).

Figure 6B, repeated. Disease free survival in trial N9831, concurrent vs. sequential trastuzumab arms

<sup>viii</sup> The median duration of trastuzumab treatment was likely to be longer in the concurrent arm (C) than the sequential arm (B). This was due to the different timing of trastuzumab sequencing, with the sequential arm commencing trastuzumab 12 weeks later. Hence patients in the concurrent arm were exposed to higher cumulative doses of trastuzumab at the time of the interim analysis—which may have contributed to both the concurrent arm’s apparent improved disease progression and higher cardiotoxicity.



Perez EA. Further Analysis of NCCTG-N9831. Slide presentation at ASCO annual meeting 2005, available online at [http://www.asco.org/ac/1.1003\\_12-002511-00\\_18-0034-00\\_19-005815-00\\_21-001.00.asp](http://www.asco.org/ac/1.1003_12-002511-00_18-0034-00_19-005815-00_21-001.00.asp), accessed January 2007.

However the concurrent arm of N9831 (Arm C) had a higher incidence of cardiovascular events than the sequential arm (Arm B), with similar caveats to the efficacy data<sup>viii</sup>, described later.

The Arm C vs. Arm B comparison in N9831 represents an internally valid direct head-to-head comparison of concurrent vs. sequential treatment regimens, and these early data<sup>29</sup> are the only available to directly compare the two regimens. These results showed statistically significant improvement in DFS with concurrent over sequential treatment. According to the Romond et al. joint analysis of trials B31 and N9831, after reviewing the results of the first joint interim efficacy analysis, the data monitoring committee overseeing trial N9831 requested an unplanned comparison of groups B and C and subsequently recommended disclosure of the results. Romond et al considered that though early, the comparison suggested delayed (i.e. sequential) administration of trastuzumab may be less effective than concurrent administration.<sup>5</sup>

Another large study of concurrent treatment, BCIRG 006, also demonstrated a similar benefit post-anthracycline in DFS over median 3 years of follow-up (HR 0.61 (0.48-0.76) for Arm AC-TH).

Further evidence, if less direct, suggesting improved efficacy of concurrent over sequential treatment can be seen in the results for both disease free survival and overall survival for all trials, and pooled results for regimens, shown earlier in tables 1 to 3 and figures 4 to 5B. These indicate overall (pooled) HRs for disease recurrence of 0.53 (0.46-0.62) for concurrent treatment post-anthracycline and 0.42 (0.21-0.83) for concurrent treatment pre-anthracycline, compared with (less efficacious) 0.70 (0.61-0.81) for sequential treatment post-anthracycline. Similarly, HRs for overall mortality were 0.63 (0.50-0.80) and 0.41 (0.16-1.08) for concurrent treatment post- and pre-anthracycline respectively, compared with 0.72 (0.55-0.94) for sequential treatment post-anthracycline.

### ***The impact of sequencing on the cardiotoxicity of 12 month treatment***

All three studies of concurrent treatment (BCIRG 006, N9831 and B31) used a regimen of trastuzumab administered concurrently with a taxane in addition to extended (up to 12 months) treatment on completion of anthracycline-containing chemotherapy. Anthracycline treatment has a dose-related cardiotoxic effect on the heart, and as a result 15-20% of patients did not meet the cardiac criteria required for trastuzumab treatment initiation where trastuzumab was initiated after patients received cardiotoxic anthracycline treatment.<sup>30</sup>

The concurrent treatment regimens used in N9831 (Arm C) and NSABP B31 (i.e. anthracycline chemotherapy followed by 12 months trastuzumab started concurrently with a taxane) demonstrated a higher risk of cardiotoxicity than the sequential treatment arms (HERA and Arm B of N9831). The pooled analysis of N9831 and NSABP B31<sup>5</sup> reported that 14.2% of trastuzumab treated patients discontinued treatment before 52 weeks because of a confirmed asymptomatic decline in left ventricular ejection fraction (LVEF). Another 4.7% discontinued because of symptoms of congestive heart failure (CHF) or another adverse cardiac effect. In addition, 30.5% of patients required at least one trastuzumab dose delay in because of a decrease in LVEF or cardiac symptoms. The cumulative incidence for New York Heart Association (NYHA) class 3 or 4 CHF/death from cardiac causes at 3 years was 0.8% and 0% respectively in the control group, and 4.1% and 2.9% in the trastuzumab group.

The above information suggests that, although more efficacious than 12 month sequential treatment (overall HR for DFS = 0.49 for concurrent compared with HR = 0.66 for sequential), the 12 month concurrent regimen (post anthracycline) is less safe for patients because of the associated risk of cardiotoxicity. However, it should be noted that patients in HERA were randomised into the study after completion of their chemotherapy; therefore patients with anthracycline-induced cardiac toxicity would not have been enrolled in the study and therefore were not captured in the denominators. Hence the HERA trial population can generally be considered to be at reduced risk of further cardiotoxicity than the B31/N9831 or BCIRG 006 trial populations, which may have contributed to the lower rates for cardiotoxicity seen in HERA.

### ***Durability of response***

To reiterate, the durability of response of sequential therapy has not been demonstrated. Extended follow-up of the HERA study population is unlikely to quantify the sequence's long-term benefits and risks, due to the partial loss of the control arm (with potential bias from non-randomisation). As discussed earlier, the 2-year follow-up HERA results showed reduced benefit in disease-free survival, when the interim analysis (Piccart-Gebhart et al 2005) reported a disease-free survival hazard ratio over 12-months' median follow-up of 0.54 (95% CI 0.43-0.67), but the 2-year median follow-up data (Smith et al 2007) showed a HR of 0.64 (0.54-0.76).

The 23-month follow-up report for the HERA study included period-specific hazards by time, showing a change in disease recurrence over time, with convergence of hazards and overlapping of their confidence intervals beyond ~18 months – see Figure 10 below:

Figure 10. Annualised DFS hazard rates, HERA one-year trastuzumab treatment vs. observation arms

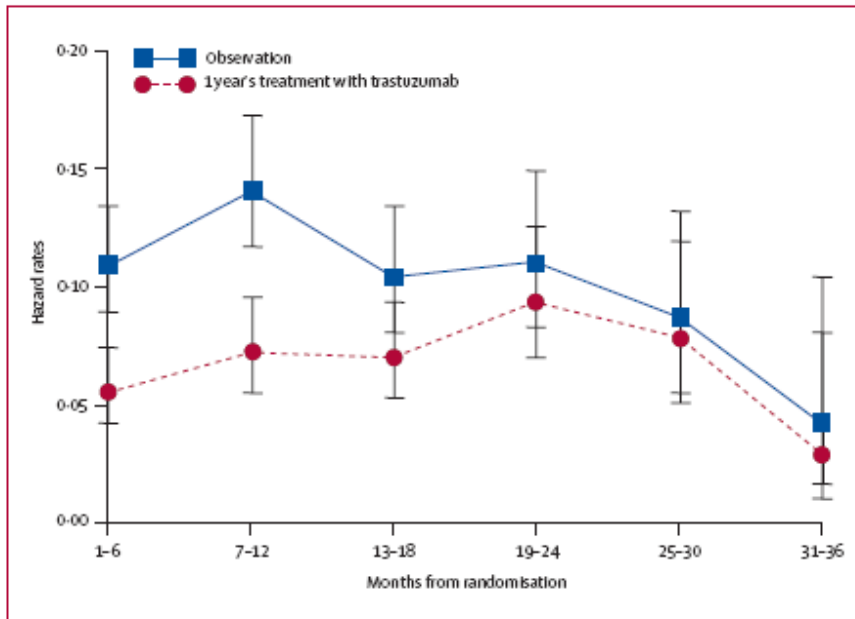


Figure 4: Annualised disease-free survival hazard rates for 1 year of trastuzumab vs observation

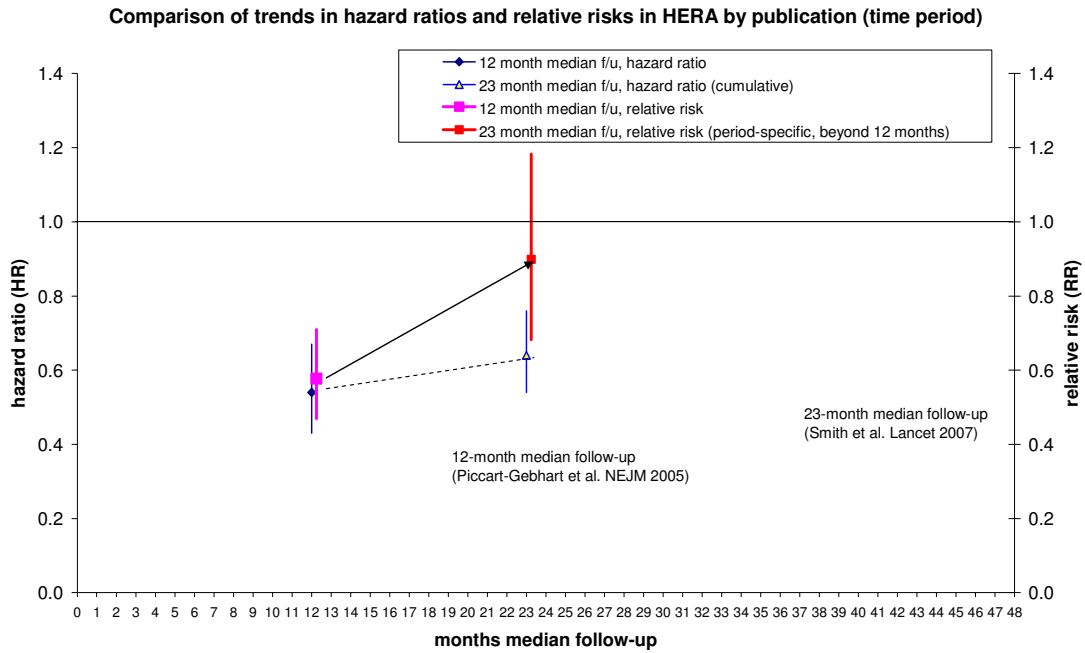
Source: Smith et al 2007

Ideally the reduction in DFS benefits over time in the HERA trial should be confirmed by testing proportional hazards by time.<sup>ix 31</sup> However, in the absence of any available such analysis of individual patient data, the reduction in DFS benefits over time reported in the HERA trial (the 0.54 HR becoming 0.64)<sup>2 4</sup> is statistically significant on testing for interaction<sup>32 33</sup> by the two time periods ( $p=0.02$ ).<sup>28</sup>

The effects of the discord between the cumulative hazard ratio reported in the 23-month HERA f/u and period-specific effects from events occurring after the 12-month interim f/u can be seen in the following graph (Figure 11). This demonstrates a greater waning of effects with period-specific relative risk. The similarities between the published 12-month median f/u hazard ratio (HR 0.54) and the calculated relative risk for the same time period (RR 0.58) give some comfort around the above testing for statistical interaction.

Figure 11. Cumulative hazard ratios and period-specific relative risks in HERA by time

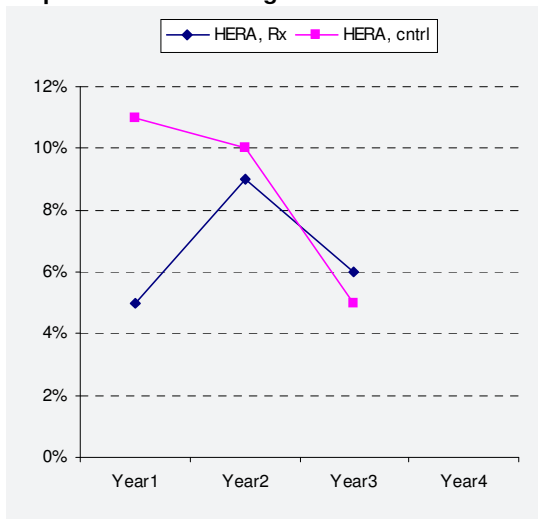
<sup>ix</sup> Testing for proportional hazards over time will be to determine whether the effect of treatment is constant or varies significantly over time



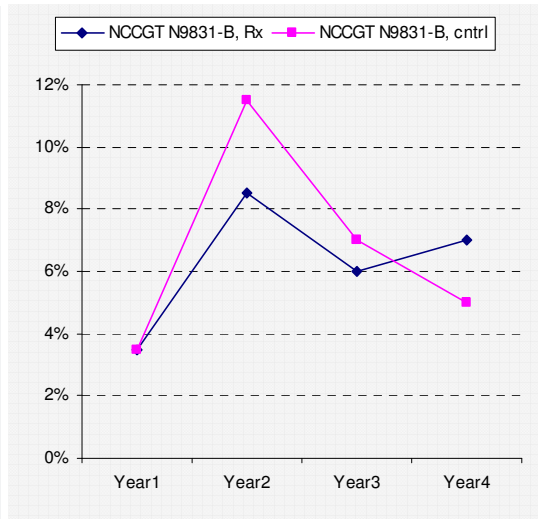
For sequential regimens, the data therefore suggest convergence in period-specific rates perhaps by 3 years, i.e. trastuzumab no longer has any effects on disease events. For concurrent regimens, the data suggest a mixture of continuing divergence and some convergence by 4 years—see the following graphs in Figure 12 that compare death or relapse rates per year for trastuzumab versus standard treatment in the five trials.

Figure 12. Estimated annual prevalence rates in trastuzumab trials, from visual abstraction of available disease free survival curves

**Sequential 12 month regimens**

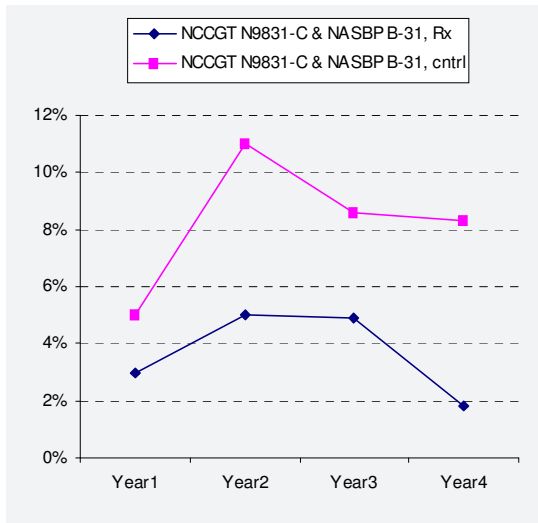


HERA 12-month regimen

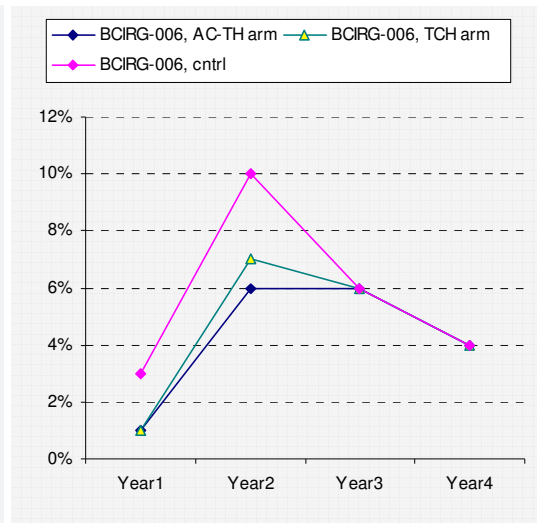


N9831 Arm B vs. Arm A

**Concurrent 12 month regimens**

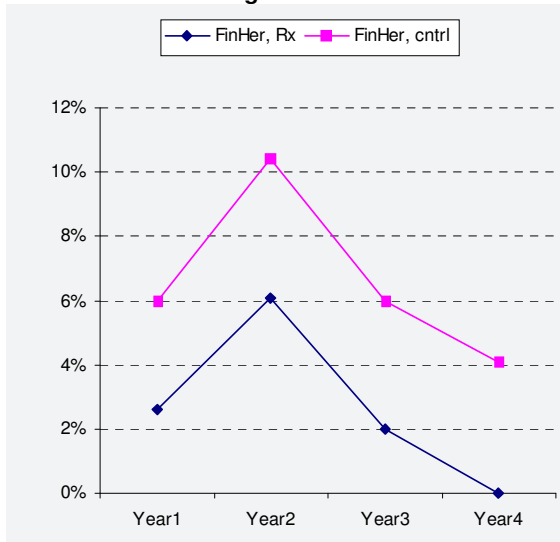


Joint analysis N9831 arm C vs. arm A, B31 (Romond)



BCIRG 006 AC-TH

### Concurrent 9 week regimens



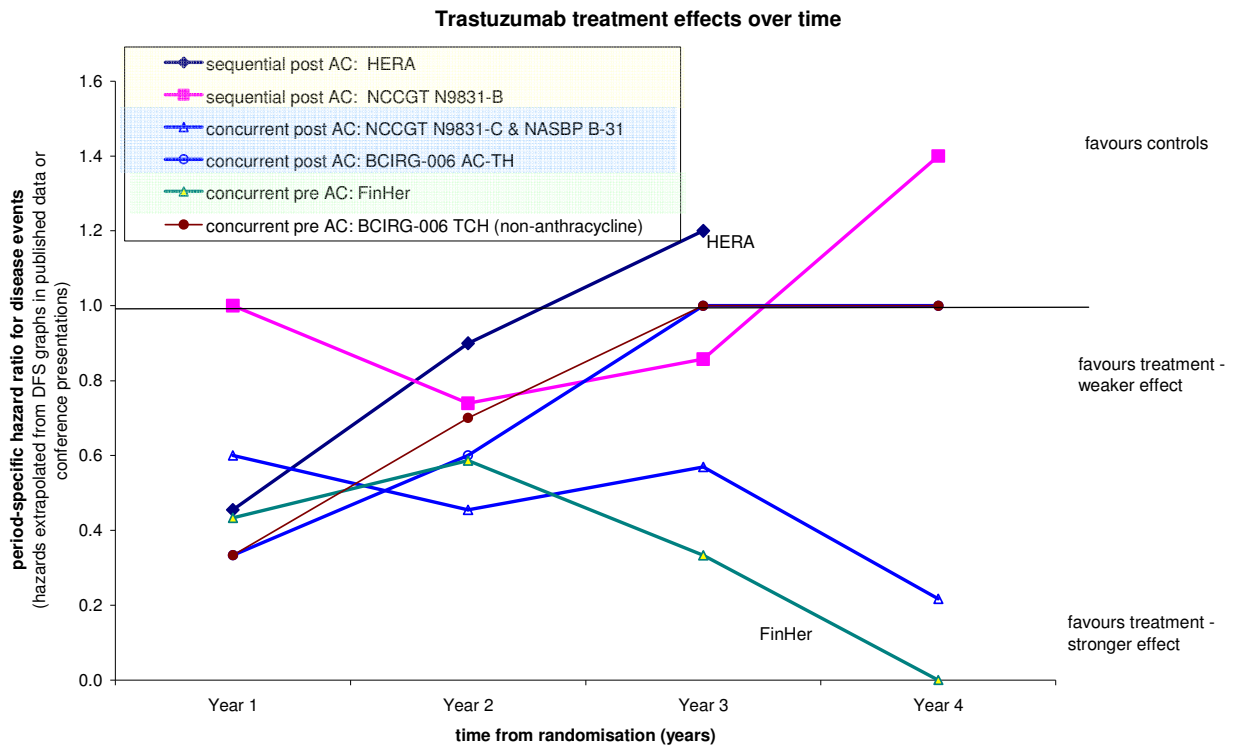
FinHer

These results must however be treated with caution, given the small numbers of patients and events as time progresses (and hence instability with greater uncertainty in the later years).

Using the above data, the following figure shows changes in hazard ratios over time for relevant trials, calculated from the published or reported DFS survival graphs. This suggests apparent waning of effects with the HERA trial (with statistical heterogeneity by time period<sup>28</sup>), also for trial N9831 Arm B (sequential)<sup>x</sup> and the BCIRG 006 trial (concurrent) with significant time-related statistical interaction ( $p=0.01$ )<sup>34</sup>, but not for other concurrent regimens (B31/N9831 Arm C joint analysis<sup>x</sup>, FinHer<sup>x</sup>) (Figure 13).

<sup>x</sup> Tests for statistical interaction by time period are available for HERA and BCIRG 006, as both these studies have available event numbers obtainable from interim analyses. No such tests for statistical interaction by time period are performable for trials N9831, B31 and FinHer, as these studies have reported their interim results only once. Ideally such testing should be performed using individual patient data and testing proportional hazards by time.

Figure 13. Efficacy of trastuzumab over time (period-specific hazard ratios for DFS)



### 3. Concurrent treatment pre-anthracyclines (FinHer and E2198)

#### *FinHer*

FinHer was an open-label RCT, comparing docetaxel with vinorelbine for the adjuvant treatment of early stage breast cancer (n=1010), where women with HER2-positive cancers were also randomised to receive nine weeks concomitant treatment with trastuzumab or no such treatment (n=232). Examining trastuzumab in FinHer is in effect a nested RCT for HER2-positive patients [comparing trastuzumab (with docetaxel or vinorelbine) vs. no trastuzumab (but still with docetaxel or vinorelbine)], sited within the wider RCT [comparing docetaxel with vinorelbine (with randomisation to trastuzumab or no trastuzumab for those patients who were HER2+ve)]. Hence the trastuzumab component of the trial has randomised HER 2 positive patients to trastuzumab or no trastuzumab and to docetaxel or vinorelbine (see diagram).

no.patients in FinHer subgroups

		docetaxel	vinorelbine	total
HER2+ve	trastuzumab	54	62 (61)*	106 (105)
	no trastuzumab	58	58	116
HER2 -ve	trastuzumab	0	0	0
	no trastuzumab	390 (389)**	388	778 (777)
<b>total</b>		<b>502 (501)</b>	<b>508 (507)</b>	<b>1010 (1008)</b>

\* 1 excluded from analysis

\*\*1 excluded from analysis because had HER2+ve CAB

Results to date, as reported in the NEJM in 2006, were from a pre-planned early interim analysis, pending final analysis at five years median follow-up (or 150 events overall, whichever occurring earlier). Data from final analysis may be available from mid-2007 (the early efficacy analysis (NEJM 2006) was for median follow-up three years and took place in May 2005) (Joensuu et al NEJM 2006<sup>9</sup>).<sup>35</sup>

The shorter duration concurrent regimen for adjuvant trastuzumab in the FinHer study was based on two theories:

1. trastuzumab acts synergistically when given concurrently with chemotherapy (indicated by data from the metastatic disease setting; note that this synergism hypothesis has subsequently been supported by the trial N9831 comparison of sequential vs. concurrent<sup>7</sup>); and
2. cardiotoxicity may be better managed when trastuzumab treatment is given concurrently with taxanes but prior to cardiotoxic anthracycline chemotherapy.<sup>36</sup>

The efficacy results of the FinHer trial to date suggest that administration of trastuzumab for nine-weeks concurrently with a taxane is effective in the treatment of HER2-positive early stage breast cancer. Trastuzumab was administered weekly concurrently either with three-weekly docetaxel or weekly vinorelbine, followed by three cycles of FEC (5-fluorouracil/ epirubicin/ cyclophosphamide) in each arm. Docetaxel was associated with improved recurrence-free survival as compared to vinorelbine (hazard ratio 0.58, 95% CI 0.40 to 0.85), and the addition of trastuzumab to chemotherapy was associated with improved recurrence-free survival as compared to the same chemotherapy administered without trastuzumab (hazard ratio 0.42, 95% CI 0.21 to 0.83).

During a median follow-up time of 3 years reported to date, these treatments have not been associated with detectable cardiac toxicity, although caution is required with this interpretation.<sup>xi</sup>

<sup>xi</sup> The low cardiotoxicity observed in FinHer could also be explained by the relatively low cumulative dose of anthracycline chemotherapy (180 mg/m<sup>2</sup> epirubicin while the maximum tolerated cumulative dose of epirubicin is of

The PTAC considered that the number of patients treated in the FinHer study was not insignificant and that the data from that trial were valuable.

#### Possible synergism of docetaxel with trastuzumab

The main aim of FinHer was to determine which treatment was the most effective—vinorelbine or docetaxel—in early stage breast cancer (with and without HER2 amplification). In patients that did and did not have HER2 amplification (i.e. all patients), docetaxel had better efficacy in terms of disease progression—the HR for DFS favoured docetaxel over vinorelbine (HR 0.58 (0.4-0.85)). The total patient numbers initiating treatment in this study were 507 (vinorelbine) and 502 (docetaxel). Overall survival (OS) favoured docetaxel, but this was not statistically significant (HR 0.66 (0.38-1.17)).

The HER2-positive trastuzumab treated patients group did not undergo a prospectively-defined subgroup analysis to examine its effects according to underlying docetaxel or vinorelbine chemotherapy. However, in response to concerns raised about mixing results for docetaxel with vinorelbine for the trastuzumab-treated HER2-positive patients, informal post-hoc analysis of these patients who were treated with docetaxel and trastuzumab reveals persisting statistically significant efficacy for this subgroup, despite a halving patient numbers. Such analysis indicates a relative risk for disease recurrence of 0.27 (0.08-0.90)<sup>xii</sup>—where docetaxel + FEC + trastuzumab vs. docetaxel +FEC (no trastuzumab) = 3/54 vs. 12/58 events/patients = 5.6% vs. 20.7%. This compares with a worse and statistically non-significant result for vinorelbine + FEC + trastuzumab patients in FinHer (RR for trastuzumab vs. no trastuzumab of 0.57 (0.27-1.20)).<sup>37</sup>

It is acknowledged that this is a small, retrospectively-defined post-hoc subgroup analysis with no evidence of statistical heterogeneity on formal testing,<sup>37</sup> and as such the results should be treated with caution. However, although such results cannot statistically be validly extrapolated to estimate expected results in clinical practice, they do support the hypothesis, along with in vitro evidence,<sup>38</sup> that docetaxel and trastuzumab have a synergistic effect when they are used together as per the FinHer regimen and that the real-life risk reductions are likely to be more, if anything, than seen with sequential 12 month treatment.

#### ***Reliability and validity of the FinHer trial and nine-week concurrent treatment***

The data from FinHer show that that the combination of docetaxel plus 9-week concurrent trastuzumab is effective and well tolerated in the treatment of HER2-positive early breast cancer. However, concerns have been raised that the evidence for 9-week treatment is less reliable because of small patient numbers showing effects only on disease free survival in only one clinical trial (the FinHer study). In other words, the subgroup of the FinHer study that evaluated adjuvant trastuzumab was small, with 232 patients randomized to receive or not to receive trastuzumab, overall survival (OS) reported to date has not been statistically significant (HR 0.41), and the data have not been formally confirmed by other trials. These three issues are addressed as follows:

---

720mg/m2), and perhaps by less sensitive LVEF testing/thresholds. In the B31/N9831 studies doxorubicin was administered at a cumulative dose of 240 mg/m2 while its maximum tolerated cumulative dose is only 500 mg/m2. Epirubicin is generally presented as a less cardiotoxic agent than doxorubicin. Indeed, as stated in the FinHer paper, the small size and the short duration of the follow-up are limitations of the study and the optimal duration of adjuvant trastuzumab therapy is not known and may be clarified only in further randomized trials. (KCE2006)

<sup>xii</sup> chi-square 0.02 (df=1), Fisher 2-tail exact probability 0.03

## Patient numbers

In terms of patient numbers in the FinHer trial (the wide 95% CI, 0.21-0.83, which reflected the degree of uncertainty from small numbers), the trial gave results that were statistically significant despite its smaller size. This reflects the strong efficacy of short duration concurrent treatment in this setting (more than halving disease recurrence), so that fewer patients were needed to confirm such a strong effect.

Because this trastuzumab regimen was so effective, HER2-positive patients studied in FinHer (assigned to trastuzumab or no trastuzumab) could have been as few as 145 for the results to still be statistically significant (calculated using binomial methods<sup>39</sup> from the central estimate relative risk).<sup>40</sup>

Concerns about the reliability of the FinHer data because of small patient numbers ignore the large effects (strong efficacy) seen in the study, and represent a statistical misunderstanding that study populations/sample sizes (denominators) drive variability of effect; in fact it is event numbers (numerators) that are the more important.<sup>41 42 xiii</sup> Large treatment effects—likely to be more clinically worthwhile—but with wider confidence intervals (greater imprecision) should not be ignored essentially because of less power (this is similar to where ‘absence of evidence is not evidence of absence’<sup>43</sup>). Conversely, more precise treatment effects from larger samples (with narrower confidence intervals) but with smaller treatment effects are likely to have less impact clinically (i.e. greater confidence but a lesser effect on outcomes); such evidence should not override less precise evidence as of right.

Such concerns about the patient numbers in FinHer<sup>44</sup> seem analogous to post-hoc power calculations, where in fact once results are available, a trial yields a treatment effect and confidence interval for the results; the power of the trial is expressed in that confidence interval, and hence the power is no longer a meaningful concern.<sup>41</sup> The first modern RCT—streptomycin for tuberculosis, undertaken in 1948—had only 107 patients<sup>45</sup>, and is a classic example of a big effect not needing big patient numbers. Recently, a phase III RCT involving 208 patients (TAnDEM) was the basis of the approval in the European Union of trastuzumab for use in combination with aromatase inhibitors for treatment of postmenopausal patients with HER2 and hormone receptor-co-positive metastatic breast cancer<sup>46</sup>—being fewer patients than in the FinHer trastuzumab nested RCT.

In addition, similar scrutiny to study size and effects can also apply to the HERA and other sequential data (i.e. all the information available for the sequential 12 month regimen, being the regimen strongly advocated for<sup>49</sup>). The hazard ratios for all 12 month sequential data combined (HERA and N9831 Arm B) and for FinHer are highly similar across all ranges, suggesting few differences despite FinHer’s smaller numbers. Combining HERA 2-year follow-up and N9831 Arm B data reduces effectiveness by 1/3<sup>rd</sup> to reach a hazard ratio for 12 months’ sequential

---

<sup>xiii</sup> When examining variability of effects, the influence of numbers of events rather than numbers of patients can be seen by examining the standard formulae for variance of relative risk for discrete dichotomous (binary) outcomes (Rothman 1986):

95% CI for RR =

$$e^{(\log_e(RR) \pm Z \times \sqrt{(\frac{1}{n_1} - \frac{1}{N_1}) + (\frac{1}{n_2} - \frac{1}{N_2})})}$$

where RR (relative risk) =  $(n_1/N_1)/(n_2/N_2)$ , Z (standardised normal variate (with  $\mu = 0, \sigma = 1$ ) = 1.96 for a 95% CI  
 $n_1$  = no. events in experimental group,  $n_2$  = no. events in observational group,  
 $N_1$  = no. subjects in experimental group,  $N_2$  = no. subjects in observational group  
 $n_1/N_1$  = incidence rate for experimental group,  $n_2/N_2$  = incidence rate for observational group

treatment of 0.70 (0.61-0.81); the upper confidence limit for this hazard ratio (i.e. the minimum extent that disease recurrence can confidently be expected to reduce) is similar to FinHer's:

- Historically (and the basis for funding decisions internationally for 12 months' sequential trastuzumab treatment), HERA's interim hazard ratio (HR) was 0.54 (95% CI 0.43-0.67);
- However, the combined updated HERA/N9831-B HR increases the hazard ratio for sequential treatment to **0.70 (0.61-0.81)**, thus a relative hazard reduction (RHR) of **30%** (95% CI **19%-39%**);
- This compares with the FinHer HR of **0.42 (0.21-0.83)**, being a RHR of **58%** (**17% - 79%**).

Hence, the most realistic estimates suggest the FinHer regimen should be at least as effective as sequential regimens (RHR 58% vs. 30%). More importantly, at the very worst—i.e. the minimum extent that disease recurrence can confidently be expected to reduce—and even accounting for its smaller number of patients, the FinHer results were nearly as effective as sequential regimens (17% vs. 19%)—see graphs below (figures 14 and 15).

Figure 14.

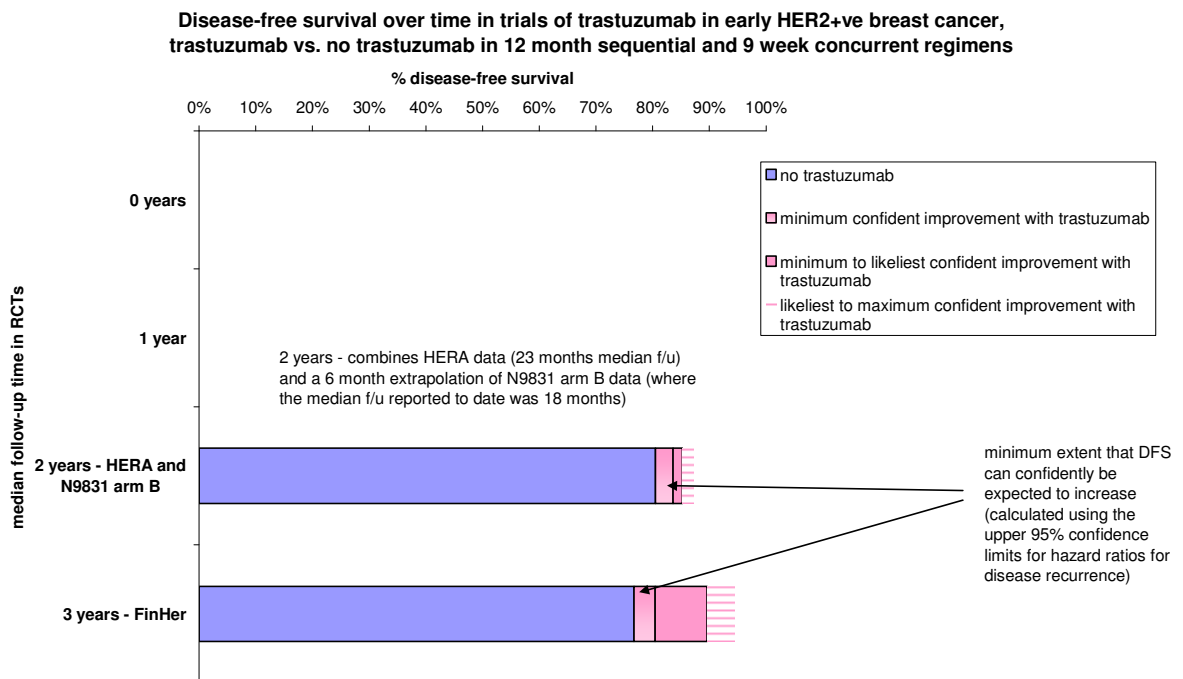
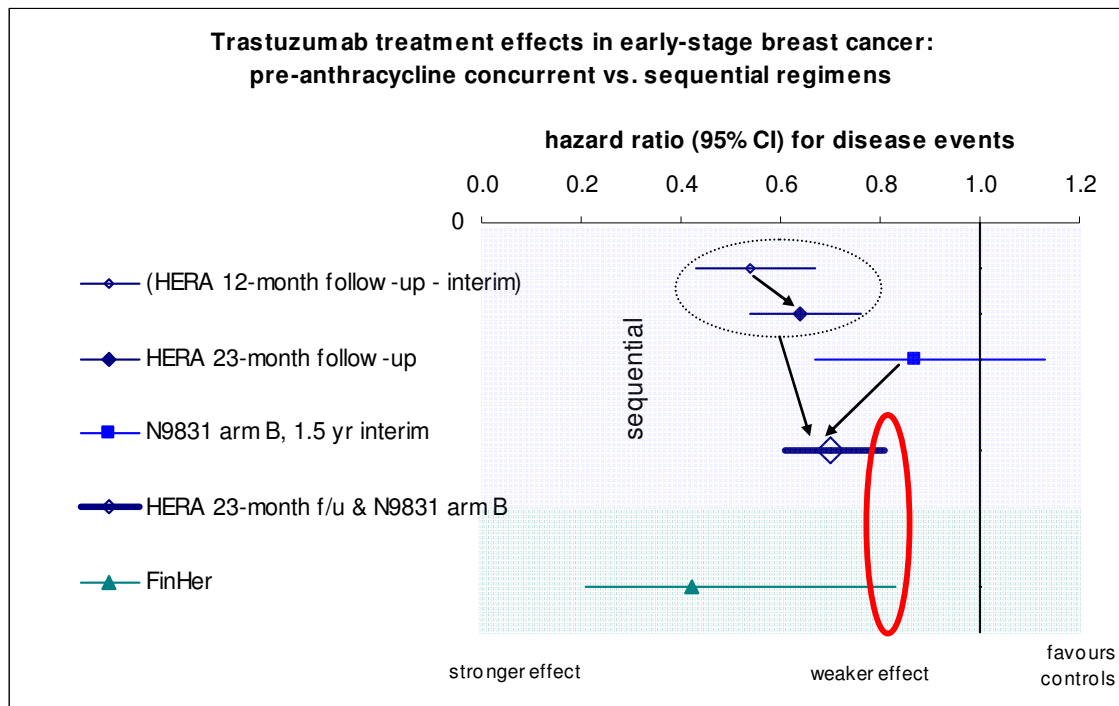


Figure 15



As noted previously, PTAC considered that the number of patients treated in the FinHer study was not insignificant and that the data from the FinHer trial were valuable.

#### Overall survival

As stated above, the FinHer results have not confirmed benefits in terms of statistically significant improvement in overall survival. At three years median follow up, the overall survival (OS) benefit in the trastuzumab treatment arm compared with those that did not receive trastuzumab, although strong (HR 0.41), was not statistically significant (95% CI 0.16-1.08).

However, this non-significant overall survival result may result from the combination of the small sample size and short follow-up at the time of analysis; as noted above, the trial results as reported in the NEJM in 2006 were a pre-planned early interim analysis, pending final analysis at five years median follow-up (or 150 events overall, whichever occurring earlier).<sup>47 xiv</sup>

Given the trend to a strong effect on OS (the 0.41 median HR), it is possible to speculate that this trend in FinHer may have become statistically significant within 3 years and 8 months median follow-up (calculated using binomial methods using constant relative risk of 0.43), i.e. by around January 2004. Although this may be an underestimate (given for instance the improved baseline survival seen in out years elsewhere), it still seems possible that, based on the trend in OS seen in the early efficacy FinHer results, a statistically significant improvement in OS may become evident in the final 5-year median follow-up analysis of FinHer expected later in 2007.<sup>48</sup>

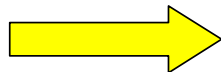
<sup>xiv</sup> Data from final analysis of FinHer may be available from mid-2007 (where the early efficacy analysis (NEJM 2005) was for when median follow-up exceed three years and was able to take place in May 2005) (Joensuu et al NEJM 2006).

In addition, the lack of overall survival benefit to date with FinHer is similar to the early interim analysis of the HERA study at 12 months median follow-up—the basis of New Zealand’s current provisional registration for early stage trastuzumab<sup>18</sup>—which also did not show statistically significant survival improvements.<sup>2</sup>

Studies of trastuzumab in the early breast cancer setting have now repeatedly demonstrated that statistically significant improvements in DFS do correlate with statistically significant improvements in overall survival.<sup>4 5 8</sup> Therefore, it seems reasonable to consider that the statistically significant improvements in DFS seen with the 9 week FinHer regimen could represent a surrogate marker for likely improvements in overall survival.

12 months’ sequential treatment was widely advocated in 2005/06 well before the availability of the HERA 23-month follow-up data that demonstrated statistically significant effects in OS. Instead, the OS effect seen for 12 month concurrent treatment in the Romond joint analysis of trials B31 and N9831 Arm C vs. A (NEJM 2005) was applied to the 12 month sequential setting.<sup>49</sup> However, HERA and the two concurrent trials (Romond) used quite different dosing/sequencing, and only 26% in HERA received taxanes and 6% received no anthracyclines, meaning direct comparison of those results with the Romond trials is problematic.<sup>22</sup> Similarly, the significant effects on overall survival from the joint analysis of trials B31 and N9831 Arm C (Romond) could be applied to the 9 week concurrent regimen. For ease of reference, the relationships between regimens can be seen in the following diagram:

		duration of treatment	
		9 weeks	12 months
sequence of treatment	concurrent with anthracycline	FinHer	N9831 C vs A, B31
	sequential to anthracycline		HERA, N9831 B vs A



= the direction that evidence of statistically significant overall survival in Romond 2005 was extrapolated or could be extrapolated to other settings

Given these observations and assumptions, it therefore seems reasonable to consider that the statistically significant improvements in DFS seen with the 9 week FinHer regimen could represent a surrogate marker for likely improvements in overall survival, whilst awaiting further longer-term results from FinHer.

Data from other clinical trials (E2198)

The results for short-term trastuzumab treatment in FinHer have not been replicated and published in a similar trial setting. However, although significant effects have been reported in only the one trial, data supporting the concept of short-term concurrent trastuzumab treatment is available from a second study. The results of ECOG E2198<sup>10</sup>, a trial of short-duration trastuzumab therapy given concurrently with paclitaxel, were presented as a poster at San Antonio Breast Cancer Symposium (SABCS) 2006. This pilot study, which assessed cardiac safety as the primary outcome, compared short duration trastuzumab (10 weeks) given concurrently with paclitaxel prior to anthracycline treatment, with the same treatment plus an additional 52 weeks trastuzumab after completion of anthracycline treatment. The 5-year follow-up reported similar clinical outcomes from the short duration concurrent regimen as with extended (12 month) trastuzumab treatment.

In detail, E2198 was a small randomized adjuvant trial of adjuvant chemotherapy in early stage breast cancer that examined the cardiac effects of HP (trastuzumab [H] plus 3 weekly administered paclitaxel [P]) followed by AC (doxorubicin [A] plus cyclophosphamide [C]) among 234 breast cancer patients with HER2-positive stage II disease. The patients were randomly allocated to receive either HP175 q3w X 4 followed by AC q3w x 4 (Arm B; HPx3ACx4), or to the same regimen followed by H for 52 weeks (Arm C; HPx3ACx4H for 52 wks). The median follow-up time at reporting was 64 months. Disease-free survival at 5 years was similar for Arms B and C (76% vs. 73%, 32/115 vs. 19/112 respectively, P= 0.55), and there was no difference in overall survival between the study arms (5-year survival was 88% for Arm C (22/115) and 83% for Arm B (12/112), P=0.29). Seven study participants were diagnosed with congestive heart failure (Arm B, n=3; Arm C, n=4), all within 3 years from randomisation. Note that the poster presentation did not describe outcomes for the standard treatment arm (control Arm A).

Although the study did not set out to determine efficacy, and was not powered to determine equivalence, the results lend weight to the case for efficacy of short duration concurrent trastuzumab therapy before anthracycline containing chemotherapy.

The PTAC considered, inter alia, that the results of ECOG E2198 supported the efficacy of short duration concurrent trastuzumab therapy when administered before anthracycline containing chemotherapy, as demonstrated in the FinHer study.

#### 4. The SOLD study

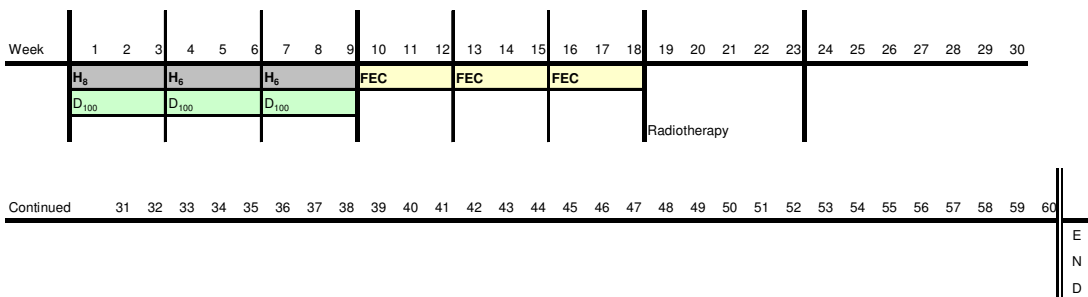
The Synergy or Long Duration (SOLD) study plans to assess the incremental efficacy and risks of adding an extended period of sequential trastuzumab to a short course of concurrent therapy (prior to anthracyclines)<sup>xv</sup>, to help determine optimum treatment length and sequence for the use of trastuzumab in early stage HER2-positive breast cancer. The trial, planned to enroll 3000 patients internationally, will compare two arms:

- nine weeks' trastuzumab, concurrent with chemotherapy (the FinHer treatment regimen), compared with
- the same FinHer treatment regimen plus an additional 42 weeks of trastuzumab (see Figure 16 below).

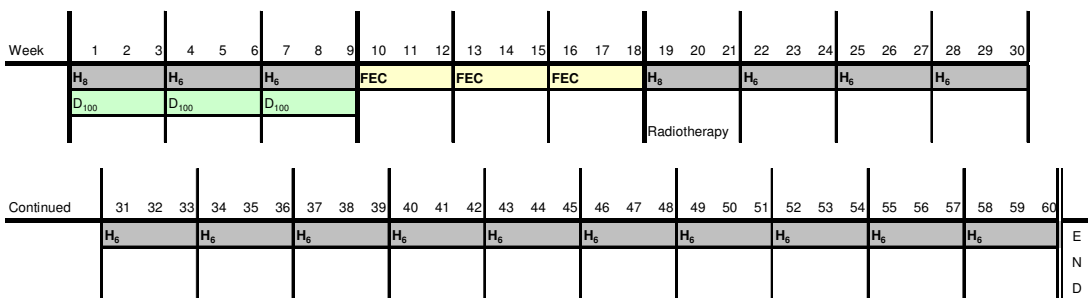
Figure 16. SOLD Study Schema, with detail

#### The Synergy Or Long Duration (SOLD) Study (Arm A)

A randomised phase III study comparing trastuzumab plus docetaxel (HT) followed by 5-FU, epirubicin, and cyclophosphamide (FEC) ( $H_{3\text{-weekly}} T \times 3?FE_{75}C \times 3$ ) to the same regimen followed by single-agent trastuzumab ( $H_{3\text{-weekly}} T \times 3?FE_{75}C \times 3?H_{3\text{-weekly}} \times 14$ ) as adjuvant treatments for breast cancer. Radiotherapy is administered at the close of the studies.



#### The Synergy Or Long Duration (SOLD) Study (Arm B)



**FEC** 5-FU, epirubicin, and cyclophosphamide  
**D<sub>100</sub>** Docetaxel 100mg/m<sup>2</sup>  
**H<sub>8</sub>** Herceptin 8 mg/kg  
**H<sub>6</sub>** Herceptin 6 mg/kg

<sup>xv</sup> The SOLD study proposes two adjuvant regimens with different duration of trastuzumab administration are compared in the treatment of early HER2-positive breast cancer. Trastuzumab will be administered in both arms for 9 weeks in combination with docetaxel to exploit the putative synergism between these drugs and to ensure effective adjuvant therapy to all study participants regardless of the result of random allocation. All patients will also receive 3 cycles of anthracycline-containing chemotherapy, and those who have hormone receptor-positive breast cancer will be treated with adjuvant hormonal therapy for a minimum of 5 years. (Joensuu—SOLD protocol)





## References and endnotes

- <sup>1</sup> RCTs identified by: Belgian Health Care Knowledge Centre (KCE). Trastuzumab in early stage breast cancer. KCE reports vol. 34C, 2006. [http://kce.fgov.be/index\\_en.aspx?ID=0&SGREF=5211&CREF=7198](http://kce.fgov.be/index_en.aspx?ID=0&SGREF=5211&CREF=7198)
- <sup>2</sup> Piccart-Gebhart M.J, Procter M, Leyland-Jones B, et al. Trastuzumab after Adjuvant Chemotherapy in HER2-Positive Breast Cancer. *N Engl J Med* 2005;353(16):1659-1672. <http://content.nejm.org/cgi/content/full/353/16/1659>
- <sup>3</sup> On-line Supplementary Appendix to: Piccart-Gebhart MJ, Procter M, Leyland-Jones B, et al. Trastuzumab after adjuvant chemotherapy in HER2-positive breast cancer. *N Engl J Med* 2005;353:1659-72. <http://content.nejm.org/cgi/data/353/16/1659/DC1/1>
- <sup>4</sup> Smith I, Procter M, Gelber RD, Guillaume S, Feyereislova A, et al. 2 year follow up of trastuzumab after adjuvant chemotherapy in HER2-positive breast cancer: a randomised controlled trial. *Lancet* 2007;369:29-36. <http://www.thelancet.com/journals/lancet/article/PIIS0140673607600282/fulltext>
- <sup>5</sup> Romond, EH, Perez EA, Bryant J, et al. Trastuzumab plus adjuvant chemotherapy for operable HER-2 positive breast cancer. *N Engl J Med* 2005;353(16):1659-1672. <http://content.nejm.org/cgi/content/full/353/16/1673>
- <sup>6</sup> On-line Supplementary Appendix to: Romond EH, Perez EA, Bryant J, et al. Trastuzumab plus adjuvant chemotherapy for operable HER2-positive breast cancer. *N Engl J Med* 2005;353:1673-84. <http://content.nejm.org/cgi/data/353/16/1673/DC1/1>
- <sup>7</sup> Perez EA. Further Analysis of NCCTG-N9831. Slide presentation ASCO annual meeting 2005, available online at [http://www.asco.org/ac/1,1003,12-002511-00\\_18-0034-00\\_19-005815-00\\_21-001,00.asp](http://www.asco.org/ac/1,1003,12-002511-00_18-0034-00_19-005815-00_21-001,00.asp)
- <sup>8</sup> Slamon D, Eiermann W, Robert N, Pienkowski T, Martin M, et al, on behalf of the BCIRG 006 Investigators. Phase III Trial Comparing AC-T with AC-TH and with TCH in the Adjuvant Treatment of HER2 positive Early Breast Cancer Patients: Second Interim Efficacy Analysis. Slide presentation ASCO annual meeting 2006, available online at [http://www.bcirg.org/NR/rdonlyres/eqkdodg2dy7t557o7s6uvj7ytp6g6gfg5gmh2ely6hnh5pjlaz3nd6jddlnao7qoikej3edohsijviisfvp367uuc/BCIRG\\_006+2nd+Interim+Analysis.pdf](http://www.bcirg.org/NR/rdonlyres/eqkdodg2dy7t557o7s6uvj7ytp6g6gfg5gmh2ely6hnh5pjlaz3nd6jddlnao7qoikej3edohsijviisfvp367uuc/BCIRG_006+2nd+Interim+Analysis.pdf)
- <sup>9</sup> Joensuu H, Kellokumpu-Lehtinen PL, Bono P, Alanko T, Kataja V, et al; FinHer Study Investigators. Adjuvant docetaxel or vinorelbine with or without trastuzumab for breast cancer. *N Engl J Med*. 2006;354(8):809-20. <http://content.nejm.org/cgi/content/full/354/8/809>
- <sup>10</sup> Sledge GW, O’Niell A, Thor AD, Kahanic SP, Zander SP, Davidson NE. Adjuvant trastuzumab: long term results of E2198. Poster presentation 2075, SABCs 2006 [http://www.abstracts2view.com/sabcs06/view.php?nu=SABCs06L\\_561](http://www.abstracts2view.com/sabcs06/view.php?nu=SABCs06L_561)
- <sup>11</sup> Numbers of patients in the study arms reported in the HERA trial differ between the first (12 month median follow-up) and second (23 month median follow-up) interim publications of the trial (Piccart-Gebhart et al. *NEJM* 2005, Smith et al. *Lancet* 2007)
- <sup>12</sup> Schulz K F. Assessing allocation concealment and blinding in randomised controlled trials. *Why bother? Evidence Based Medicine* 2000; 5:36-37.
- <sup>13</sup> Jackson R, Ameratunga S, Broad J, Connor J, Lethaby A, Robb G, Wells S, Glasziou P, Heneghan C. The GATE frame: critical appraisal with pictures. *Evid Based Med*. 2006 Apr;11(2):35-8. <http://ebm.bmj.com/cgi/content/full/11/2/35>.
- <sup>14</sup> S Wells. Summary of GATE critical appraisals of trastuzumab trials. EPIQ, University of Auckland, May 2007. Commissioned by PHARMAC. Further critical appraisal of the five relevant clinical trials (HERA, B31, N9831, BCIRG006, and FinHer) using the full Graphic Appraisal Tool for Epidemiology (GATE) framework (<http://ebm.bmj.com/cgi/content/full/11/2/35>), to summarise and assess the strength and quality of all the available relevant RCTs. The analysis was restricted to critically appraising the five individual RCTs as published or otherwise reported; it was not intended to integrate the epidemiological evidence with patient preferences, policy issues or clinical considerations (as occurs with page 4 of the GATE instrument), nor provide a meta-analysis or systematic review or other formal policy advice. EPIQ obtained both specialist clinical epidemiological peer review and specialist oncologist content review for each of the GATE appraisals.
- <sup>15</sup> The GATE frame is a visual framework, designed for critical appraisal, developed by EPIQ. GATE emerged out of the Evidence-Based Medicine (EBM) Working Group users’ guides for appraising for evidence practice (the 28-article series published in *JAMA* by Drs Sackett, Oxman, Guyatt, Cook, Naylor, etc. (<http://www.cche.net/usersguides/main.asp>). The framework graphically represents the generic structure of

---

epidemiological studies and substantively helps systematise critical appraisal in a comprehensive but intuitive way. Details on the GATE framework are available at the EPIQ website ([www.epiq.co.nz](http://www.epiq.co.nz)) and Jackson et al <http://ebm.bmj.com/cgi/content/full/11/2/35>.

<sup>16</sup> Hind D, Pilgrim H, Ward S. Questions about adjuvant trastuzumab still remain. *Lancet* 2007; 369:3-5. <http://www.thelancet.com/journals/lancet/article/PIIS014067360760004X/fulltext>

<sup>17</sup> Hypothetically, all (or almost all) breast cancer cells need to be eradicated for cure. Concomitant administration of the most effective agents available is an obvious strategy to achieve complete eradication of all subclinical cancer. Based on in vitro data, the combination of trastuzumab and docetaxel may be one of the most synergistic ones of all the trastuzumab combinations available to date. These in vitro findings are supported by the clinical data suggesting high activity of the docetaxel plus trastuzumab regimen in the adjuvant, preoperative systemic, and metastatic setting. Long duration of adjuvant administration of single-agent trastuzumab might also result in cancer cell eradication and gradual death of dormant cancer cell populations, although the bulk of evidence suggests that trastuzumab administered in combination with chemotherapy is more effective than trastuzumab given as a single agent. (Joensuu—SOLD protocol)

<sup>18</sup> New Zealand datasheet (Medsafe): Herceptin® Trastuzumab 150 mg and 440 mg powder for concentrate for solution for infusion, <http://www.medsafe.govt.nz/profs/Datasheet/h/Herceptininf.htm>

<sup>19</sup> Australian datasheet: HERCEPTIN® PRODUCT INFORMATION, dated 21 April 2006. <http://www.pbs.gov.au/html/healthpro/search/results?term=herceptin&publication=&form-type=simple>

<sup>20</sup> US datasheet: Initial US Approval: September 1998, Revision Date: November 2006. <http://www.gene.com/gene/products/information/oncology/herceptin/insert.jsp>

<sup>21</sup> New Zealand datasheet (Medsafe): Herceptin® Trastuzumab 150 mg and 440 mg powder for concentrate for solution for infusion, <http://www.medsafe.govt.nz/profs/Datasheet/h/Herceptininf.htm>

<sup>22</sup> Herceptin and early breast cancer: a moment for caution. *Lancet*. 2005 Nov 12;366(9498):1673. <http://www.thelancet.com/journals/lancet/article/PIIS0140673605676702/fulltext>

<sup>23</sup> 'Although the [Romond 2005] "joint analysis was developed and analysed" by both trial teams, it is unclear whether this synthesis was planned in advance of the start of both trials. The report merely notes that the US FDA and National Cancer Institute approved the joint analysis plan, which may reflect the expectation that neither trial alone would demonstrate a positive result' (*Lancet* 2005).

<sup>24</sup> Geddes J, Szatmari P, Streiner D. The worm turns: publication bias and trial registers revisited. *Evid Based Ment Health*. 2004 Nov;7(4):98-9. <http://ebmh.bmj.com/cgi/content/full/7/4/98>

<sup>25</sup> Chalmers I. Underreporting research is scientific misconduct. *JAMA* 1990;263:1405-1408

<sup>26</sup> Antes G, Chalmers I. Under-reporting of clinical trials is unethical. *Lancet* 2003;361:978-979.

<sup>27</sup> HERA's previous 12-month median f/u results for overall survival (Piccart-Gebhart et al 2005) did not reach statistical significance. Advocates for the HERA regimen (Hortobagyi 2005) relied instead on extrapolating from the Romond overall survival results.

<sup>28</sup> PHARMAC analysis of the two HERA publications (Piccart-Gebhart 2005, Smith 2007). The 23-month median HERA f/u (Smith 2007) comprises cumulative hazards over the entire 23 months, with 218+321=539 first disease recurrence events – in turn comprising events and hazards reported as accruing by the first 12 months median f/u (Piccart-Gebhart) (127+220=347 events) and then the complement of events for the remaining time period (539-347=192), i.e. events occurring in the time since the first analysis. Analysis therefore involves subtracting events recorded as accruing in the first time period (12 months median follow-up) from the cumulative events recorded as accruing over the whole time period analysed to date (23 month median follow-up). This derives numbers of events occurring in the time since the first analysis, in order to calculate period-specific rates and relative risks. In turn this allows testing for statistical interaction (comparing the two time periods) using standard binomial methods:

<b>HERA - 23-month vs. 12 month median follow-ups</b>					difference	calculated relative risk (RR)			RR reduction (RRR)	
numbers				i.e. tmt effect (ARR)		published RR HR (95% CI)	-95% CI	+95% CI		
	trastuzum ab	std tmt	total							
<b>Input data:</b>										
12 mth f/u	n events	127	220	347						
0-12 mths	N patients	1694	1693	3387						
	%	7.5%	13.0%		5.5%	0.54 (0.43-0.67)				
23 mth f/u	n events	218	321	539						
0-23 mths	N patients	1703	1698	3401						
cumulative	%	12.8%	18.9%		6.1%	0.64 (0.54-0.76)				
<b>calculations</b>										
0-12 mths	n events	127	220	347						
	N patients	1694	1693	3387						
	%	7.5%	13.0%		5.5%	0.54	0.58	0.47	0.71	42.3%
	std error of %	0.6%	0.8%		0.8%					
13-23 mths	n events	91	101	192						
=23mth f/u minus 12m	N patients	1703	1698	3401						
	%	5.3%	5.9%		0.6%	0.90	0.68	1.18	10.2%	
	std error of %	0.5%	0.6%		0.3%					
	difference in treatment effects				4.9%					
	approximate 95% CI for difference					4.9%	3.3%	6.5%		
	standard error for difference in treatment effects				0.8%					
	ratio of difference/SE				5.915					
	p-value				0.02	+ve evidence of heterogeneity				

<sup>29</sup> The N9831 Arm C vs. Arm B data await further analysis and publication, as was accorded the Arm C vs. Arm A comparison when this was combined with data from another study in the Romond 2005 joint analysis publication.

<sup>30</sup> Belgian Health Care Knowledge Centre (KCE). Trastuzumab in early stage breast cancer. KCE reports vol. 34C, 2006. [http://kce.fgov.be/index\\_en.aspx?ID=0&SGREF=5211&CREF=7198](http://kce.fgov.be/index_en.aspx?ID=0&SGREF=5211&CREF=7198), pp 51, 72.

<sup>31</sup> Testing for proportional hazards over time will be to determine whether the effect of treatment is constant or varies significantly over time.

<sup>32</sup> Matthews JNS, Altman DG. Interaction 3: How to examine heterogeneity. *BMJ* 1996;313:862. <http://www.bmj.com/cgi/content/full/313/7061/862>.

<sup>33</sup> Altman DG, Bland JM. Interaction revisited: the difference between two estimates. *BMJ* 2003;326:219. <http://www.bmj.com/cgi/content/full/326/7382/219>.

<sup>34</sup> PHARMAC analysis of the two conference slide presentations for BCIRG 006, subtracting events recorded as accruing in the first time period (23 months median follow-up) from the cumulative events recorded as accruing over the whole time period analysed to date (36 month median follow-up) to derive numbers of events occurring in the time since the first analysis, then testing for interaction using standard binomial methods (Matthews & Altman 1996, Altman & Bland 2003):

BCIRG 006 concurrent anthracyclines - 36-month vs. 23 month median follow-ups										
		numbers			difference		calculated relative risk (RR)			RR reduction (RRR)
		trastuzum ab	std tmt	total	i.e. tmt effect (ARR)	published RR HR (95% CI)	-95% CI	+95% CI		
<b>Input data:</b>										
23 mth f/u	n events	77	147	224						
0-23 mths	N patients	1074	1073	2147						
	%	7.2%	13.7%		6.5%	0.49				
36 mth f/u	n events	128	192	320						
0-36 mths	N patients	1074	1073	2147						
cumulative	%	11.9%	17.9%		6.0%	0.61				
<b>calculations</b>										
0-12 mths	n events	77	147	224						
	N patients	1074	1073	2147						
	%	7.2%	13.7%		6.5%	0.49	0.52	0.40	0.68	47.7%
	std error of %	0.8%	1.0%		1.1%					
24-36 mths	n events	51	45	96						
=36mth f/u	N patients	1074	1073	2147						
minus 23m	%	4.7%	4.2%		-0.6%		1.13	0.77	1.68	-13.2%
	std error of %	0.6%	0.6%		0.3%	(uses absolute value (+0.6%) for ARR (-0.6%))				
difference in treatment effects					7.1%					
approximate 95% CI for difference							7.1%	4.9%	9.3%	
standard error for difference in treatment effects					1.1%					
ratio of difference/SE					6.363					
p-value					0.01	+ve evidence of heterogeneity				

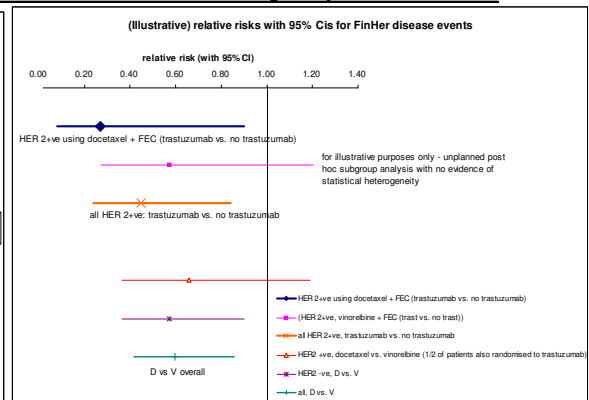
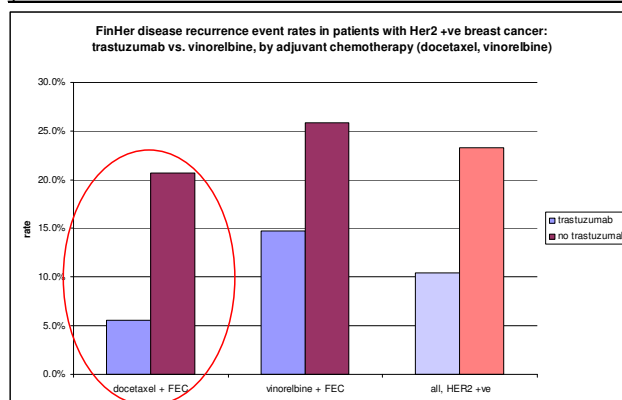
<sup>35</sup> The FinHer study was designed to have a power of 0.80 to detect an increase in five-year recurrence-free survival from 70 percent to 80 percent in the docetaxel-plus-FEC group as compared with the vinorelbine-plus-FEC group (with use of a two-sided test at a significance level of 0.05); approximately 150 events were required for this purpose. The study's designers estimated at the time that 30% of the participants would have breast cancer with HER2/neu amplification and that the study would be able to detect a difference in their five-year recurrence-free survival of 50 percent to 67 percent at a power of 0.80 when approximately 1000 patients were enrolled. Protocol-defined safety analyses took place in March 2001, September 2001, and December 2002. The protocol specified that safety and early efficacy analyses were to be carried out when the median follow-up time exceeded three years; this point was reached in May 2005, hence the NEJM paper published in February 2005. The final analysis is scheduled to be performed when 150 events have occurred or the median follow-up time exceeds five years. For the primary variable, a p-value of less than 0.029 was considered to indicate significance, in order to maintain an overall type 1 error of 0.05 for the interim and final analysis.

<sup>36</sup> Hypothetically, all (or almost all) breast cancer cells need to be eradicated for cure. Concomitant administration of the most effective agents available is an obvious strategy to achieve complete eradication of all subclinical cancer. Based on in vitro data, the combination of trastuzumab and docetaxel may be one of the most synergistic ones of all the trastuzumab combinations available to date. These in vitro findings are supported by the clinical data suggesting high activity of the docetaxel plus trastuzumab regimen in the adjuvant, preoperative systemic, and metastatic setting. Long duration of adjuvant administration of single-agent trastuzumab might also result in cancer cell eradication and gradual death of dormant cancer cell populations, although the bulk of evidence suggests that trastuzumab administered in combination with chemotherapy is more effective than trastuzumab given as a single agent. (Joensuu—SOLD protocol)

<sup>37</sup> PHARMAC analysis of the FinHer publication:

**FinHer HER2 +ve pts - docetaxel +/- trastuzumab vs vinorelbine +/- trastuzumab**

		numbers		total	difference	calculated relative risk (RR)			RR reduction (RRR)
		trastuzum ab	no trastuzum ab		i.e. tmt effect (ARR)	RR	-95% CI	+95% CI	
docetaxel + FEC	n events	3	12	15					
	N patients	54	58	112					
	%	5.6%	20.7%		15.1%	0.27	0.08	0.90	73.1%
	std error of %	3.1%	5.3%		6.8%				
vinorelbine + FEC	n events	9	15	24					
	N patients	61	58	119					
	%	14.8%	25.9%		11.1%	0.57	0.27	1.20	43.0%
	std error of %	4.5%	5.7%		5.8%				
difference in treatment effects					4.0%				
approximate 95% CI for difference						4.0%	-13.4%	21.5%	
standard error for difference in treatment effects					8.9%				
ratio of difference/SE					0.453				
p-value					0.50	no evidence of heterogeneity			



<sup>38</sup> Pegram MD, Lopez A, Konecny G, Slamon DJ. Trastuzumab and chemotherapeutics: drug interactions and synergies. *Semin Oncol.* 2000;27(6 Suppl 11):21-5

<sup>39</sup> Rothman KJ. *Modern epidemiology.* 1st edition. Boston, MA: Little, Brown, 1986

<sup>40</sup> By contrast, in order for HERA to achieve its HR of 0.64 at the 23 month median follow-up, a minimum of around 560-570 patients would have been required. HERA did so with its 3,401 patients, because it was powered to detect changes in overall survival, hence its narrower CI for DFS, albeit with lessened efficacy for point estimates than seen with FinHer.

<sup>41</sup> Schulz KF, Grimes DA. Sample size calculations in randomised trials: mandatory and mystical. *Lancet.* 2005 Apr 9;365(9467):1348-53. <http://www.thelancet.com/journals/lancet/article/PIIS0140673605610343/fulltext>

<sup>42</sup> Sackett DL, Cook DJ. Can we learn anything from small trials? *Ann N Y Acad Sci.* 1993 Dec 31;703:25-31; discussion 31-2.

<sup>43</sup> Altman DG, Bland JM. Absence of evidence is not evidence of absence. *BMJ.* 1995 Aug 19;311(7003):485. <http://www.bmj.com/cgi/content/full/311/7003/485>

<sup>44</sup> Concerns about FinHer's small number of patients seem to relate to the power (1-β) of the study. Power is the probability of detecting a statistically significant difference when a difference of a given magnitude really exists (analogous to sensitivity or 'true-positives'). Power derives from β error, type II error (β) being the probability of not detecting a statistically significant difference when a difference of a given magnitude in reality exists—i.e. the chance of a false-negative result; mathematically, power is the complement of β (1-β) and represents the probability of avoiding a false-negative conclusion. However, rather than just power, sample sizes depend on four components, being

type I error ( $\alpha$ ), power ( $1-\beta$ ), the event rate in the control group, and the treatment effect (or analogously the event rate in the treatment group) (Schulz & Grimes 2005). Some small trials are so definitively positive that they are sufficient to identify the best therapy (Sackett & Cook DJ 1993).

It has been observed that preoccupation with sample size overshadows the more pertinent concerns of elimination of bias, and that unbiased trials with imprecise results trump no results at all (Schulz & Grimes 2005): “Trials should be methodologically strong, thus eliminating bias. Unfortunately, the adequate-power mantra frequently overwhelms discussion on other methodological aspects, e.g. inadequate randomisation usually yields biased results which cannot be salvaged even if a huge sample size generates great precision. By contrast, if investigators design and implement a trial properly, that trial essentially yields an unbiased estimate of effect, even if it has lower power (and precision).”

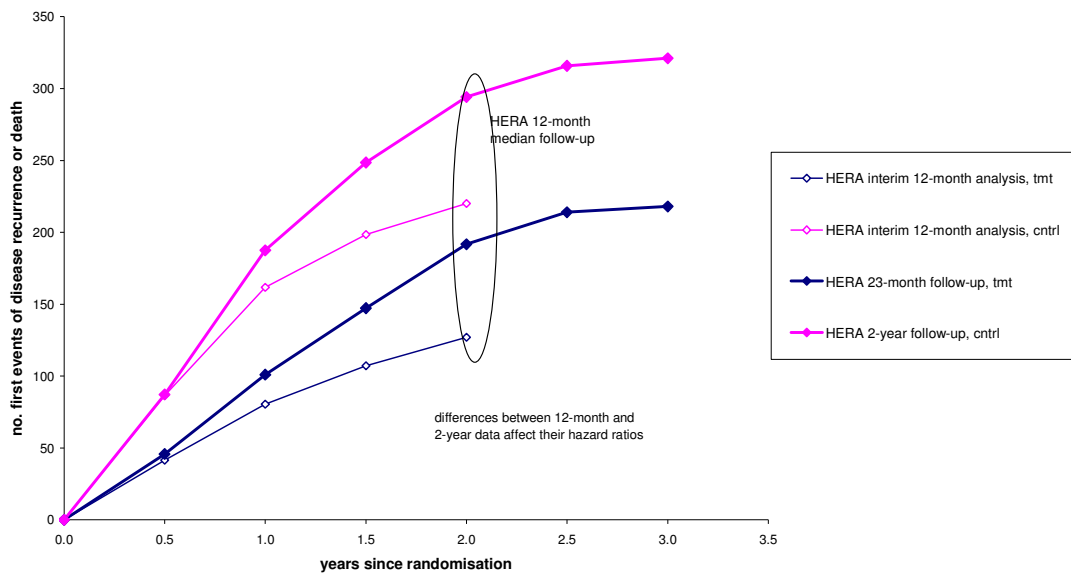
<sup>45</sup> Yoshioka A. Use of randomisation in the Medical Research Council's clinical trial of streptomycin in pulmonary tuberculosis in the 1940s. *BMJ*. 1998 Oct 31;317(7167):1220-3. <http://www.bmj.com/cgi/content/full/317/7167/>

<sup>46</sup> Roche media release 3 May 2006. Herceptin approved in Europe for use in combination with an aromatase inhibitor for the treatment of patients with HER2 and hormone receptor-co-positive metastatic breast cancer. <http://www.roche.com/med-cor-2007-05-03>

<sup>47</sup> Data from final analysis may be available from mid-2007 (where the early efficacy analysis (NEJM 2005) was for when median follow-up exceed three years and was able to take place in May 2005) (Joensuu et al NEJM 2006).

<sup>48</sup> The impact of the timing of analysis and hence the potential for extra events to accrue and hence improve precision can be seen by examination of the 23 month HERA results compared with the 12 month HERA results (survival analysis using published hazards (Smith 2007) and numbers of patients at risk (Piccart-Gebhart 2005, Smith 2007)). Similar patterns of increased events would be expected with the FinHer final analysis—both for DFS and overall survival.

**Disease events HERA: 12-month interim vs. 23-month median follow-up data**  
 survival analysis using published hazards (Smith 2007) and nos. patients at risk (Piccart-Gebhart 2005, Smith 2007)



<sup>49</sup> Hortobagyi GN. Trastuzumab in the treatment of breast cancer. *N Engl J Med*. 2005 Oct 20;353(16):1734-6.