Author's response to reviews

Title: Resistance gene expression determines the in vitro chemosensitivity of non-small cell lung cancer (NSCLC)

Authors:

Sharon Glaysher (sharonglaysher@googlemail.com)
Dennis Yiannakis (Dennis.Yiannakis@phnt.swest.nhs.uk)
Francis G Gabriel (francis.gabriel@porthosp.nhs.uk)
Penny Johnson (penny.johnson@porthosp.nhs.uk)
Marta E Polak (marta.polak@googlemail.com)
Louise A Knight (louise.knight@porthosp.nhs.uk)
Zoe Goldthorpe (zoegold@hotmail.com)
Katharine Peregrin (kate.peregrin@googlemail.com)
Mya Gyi (mya.gyi@porthosp.nhs.uk)
Paul Modi (paulmodi@doctors.org.uk)
Joe Rahamim (Sarah.Sweetlove@phnt.swest.nhs.uk)
Mark E Smith2 (mark.smith@phnt.swest.nhs.uk)
Khalid Amer (Khalid.Amer@suht.nhs.uk)
Bruce Addis (bruce.addis@suht.nhs.uk)
Matthew Poole (Matthew.Poole@port.ac.uk)
Ajit Narayanan (ajit.narayanan@aut.ac.nz)
Tim J Gulliford (tim.gulliford@porthosp.nhs.uk)
Peter E Andreotti (pandreotti@asdxbiosystems.com)
Ian A Cree (ian.cree@porthosp.nhs.uk)

Version: 2 Date: 16 June 2009

Author's response to reviews: see over
To the Editor

Thank you for your email response to our paper. ArrayExpress do not normally take PCR data and although this is an array experiment, the format is that of a 384 well plate and there are only 96 genes per cases. As an alternative and potentially more useful option, we have created a separate file with the raw data which can be published on-line with the paper as supporting information. We hope that this is acceptable. We have re-structured the abstract, and made the copy editing corrections in Word, though some of these are probably due to formatting issues between versions of Word and may require the some editorial adjustment. We have also added competing interests and authors' contribution sections as requested. I hope that these meet the needs of the journal.

We are grateful for the helpful comments made by both reviewers, and would respond as follows:

Dr Paul Boutros
We are pleased that you thought this to be an important topic, and we hope that it establishes a proof of principle which will be of use to others working in this field. We note that the original paper, your concerns and this response will be available to authors who wish to see more information, and they are welcome to contact me for further clarification, though I hope that our detailed response answers your concerns.

To answer these in order:
1. The ATP-TCA has been used in a large number of publications, and the IndexSUM method is well-established. We have quoted the first publication to use it in 1993 in the list of references. The IndexSUM is defined as the sum of the surviving fraction of cells at each dilution tested and allows description of the activity of a drug by a single measurement. It is skewed towards measurement of effect at lower concentrations, which provides better discrimination between drugs or patients than any other parameter we or others using this technique have measured over the years.

2. The use of four housekeeping genes is based on our previous paper (ref 9), and we therefore included these genes again in our Taqman array. In preliminary work, we examined the correlation between the results obtained for each gene and showed that PBGD was the least variable gene. As the others showed greater variability, the use of PBGD alone was preferred over the use of a geometric mean. This is assisted by the fact that PBGD comes up at a Ct value close to most of the genes studied. We are not sure that it would be beneficial to include a figure to show this here, but can do so if required and have made the point clearer in the text.

3. The statistical analysis of qRT-PCR data is worthy of a paper in its own right, and we received assistance from experts at Applied Biosystems who examined the data independently. We realised that the standard method of analysis (deltaCt) was inappropriate for linear regression as it skews the data obtained, and therefore those to use a gene expression ratio, normalising the data by logarithmic transformation. The ATP-TCA data were indeed treated as a continuous dependent variable and tend to be normally distributed. No intercept was term was included, and genes were added by forward regression according to their univariate correlations following entry of each gene. The F-value is essentially the threshold for inclusion of a gene into the forward linear regression model within SPSS: the default is 0.5, and was relaxed to allow a greater number of genes to enter the model based on initial assessment of the most appropriate model size using Akaike Information Criterion versus model size. The Akaike Information Criterion (AIC) is a
function of model error and size which penalises large models, and the lowest number is regarded as best. The PRESS (prediction residual sum of squares) method used is an adjusted regression method which was employed to prevent overfitting, as it is a ‘leave one out method’. We can include AIC data to show the optimal model size if required, though I am concerned that the inclusion of too much statistical analysis could be counter-productive, and have included a simplified but hopefully clearer explanation in the statistics section.

4. This is true, but in our defence, this is revisiting published data and we see no point in enlarging the paper further by evaluating synergy formally in the ATP-TCA, which is not the key focus of the paper.

5. The per patient curves shown in figure 1 are there to show the degree of heterogeneity from the raw data – the distribution of responses observed is provided as figure 2 using IndexSUM derived from these data. We have certainly used summary curves with confidence intervals previously, but in our opinion, the method chosen shows the degree of heterogeneity much more clearly and we would like to keep it.

6. The analysis suggested has been added to the results section, with an additional figure. The TCA Ovarian Cancer Trial data (Cree IA et al, Anti-Cancer Drugs 2007; 18: 1093-1101) examined this in a different tumour type, but with greater direct clinical relevance.

7. The deposition of raw PCR data in the ArrayExpress database is not currently feasible, according to their helpline, and we have instead put together a spreadsheet to be published alongside the paper. We have also added a further table which includes the list of genes. The authors take this opportunity to remind the reader that these data are for research use only – the correlation of ATP-TCA data with clinical outcome in lung cancer is unknown and the use of surgical specimens has introduced an inevitable bias in the type and size of tumours studied. Use of these models in patients is not advised and we would regard such use as unethical unless full validation against clinical data as part of a clinical trial is performed.

8. The individual correlations for each model are included in table 4. No single regressions reached significance and in any case, are rendered meaningless by the Bonferroni correction which would be required. We did not allude to this in the text, though there may have been some confusion from a comment in the results section and this has been amended to make it clear that we only performed multiple linear regression.

9. Minor revisions – we are grateful for these and have made the necessary adjustments in the text.

Dr Yataro Daigo
Thank you for your useful comments. Our data simply provide proof of principle and confirm that there is correlation between in vitro chemosensitivity and expression of known resistance mechanisms – this is a long way from personalised signatures, but an essential step, in our view. To respond to each comment:

1. The ATP-TCA needs around 10% neoplastic cells, or they tend not to survive in culture: non-neoplastic cells die during the six-day culture period due to the use of an adherence-free system and selective medium. One of the problems of RNA extraction is that the sample contains material derived from both neoplastic and non-neoplastic cells. We reduced this as far as possible by careful selection of blocks and were helped by the fact that most of the normal lung around the tumour is relatively poorly cellular. We would
estimate that at least 70% of the RNA was of neoplastic origin, but have no markers to confirm this.

2. We attempted the subgroup analysis suggested, but found no difference. It is likely that there were insufficient numbers for this sub-group analysis and we could not obtain any meaningful results: this is noted in the results section.

Yours sincerely

Ian A Cree, for the authors.