Author's response to reviews

**Title:** Importance of pre-analytical steps for transcriptome and RT-qPCR analyses in the context of the phase II randomised multicentre trial REMAGUS02 of neoadjuvant chemotherapy in breast cancer patients

**Authors:**

Patricia de Cremoux (patricia.de-cremoux@curie.net)
Fabien Valet (fabien.valet@curie.net)
David Gentien (david.gentien@curie.fr)
Jacqueline Lehmann-Che (jacqueline.lehmann-che@sls.aphp.fr)
Véronique Scott (scott@igr.fr)
Carine Tran-Perennou (carine.perennou@curie.net)
Catherine Barbaroux (catherine.barbaroux@curie.net)
Nicolas Servant (nicolas.servant@curie.fr)
Sophie Vacher (s.vacher@st-cloud-huguenin.org)
Brigitte Sigal-Zafrani (brigitte.sigal@curie.net)
Marie-Christine Mathieu (mathieu@igr.fr)
Philippe Bertheau (philippe.bertheau@sls.aphp.fr)
Jean-Marc Guinebretière (jm.guinebretiere@stcloud-huguenin.org)
Bernard Asselain (bernard.asselain@curie.net)
Michel Marty (m.marty@sls.aphp.fr)
Frédérique Spyratos (f.spyratos@curie.net)

**Version:** 2  **Date:** 19 January 2011

**Author's response to reviews:** see over
Quoted: Reviewer's report

Title: Importance of pre-analytical steps for transcriptome and RT-qPCR analyses in the context of the phase II randomised multicentre trial REMAGUS02 of neoadjuvant chemotherapy in breast cancer patients

Version: 1 Date: 29 October 2010

Reviewer: John Martens

**Major Compulsory Revisions:**

1 - I think the authors need to include more details on how the gene expression datasets were merged. Were they united without normalisation of the separate datasets first. Furthermore, were any other algorithms used to better align the different datasets. There are several procedures to do this and it is not described whether or how this was done. Currently, the description of how that microarray datasets were merged is insufficient to judge the firm conclusion that more stringent criteria are need for integrating microarrayRNA data compared to real time results.

The method section was rewritten in a more detailed and reproducible way. It is now well indicated that “Microarrays data provided by the four centres were normalized together using the GCRMA procedure”. No others algorithms were used to better align the different datasets. In addition, Microarray data were uploaded on GEO data base, with accession number GSE26639. Here is the link for reviewer access:


A supplementary methods section was added in order to precise the functions and the corresponding R packages that were used at each step. Details of normalisation algorithm and procedure used are now fully detailed in this supplementary methods section. From data on GEO database, each step of analysis can now be performed using the R functions provided by this supplementary method section.

**Minor Essential Revisions:**

2 - Inclusion of a REMARK-type of diagram would be helpful for the reader. What were exact inclusion criteria and what were the subsequent exclusion criteria (step by step) why from the initial included 340 biopsy-ed patients only 226 and 230 were in the end used in the microarray and real time studies that were presented.

We agree with the reviewer. We proposed now in Figure 1, a flow chart diagram to highlight inclusion and exclusion criteria and facilitate the reader progression.

3 - What type of clustering was performed?

All clustering details (similarity measure and agglomerative criteria) were now given in the method section: a hierarchical clustering was performed. Spearman correlations were used as similarity measures, and we used a Ward algorithm as the agglomerative criterion. In addition, details on the correlations matrix calculation and the R functions used are given in Transcriptome and RT-qPCR analysis section (supplementary methods section).
4 - Have the authors tested for the robustness of clustering?
Indeed, we performed analyses for testing the robustness of clustering. However, due to the large number of variables (5000 probe sets), we were not able to efficiently use stability measures that were shown unsuitable for large datasets and time consuming [1–2]. However, we investigated through a re-sampling approach the stability of the centre 4 effect. Details are given in Cluster stability section (supplementary methods section) and in Table 4s and 5s. Our results showed that whatever the numbers of clusters that could be defined, centre 4 patients were more likely to be classified within the same cluster.

1 - Giancarlo, R., D. Scaturro, and F. Utro, Computational cluster validation for microarray data analysis: experimental assessment of Clest, Consensus Clustering, Figure of Merit, Gap Statistics and Model Explorer. BMC Bioinformatics, 2008. 9: p. 462.


5 - Table 3 and 4 are less essential and could be presented in the supplement.
We agree with the reviewer that some tables are less important and could be placed in the supplement. Within this aim, Table 2 and 3 were now moved to supplementary data (now Table 2s and 3s).
Reviewer's report  
Title: Importance of pre-analytical steps for transcriptome and RT-qPCR analyses in the context of the phase II randomised multicentre trial REMAGUS02 of neoadjuvant chemotherapy in breast cancer patients  
Version: 1 Date: 31 October 2010  
Reviewer: Jorge Reis Filho

Major reviews

1 - This study is NOT MIAME guidelines compliant and the microarray-based gene expression data were NOT deposited in a public repository. At this day and age, this is utterly unacceptable and needs to be rectified prior to any resubmission.  
Microarray data were uploaded on GEO with accession number GSE26639. Here is the link for reviewer access:  

2 - The microarray analysis methods are described in such a sketchy fashion that this reviewer would be unable to reproduce the analyses described in this manuscript. Following the results of Ioannidis et al. (Nat Genet 2009), who were unable to reproduce the results of several high profile studies on gene expression profiling owing to lack of analysis details, the authors have the onus to provide sufficiently detailed analysis methods. This would be best achieved by providing a supplementary Sweave document (or any form of auto-executable file), similar to those described by Baggerly and Coombes.  
The method section was rewritten in a more detailed and reproducible way. It is now well indicated that “Microarrays data provided by the four centres were normalized together using the GCRMA procedure”. No others algorithms were used to better align the different datasets. A supplementary methods section was added in order to precise the functions and the corresponding R packages that were used at each step. Details of normalisation algorithm and procedure used are now fully detailed in this supplementary methods section. From data on GEO database, each step of analysis can now be performed using the R functions provided by this supplementary method section.

3 - According to the authors, “Tumour cellularity was evaluated on frozen sections of the biopsies dedicated to RNA extraction by local pathologists identifying epithelial tumour cell vs. stromal cells, inflammatory cells and necrosis”. This assessment was of crucial impact on the sample selection, given that only samples with >30% of tumour cells were “kept for further analysis”. Despite the importance of this step, the methodology for the assessment of tumour cellularity is subjective. In fact, it is unclear whether the local pathologists assessed percentage of nuclei, percentage of section area or percentage of cytoplasmic area. This needs to be clarified. Furthermore, the method employed by the pathologists to determine these percentages needs to be fully described.  
We agree with the referee. Indeed, evaluation of tumor cellularity by the pathologists on the frozen section of the biopsies dedicated to RNA extraction has major importance for further analyses. This preanalytical step was largely discussed in the group by the pathologists at the time of the design of the trial. Frozen cores were embedded in OCT and sectioned at -20°C in a cryostat. Sections (5 µm in thickness) were stained with haematoxylin-eosin-safran to assess the percentage of invasive malignant cells. Interpretation of the frozen biopsies was performed by local staff breast pathologists. Recording criteria were defined at a consensus meeting prior to the study. The percentage of invasive and in situ malignant cells was recorded versus the amount of benign epithelial cells, stromal and inflammatory cells nuclei.
Semi-quantitative evaluation was performed. For patients in whom multiple biopsies were available, the biopsy with the highest invasive content was used for microarray analysis. The description is now clarified in the text in the 2nd paragraph of “Tissue samples” section.

4 - In the RNA extraction protocols, it is unclear if a step to eliminate completely any trace of ‘contaminant’ DNA was introduced. If that was not the case, how would the authors exclude that some some of the SYBR Green results were not derived from amplification of the genomic DNA? Please clarify and provide the required controls.

In the RNA extraction protocol, we did not include any step to eliminate completely any trace of contaminant DNA. The reason was that we and others experienced that DNAse may alter the quality or quantity of RNA. But we included two steps allowing to be sure to have no DNA contaminants:

a. First all samples were tested for albumin DNA contaminants using an intronic albumin gene design. No amplification of albumin DNA was observed in our series of samples.

b. In addition, target genes’ primers and probes were designed to amplify only exonic sequences. The design was done on two adjacent exons of the gene and we systematically chose an intermediate intron > 1000 bp. The amplicon size was for all systems around 100 bp and the qPCR efficiency was always >95% for all analyzed genes. These conditions did not allow amplifying any contaminant DNA, if present.

This was more clearly explained in the text in “RNA isolation, quantification and qualification” section.

5 - The choice of housekeeping genes is questionable. Some of these housekeeping genes have been shown to have a remarkably poor performance when tumour (rather than cell line) material is used (please see Lab Invest. 2005 Jan;85(1):154-9; Mol Diagn. 2004;8(2):107-13; J Mol Med. 2005 Dec;83(12):1014-24). This reviewer would strongly encourage the authors to include at least a few more robust housekeeping genes. Indeed, this is a great issue in RT-qPCR quantification, especially for breast carcinoma analyses. Both laboratories in charge of RT-qPCR have tried to answer this question since 1999. Through more than 50 articles (I Bièche, F Spyropoulos or P De Cremoux cited in Pubmed), we have investigate the variability of 10 reference genes, showing that TBP and RPLPO had the lowest variability. In addition, the other genes were chosen to allow further comparison with the Oncotype DX algorithm.
6 - Another question that is germane here is whether the pathologists were trained a priori. Given the poor inter-observer agreement between pathologists, different pathologists may have excluded cases that had approx 30% of tumour cells differently and this may have led to biases in sample selection. This preanalytical step for tumor cell analysis was largely discussed in the group by the pathologists at the time of the design of the trial. They a priori defined the criteria for identification of tumor cellularity as defined in the answer to question 3. A consensus of lecture was initially defined by the group of pathologists and applied for all biopsies. This is added in the second paragraph of Tissue sample section.

7 - The choice of 30% of tumour cells is debatable. The results of Cleator et al. (Breast Cancer Res 2006) and Weigelt et al. (Lancet Oncol 2010) demonstrate clearly that different percentages of non-neoplastic cells in the samples affects the performance of gene classifiers. Before accepting the 30% threshold based on convenience (i.e. the same threshold used in other studies), the authors should instead do a series of spiking experiments to demonstrate what percentage of non-neoplastic cell contamination affects the performance of qRT-PCR and microarrays for classification of samples by gene expression profiling. We agree with the referee that the choice of a cutoff for cellularity is of upmost importance in transcriptomic studies and that it may induce potential bias. We have read with interest the Weigelt’s paper and the previous papers of the same group. The option of microdissection under microscope with a needle allowing isolating samples with more than 90% tumor cells is interesting but not easily applicable to biopsies in the context of a clinical trial like REMAGUS02. Concerning the Cleator’s paper, we observed that a cutoff of 20% of tumor cells was used by the authors before studying the effect of stromal component of breast tumors. Members of our group participated to the discussions preceding the MINDACT trial, the initial cutoff of 50% was changed to 30% because a too large number of cases was excluded; this experience was useful to choose the cutoff of 30% in the REMAGUS02 trial.

8 - The fact that Centre 4 is a clear outlier needs to be better explored in the discussion. Is this a mere problem of RNA extraction (Trizol vs other methods)? How would this impact on the interpretation of the results of from material obtained from current clinical trials where different extraction methods have been employed. Are there bioinformatic methods to minimise the impact of different extraction methods or is this an insurmountable hurdle? The authors need to expand on this in the discussion. We have improved the discussion concerning the material from Center 4, as indicated in the last paragraph of the Discussion section.
9 - This reviewer cannot subscribe to the authors’ conclusion that “Our data showed that strict quality criteria for RNA integrity assessment and well calibrated and standardized RT-qPCR allows multicentre analysis of genes transcripts with high accuracy in the clinical context.” In fact, this manuscript shows that even with strict quality criteria for RNA integrity assessment and well calibrated and standardised RT-qPCR, there was centre bias in the gene expression analyses. Without offering a solution for the centre-based bias, this conclusion is not supported by the authors’ data.

In the present work, using the same material, we observed a clear difference between the observations made with RT-qPCR analyses and transcriptome analyses. The centre 4 effect disappeared in the strict experimental conditions used for RT-qPCR i.e. calibration curves (absolute quantification, using 8 different points of the standard curve, one calibrator and reference genes). We have more detailed these data in the discussion section, clarifying the fact that using the same material, qPCR standardization methods may handle the differences of quality of RNA but not for the transcriptome analyses. More stringent RNA quality criteria were needed in this later case.

10 - The high number of cases excluded prior to the analysis owing to low cellularity (~15%) should also be discussed. How can microarrays and qRT-PCR be used if 15% of the samples are not adequate for this type of analysis? It is also important to determine whether there was a difference in terms of tumour size, ER, PR, HER2, grade and histological type between the samples excluded and included in the study. One could hypothesise that the group of samples with <30% of cancer cells is enriched for lobular carcinomas or ER positive tumours.

Among the cases excluded for insufficient quantity of tumor cells, the median [min – max] percentage of tumor cells was equal to 10 [0 – 20]. In addition, comparison of characteristics between samples excluded and included in the study (see tables s6 and s7) did not evidence any heterogeneity regarding tumor size, ER, PR, HER2.

However, for both transcriptome and RT-qPCR analyses, we evidenced that

i) the proportion of samples of low grade was significantly higher in excluded material (pv equal to 0.018 and pv=0.044 for transcriptome and RT-qPCR analyses respectively)

ii) the proportion of lobular carcinoma were significantly higher in excluded material (pv=0.003 and pv=0.004 for transcriptome and RT-qPCR, respectively)

These results were now added in the RNA quantification and qualification paragraph of the results section as well as in the discussion section.
11 - Out of the 290 samples with sufficient tumour cell content, 22% and 18% of samples were not available for microarray and qRT-PCR analysis, respectively. Why were the samples excluded? This needs to be explained in greater detail in the methods. The flow chart shows more clearly the causes of exclusion. It is now included in the paper as figure 1. It was probably not clear enough in the text that we kept for Biostatistics RT-qPCR analyses, only samples with complete data for the 45 genes.

12 - Out of all samples collected, only 66.4% were successfully analysed by microarrays (i.e. one third of samples were excluded). Again, the impact of this high failure rate needs to be discussed in relation to the design of clinical trials that also address translational research questions based on transcriptomic analysis. We agree that a high failure rate is an important point to be taken in consideration particularly in prospective clinical trials where transcriptome analyses are decisional. In the EORTC10994 p53 trial where very strict freezing conditions were imposed, 20% of samples were excluded. Much higher failure rates were observed in the adjuvant MINDACT trial. A paragraph was added in the second paragraph of the discussion section.

Minor reviews

13 - The term Elston-Ellis grade is most unorthodox. Chris Elston and Ian Ellis themselves have always referred to their modification of the Scarff-Bloom-Richardson system as “Nottingham Grading System”. Furthermore, the reference Elston & Ellis (Histopathology 1991) should be included in the manuscript. We added the Elston Ellis reference in the reference list and changed the sentence in the text in accordance with the referee wishes.

14 - The Clinicaltrials.gov reference number and webpage of the trial should be included in the manuscript. The reference number of Remagus 02 trial is now added in the text, in the patients and samples section (ISRC TN 10059974)

15 - Some typos were found and need to be corrected prior to resubmission. We have now corrected the text before resubmission