Using serum CA125 to assess the activity of potential cytostatic agents in ovarian cancer

M. R. Hall [1], A. Petruckevitch[1], J. Pascoe [1], M. Persic [2], S. Tahir[3], J. S. Morgan[4], C. Gourley[5], N. Stuart[6], S. M. Crawford[7], D. E. Kornbrot[8], Qian W[9] and G. J. Rustin[1]

Mount Vernon Cancer Centre[1], Northwood, United Kingdom; Derby Hospitals NHS Foundation Trust[2] Derby, United Kingdom; Mid Essex Hospital Service NHS Trust[3], Chelmsford, United Kingdom; Ipswich Hospital NHS Trust[4], Ipswich, United Kingdom; University of Edinburgh Cancer Research UK Centre[5], MRC IGMM, Western General Hospital, Edinburgh, United Kingdom; Ysbyty Gwynedd[6], Bangor, United Kingdom; Airedale General Hospital[7], Keighley, West Yorkshire, United Kingdom; University of Hertfordshire[8], Hatfield, United Kingdom; Cambridge Cancer Trials Centre[9], Addenbrookes Hospital, Cambridge, United Kingdom.

Corresponding author	Dr Marcia Hall
Address	Mount Vernon Cancer Centre, Rickmansworth Road, Northwood, Middlesex HA6 2RN
Telephone	01923-844190
Fax	01923-844840
Email	marcia.hall@nhs.net

Funding: This work was supported by the Cancer Treatment and Research Trust and the National Cancer Research Institute.

Acknowledgements:
We would like to thank the following investigators for contributing patients: Dr D E Parkin, Aberdeen Royal Infirmary; Dr N Reed and Dr R Glasspool, Beatson Cancer Centre, Glasgow; Dr J A Green, Clatterbridge Centre for Oncology; Dr A Lamont, Colchester; Dr S Pearson, Cumberland Infirmary, Carlisle; Dr A Horne Great Western Hospital, Swindon; Dr A Montes and Dr A Winship, Guy's Hospital London; Dr S Sundar, Kingston Hospital, Sutton; Dr O McNally, Musgrove Hospital Taunton; Dr A Hughes Queen Elizabeth Hospital Gateshead; Dr S Essapen, Royal Surrey County Hospital, Guildford; Dr Melanie Harvey, Salisbury District Hospital; Dr C Green, Southampton Hospital and Dr Tim Perren, St James' University Hospital, Leeds.

Abstract

Objective: New strategies are required to rapidly identify novel cytostatic agents before embarking on large randomised trials. This study investigates whether a change in rate of rise (slope) of serum CA125 from before to after starting a novel agent could be used to identify cytostatic agents. Tamoxifen was used to validate this hypothesis.

Methods: Asymptomatic patients with relapsed ovarian cancer who had responded to chemotherapy were enrolled and had CA125 measurements taken every 4 weeks, then more frequently when rising. Once levels reached 4 times upper limit of normal or nadir, they started continuous tamoxifen 20 mg daily, as well as fortnightly CA125 measurements until symptomatic progression. Due to the potentially non-linear relationship of CA125 over time it was felt that to enable normal approximations to be utilised a natural logarithmic standard transformation [ln(CA125)] was the most suitable to improve linearity above the common logarithmic transformation to base 10.

Results: From 235 recruited patients, 81 started tamoxifen and had at least 4 CA125 measurements taken before and 4 CA125 measurements taken after starting tamoxifen, respectively. The

mean regression slopes from using at least 4 ln(CA125) measurements immediately before and after starting tamoxifen were 0·0149 and 0·0093 [ln(CA125)/day] respectively. This difference is statistically significant, P = 0·001. Therefore in a future trial with a novel agent, at least as effective as tamoxifen, using this effect size, the number of evaluable patients needed, at significance level of 5% and power of 80%, is 56.

Conclusions: Further validation of this methodology is required but there is potential to use comparison of mean regression slopes of ln(CA125) as an interim analysis measure of efficacy for novel cytostatic agents in relapsed ovarian cancer.

Introduction

The activity of cytotoxic chemotherapy is conventionally assessed by clinical and radiological measurement of change in size of tumour masses. Internationally standardised radiological criteria (RECIST) are used to assess response[1]. However, many new agents with potential anti-cancer activity are predicted to be cytostatic rather than cytotoxic. These approaches may lead to stabilisation of tumour growth rather than a rapid shrinkage. The danger of using radiological response criteria to assess the clinical utility of these agents is that active agents are not always detected in phase II clinical trials. For example in patients with advanced gastrointestinal stromal tumour (GIST), responses are not always accompanied by reductions in tumour size[2]. New approaches are therefore required to assess the activity of cytostatic agents in phase II clinical trials to minimize the chance of active cytostatic agents being discarded. One way to prove a drug has cytostatic activity would be to show that tumour growth in individual patients was slower after starting on a novel agent than before. In most types of cancer there is not a reasonable time frame where it is acceptable to observe tumour growth without offering therapy. However, in ovary cancer, most patients have rising levels of the serum tumour marker, CA125, for a

period whilst remaining asymptomatic. There is an average lead-time of four months between the initial rise of the CA125 and symptomatic relapse of ovarian cancer, generally requiring treatment. It has been shown that it is preferable to delay chemotherapy treatment in such relapsing patients until they develop symptoms[3]. We postulated that this would be a potential group of patients in whom the rate of change of CA125 levels, before and after initiation of a novel drug could be an alternative method of identifying activity worthy of trial continuation.

The first stage of this project was to determine whether the rate of rise of CA125 is consistent enough to allow comparison of (rate of rise) slopes before and after the introduction of a treatment. This required sufficient CA125 measurements from patients on no therapy or placebo. A pilot study ascertained that most patients with relapsed ovarian cancer would not accept intense CA125 measurements, if there were no prospect of some therapy once the CA125 levels started rising. Although such asymptomatic patients with rising serum CA125 levels remain anxious about recommencing treatment, many accept a less toxic therapy provided they know that they will be closely monitored and able to start cytotoxic treatment again as soon as they become symptomatic.

Tamoxifen has been shown to have activity in approximately 15% of patients with relapsed ovarian cancer[4] and is an ideal agent with minimal toxicity. The tolerability of tamoxifen and it's modest activity in ovarian cancer patients (akin to placebo in most), would allow us to determine any variability in rate of rise of CA125, as well as the chance of detecting some change in the 15% of patients where it is active. Figure 1 shows graphically an outline of what was expected to occur with respect to the CA125 measured over time for such patients and a range of potential outcomes after the patients start taking tamoxifen.

Methods

We prospectively registered women between May 2005 and November 2009 with histologically confirmed *recurrent* epithelial ovarian, fallopian tube or primary peritoneal cancer who had completed and responded to treatment for first relapse (i.e second course of chemotherapy, NOT after primary adjuvant / neoadjuvant therapy), with a fall in their CA125 of at least 50% by the end of treatment. A protocol amendment in April 2009 allowed the registration of patients who had responded to 2nd, 3rd and 4th line relapse therapy as long as their CA125 had fallen by 50% by the end

of treatment. No measurable disease was required for trial registration. Patients were ECOG performance status 0-2 and had no symptoms of ovarian cancer requiring chemotherapy. None of the patients were taking any hormone therapy and all other concomitant medications were at a stable dose throughout the trial period. Adequate end organ function was required in all patients including: Hb >10g/dl, WCC > 2.5 x 10^9/l, plts >100 x 10^9/l, creatinine, AST and or ALT < 2 x upper limit of normal, and bilirubin < 1.5x upper limit normal.

For each registered patient a baseline or nadir CA125 was documented, defined as the lowest CA125 estimate, obtained at least 3 weeks after completing chemotherapy for relapse. Once registered, patients had serum CA125 concentrations measured monthly, with all serial assays of an individual being performed in the same laboratory. If a patient's CA125 level started rising, the frequency of the CA125 levels was increased to fortnightly. When (if) the CA125 level reached four times the upper limit of normal or the patient's registered lowest nadir level (if higher than normal), the patient was contacted, imaging of thorax, abdomen and pelvis was performed and they were started on tamoxifen, 20mg daily. Provided these patients

remained stable and asymptomatic whilst taking tamoxifen, CA125 measurements were continued fortnightly.

CT scans were performed every three months in accordance with local practice but responses by RECIST criteria were not expected in this study. Radiological assessment was predominantly for safety purposes in the unlikely event that women have significantly progressive disease without a rise in CA 125. Similarly, if at any point, a registered patient became symptomatic from their ovarian cancer, tamoxifen and study related CA125 measurements were stopped, they were withdrawn from the study and treated in accordance with local practice. Asymptomatic patients with a rising CA125 level whilst taking tamoxifen were allowed to continue on study until they became symptomatic.

The ethics committee approved the protocol and the study was conducted in accordance with the Principles of Good Clinical Practice and the Declaration of Helskini. All patients provided written informed consent.

Statistical Methods

CA125 values are not naturally normally distributed so a log transformation was used to enable linear associations to be examined. Linear regression lines were fitted to each patients data to assess any increase or decrease in ln(CA125) levels over the study duration, before and after treatment. It was agreed that at least 4 CA125 measurements were required to fit these linear regression lines and statistically assess the potential rise/fall of a patient's CA125 levels before and after treatment with tamoxifen. All patients were included until clinical progression to maximise evaluability and sample size.

The percentage of patients with a log linear rise was calculated to ensure that sufficient numbers existed to be able to examine the association of log linear rise before and after starting tamoxifen. The consistency of the log linear slopes was then examined to determine whether comparison before and after introduction of tamoxifen was possible. The individual slopes obtained were then compared using a paired t-test.

It could be argued that examining the change in slopes of CA125 levels over time, before and after starting tamoxifen could be misleading due to the possibility of regression to the mean. We therefore also examined the ln(CA125) slopes of other registered patients who were on trial and had four CA125 measurements before starting or without starting tamoxifen – usually because they became too symptomatic requiring immediate chemotherapy - to act as a comparator group.

Finally the number of patients required to detect a difference in ln(CA125) slopes, before and after starting tamoxifen, was calculated. This was to determine whether using a comparison of ln(CA125) slopes before and after introduction of a novel agent might allow more rapid decisions concerning the efficacy of these drugs, particularly if they are not expected to show actual tumour shrinkage measurable by RECIST.

Results

235 women were recruited to the study and provided at least one trial CA125 sample. 204 patients were recruited after their first relapse therapy. The mean (SD) age in years at the time of consenting for study registration was 64.4 (11.0).

Fifty-three percent of the study group received tamoxifen (125/235), but of these only 81 patients had at least four CA125 measurements collected before **and** after the introduction of tamoxifen. The remaining 44 patients had fewer than four CA125 measurements, either before or after, or both and were excluded from log linear slope analysis.

110 patients (47%) in the study never started tamoxifen, because they required chemotherapy treatment for their symptomatic ovarian cancer before this. Only 74 of these patients had at least four CA125 measurements before coming off study, 36 patients had fewer than four CA125 measurements prior to requiring chemotherapy and coming off study.

The mean (SD) baseline (nadir) CA125 value for the 81 patients who had at least four CA125 measurements before and after starting tamoxifen was 61.5 (106.4). Table 1 shows the average slope over time for the 81 evaluable patients before and after receiving tamoxifen. Figure 2 illustrates graphically CA125 plotted over time for these 81 patients.

The mean slopes before and after treatment both significantly increased over time in this group. A paired t-test revealed that there was a significant difference in the rates of rise across patients over time before and after treatment (Table 2), indicating that taking tamoxifen slowed the rise in CA125 levels. Figure 3 represents this graphically in a scatter plot.

14 percent (11/81) of the patient group who received tamoxifen showed a registered decline in their CA125 levels after treatment with tamoxifen. 4 patients had a 50% decrease in their CA125 after starting tamoxifen, confirmed and maintained for at least 28 days or more, thus fulfilling the GCIG criteria of CA125 response[5].

CA125 variation and regression to the mean

In order to explore the effect of regression to the mean in individual patients, possibly accounting for the slowing of the rise of CA125 after starting tamoxifen, we examined the CA125 levels in patients who did not start tamoxifen at all. These patients are likely to have had faster growing tumours than those asymptomatic and able to start tamoxifen. If regression to the mean was a significant factor, it

would occur in all groups of patients where CA125 is measured over time. The similarity of the mean slope of 0.0131 for the non tamoxifen treated group versus 0.0149 in the patients before starting tamoxifen suggests that regression to the mean is not having an impact. It is reassuring to note that the mean change in rise in slope (Table 3) of 0.0056 in those treated with tamoxifen is significantly different.

Seventy-four patients had at least 4 CA125 measurements and Table 3 shows the mean slope derived from fitting a regression line to each patient's measurements after recruitment. Comparing the mean slopes between this group and the pre-tamoxifen group showed a non-significant difference between them.

Sample size for future trials

Using the mean change in slope for tamoxifen described above, the number of evaluable patients required to detect an effect of a similar size with a novel agent is 74 and 56 with a power of 90% and 80%, respectively. However, only 81 out of 235 (34%) patients were statistically evaluable therefore 218 and 165 patients respectively would have to be registered and start monthly CA125 measurements.

Although there may be an issue with bias due to the reduced number of evaluable patients in this analysis, attempting to ascertain a patient's CA125 slope based on only 2 observations pre-treatment would not be a statistically robust approach and could be potentially misleading in terms of any assumed linear relationship, the design range of evaluable values per patient were explored and made no difference to the results per se. Using four observations before and four observations after treatment were thought adequate to establish any log linear relationship. Using alternative more complex methods led to identical conclusions on a subset of these patients[6].

Discussion

Raw CA125 measurements are not normally distributed; however, this study has shown that log transforming the CA125 levels, results in data enabling traditional linear regression methods to be used. We have shown that patients have log linear rises in CA125 levels after treatment for relapsed ovarian cancer and that these log linear rises are consistent enough for comparison before and after therapy. In this group of 81 tamoxifen treated patients, there is a significant difference ($P = 0.001$) in the mean of the slopes of the ln(CA125) in patients before and after treatment with tamoxifen. The fact that a

significance level of P<0.001 was reached with a reduced sample of 81 does not in any way jeopardise the power of this design, potentially precision may have been improved with an increased evaluability rate. The proportion of patients with a decreasing slope (14%) after treatment with tamoxifen is the same as published data which suggests that between 13 and 17% patients have a RECIST response when treated with tamoxifen[7, 8].

The description of the rise/fall of ln(CA125) over time in response to the introduction of tamoxifen suggests that this approach would be feasible in assessing the activity of novel agents. However, the nature of this "before and after" design of study, without the use of a separate control group, in attempting to quantify the true variability of patients is not without pitfalls due to potential regression to the mean[9]. We have attempted to examine this by comparing the ln(CA125) slopes of patients in the group who did NOT receive tamoxifen with the "before" ln(CA125) slope of those in the treated group. The hypothesis is that if regression to the mean is a significant factor, you would expect to see variability and different slopes in both the untreated group and the pre-treatment group. As we have shown that there is no significant difference in the mean of

these slopes, it is unlikely that the real change observed in the 81 patients before and after the start of tamoxifen was due to regression to the mean.

The results of this study have enabled us to calculate the number of patients that may be required to detect a significant difference in ln(CA125) slopes to potentially assess other novel agents in the same way. 74 and 56 evaluable patients are required for 90 and 80% power respectively, this would require recruitment of 218 or 165 patients and is very similar to the numbers required to assess novel agents using well described randomized phase II trial methods using RECIST criteria. However, these methods have been developed to assess cytotoxic agents where tumour shrinkage is readily apparent. By contrast, this method may be useful, with drugs that are likely to be cytostatic and where actual RECIST responses are not expected. Currently the only way to ascertain whether a novel agent is active in controlling tumor growth (cytostatic) in man is to do an initial randomized phase 2 trial with progression free survival as the primary endpoint. These trials require the trial drug to be given to far more than 56 patients to give a go / no go decision as to whether further development is justified.

To demonstrate how this approach could also be used in the initial selection of potentially toxic cytostatic drugs, application of the change in slope ln(CA125) methodology was theoretically applied to the randomised tamoxifen versus thalidomide trial recently reported[10]. In this trial, a target recruitment of 260 ovarian cancer patients was planned to provide a 90% chance of proving thalidomide active, however at the first interim analysis, 4.5 years after trial activation, the study was halted because the PFS and OS rates were 38% and 39% higher respectively in the *tamoxifen* arm. Using the change in ln(CA125) slope methodology, a single arm study of thalidomide, recruiting 100 patients and treating only half of these for 8-12 weeks would have saved a significant number of patients exposure to the toxicities of thalidomide to prove that it was inactive. To further reduce exposure to potentially toxic, ineffective drugs, it might be useful to explore using a doubling rather than a quadrupling of nadir CA125 and to include patients with persistently rising levels within the upper limit of normal[11].

An additional application of this methodology may be in assessment of slow release agents such as pegylated liposomal doxorubicin,

which may result in an initial rise in CA125 (reflecting initial disease progression) because of the slow release nature of therapy, prior to 'response' [12]. Utilising changes in ln(CA125) slopes, asymptomatic patients on these agents could continue on the study agent, provided that there is a reduction in the rate of rise of CA125, despite it not reaching "response" by standard GCIG CA125 criteria. Disadvantages of this methodology, aside from the disappointingly high number of registered patients required for an evaluable subset, include the possibility that drugs altering CA125 shedding might result in spurious results. *In vitro* studies will be required to exclude this possibility.

Although there is a significant difference (P = 0.001) in the mean of the slopes of the ln(CA125) in patients before and after treatment with tamoxifen, this appears to occur early after starting tamoxifen, and in most patients it is short-lived; the number of patients where the CA125 response was maintained was only 15%, identical to that described in the general literature[4]. So this method might also paradoxically alert one to drugs that only have a fleeting impact on ovarian cancer. It is thus important to use this method alongside standard assessments of response and progression free survival.

Retrospective validation of this method has not been possible as there are no large data sets where CA125 measurements were performed every two weeks for a reasonable period *prior* to starting a trial drug, as in standard practice CA125 is generally measured 2-3 monthly. However this work has shown proof of concept that utilizing the change in rate of rise of CA125 can be an early indicator of activity of new agents, particularly at an interim analysis, where a decision may be required as to whether to continue or stop early phase studies in the absence of objective RECIST responses. Prospective validation will require that agents shown to be active by changing the CA125 doubling rate are also shown to prolong PFS / OS and agents that fail to decrease the CA125 doubling rate also fail to prolong PFS.

References

1. Eisenhauer EA, Therasse P, Bogaerts J, et al. New response evaluation criteria in tumors: revised RECIST guideline (version 1.1). *Eur J Cancer*. 2009;45(2):228-47.

2. Desai J. Response assessment in gastrointestinal stromal tumour. *Int J Cancer* 2011;128(6):1251-8.

3. Rustin GJ, van der Burg M, Griffin CL, et al. Early versus delayed treatment of relapsed ovarian cancer (MRC OV05/EORTC 55955): a randomised trial. . *Lancet* 2010;376(1155-63.).

4. Williams CJ, Simera I, Bryant A. Tamoxifen for relapse of ovarian cancer. *Cochrane Review Issue 2*, 2003

5. Rustin GJ, Vergote I, Eisenhauer E, et al. Definitions for response and progression in ovarian cancer clinical trials incorporating RECIST 1.1 and CA125 agreed by the Gynecological Cancer Intergroup (GCIG). Int J Gynecol Cancer. 2011 Feb;21(2):419-23.

6. Qian W, Hall M, Rustin G, & Kornbrot D. Detection of disease change from a biological marker: using CA125 in ovarian cancer as an example. *Paper presented at the RSS 2012 International Conference, Telford, UK.* Retrieved from http://www.rssconference.org.uk/wp-content/uploads/2012/02/Abstracts-booklet.pdf

7. Markman M, Iseminger KA, Hatch KD et al. Tamoxifen in platinum refractory ovarian cancer: a Gynaecologic Oncology Group Ancillary Report. *Gynecol Oncol* 62, 4-6 (1996).

8. Perez-Gracia JL, Carrasco EM. Tamoxifen therapy for ovarian cancer in the adjuvant and advanced settings: systematic review of the literaure and implications for future research. *Gynecol Oncol*. 2002;84:201-9

9. Gore M, A'Hern R, Stankiewicz M, et al. Tumour marker levels during marimastat therapy. *Lancet* 1996;348((9022)):263-4.

10. Hurteau JA, Brady MF, Darcy KM et al. Randomised phase III trial of tamoxifen versus thalidomide in women with biochemical-recurrent-only epithelial ovarian, fallopian tube or primary peritoneal carcinoma after a complete response to first-line platinum / taxane chemotherapy with an evaluation of serum vascular growth factor (VEGF): A Gynecologic Oncology Group Study. *Gyn Oncol*. 2010;119:444-50.

11. Liu P-Y, Alberts D, Monk B, Brady M, Moon J, Markman M. An early signal of CA-125 progression for ovarian cancer patients receiving maintenance treatment after complete clinical response to primary therapy. *J Clin Oncol* 2007;25:3615-20.

12. Coleman R L, Gordon A, Barter J, Sun S, Rackoff W, Herzog TJ. Early changes in CA125 after treatment with pegylated liposomal doxorubicin or topotecan do not always reflect best response in recurrent ovarian cancer patients. *Oncologist* 2007;12(1):72-8

Figure
Click here to download high resolution image

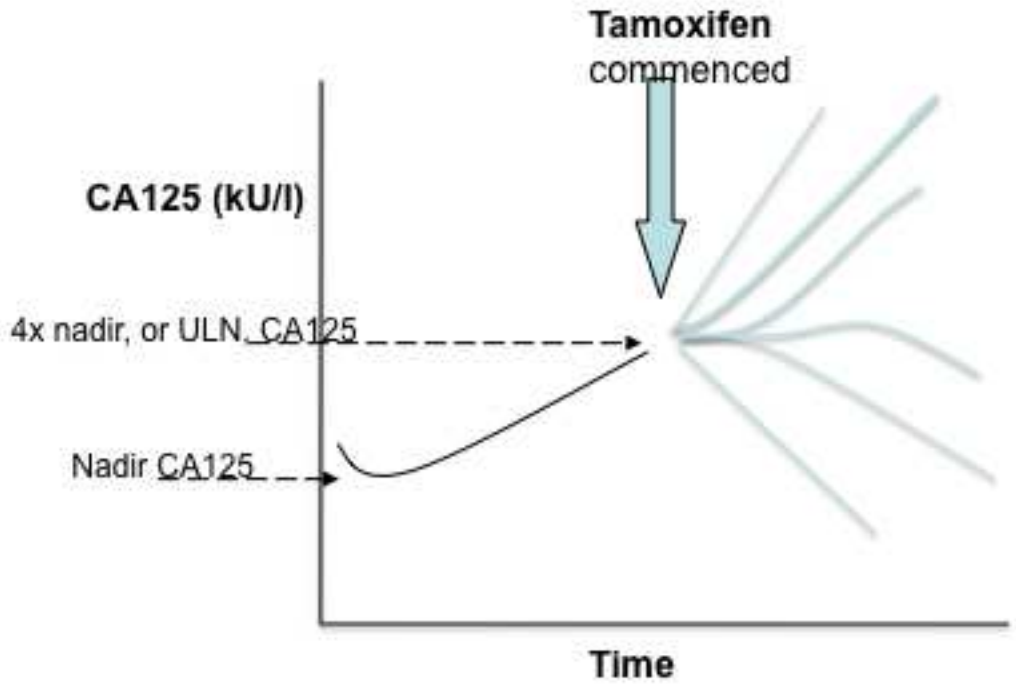

Figure 1:
Possible trajectories of CA125 levels after starting tamoxifen for rising CA125 in asymptomatic patients with relapsed ovary cancer

Table 1:

Regression slopes for ln(CA125) before and after starting tamoxifen (n=81) [ln(CA125)/day]

	Mean	Standard deviation	95% Confidence interval	P-value
Before	0.0149	0.0108	0.0125 – 0.0173	<0.0001
After	0.0093	0.0093	0.0073 – 0.0114	<0.0001

Table 2:

Paired slope comparison before and after starting tamoxifen
[ln(CA125)/day]

Number of patients	Mean change	Standard deviation	95% Confidence interval	P-value
81	0.0056	0.0146	0.0024 – 0.0088	<0.001

Table 3:

Regression slope for those patients who did not start on tamoxifen [ln(CA125)/day]

Number of patients	Mean	Standard deviation	95% Confidence interval	P-value
74	0.0131	0.0132	0.0100 – 0.0161	<0.0001

Figure 2: Spaghetti plot of observed ln(CA125) over time for each patient (vertical lines represent start of tamoxifen therapy)

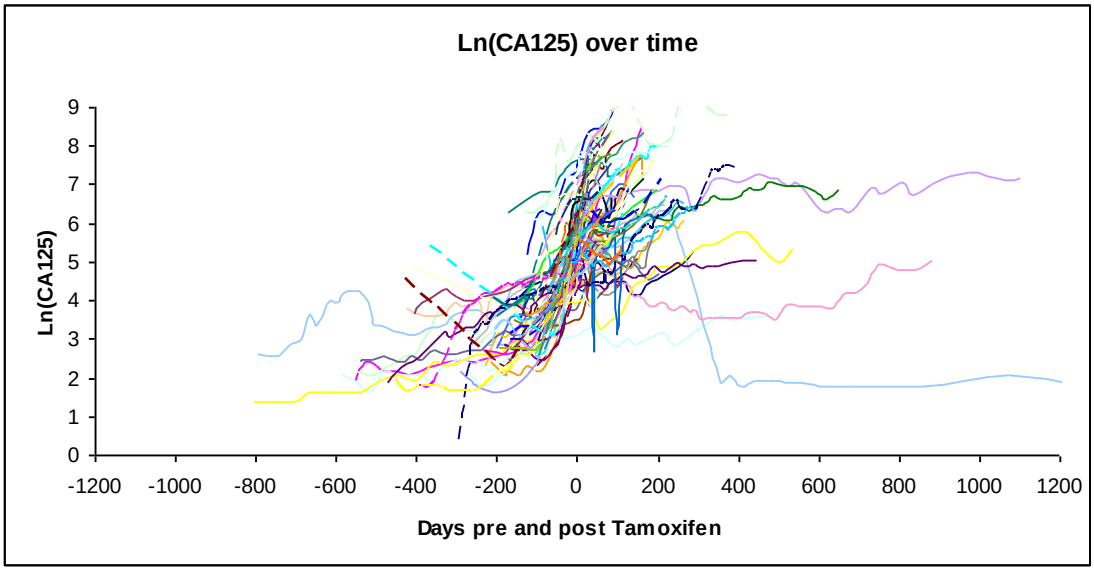

Figure 3: Scatter plot showing each patient's slope of ln(CA125) before and after treatment with tamoxifen