Research article

Cell killing and resistance in pre-operative breast cancer chemotherapy

Paolo Ubezio*1 and David Cameron2

Address: 1Biophysics Unit, Department of Oncology, Istituto di Ricerche Farmacologiche "Mario Negri", Via La Masa 19, I-20156 Milan, Italy and 2Department of Oncology, Edinburgh University, Western General Hospital, Crewe Road South, Edinburgh EH4 2XU, UK

Email: Paolo Ubezio* - ubezio@marionegri.it; David Cameron - d.cameron@ncrn.org.uk

* Corresponding author

Abstract

**Background:** Despite the recent development of technologies giving detailed images of tumours in vivo, direct or indirect ways to measure how many cells are actually killed by a treatment or are resistant to it are still beyond our reach.

**Methods:** We designed a simple model of tumour progression during treatment, based on descriptions of the key phenomena of proliferation, quiescence, cell killing and resistance, and giving as output the macroscopically measurable tumour volume and growth fraction. The model was applied to a database of the time course of volumes of breast cancer in patients undergoing pre-operative chemotherapy, for which the initial estimate of proliferating cells by the measure of the percentage of Ki67-positive cells was available.

**Results:** The analysis recognises different patterns of response to treatment. In one subgroup of patients the fitting implied drug resistance. In another subgroup there was a shift to higher sensitivity during the therapy. In the subgroup of patients where killing of cycling cells had the highest score, the drugs showed variable efficacy against quiescent cells.

**Conclusion:** The approach was feasible, providing items of information not otherwise available. Additional data, particularly sequential Ki67 measures, could be added to the system, potentially reducing uncertainty in estimates of parameter values.

Background

Mathematical modelling of cancer and cancer therapy has been attempted on all scales. There have been interesting examples ranging from the microscopic levels of single molecular interactions and protein network modelling of specific cellular functions [1], to cell cycle and in vitro cell proliferation, simulating the time course of cytostatic and cytotoxic effects of drugs [2-4], up to the macroscopic level considering tumour growth in vivo [5-13].

Mathematical models differ also in the way they use experimental data, some being more theoretical, trying to explain the general behaviour of a biological system, and others aiming to fit specific data sets.

At the level of in vivo tumour growth, a typical datum is the time course of tumour volumes, measured (by callipers or in vivo imaging) in animals in the pre-clinical stage of drug development or in humans. Because these are rel-
At relatively simple datasets, simple descriptive models of tumour growth have been used. For instance Skipper adopted a model of exponential growth, defined by the number of tumour cells at the start of the observation period and by the doubling time [14]. This model was appropriate to describe leukaemia developing in mice. The effect of a cytotoxic drug was described by the fraction of cells killed at each administration, so the complete model (tumour growth + treatment) had only three parameters.

Norton and Simon improved Skipper's model and used the Gompertz function to describe a continued decline of the growth rate as the tumour mass increases [15]. This shape fits many experimental time courses of solid tumours, particularly in animals, where the whole natural history of the tumour can be followed. Reduction of drug efficacy in massive tumours was described by a simple link between the fraction of surviving cells at each treatment and the growth rate. Similarly, Goldie and Coldman [16] and others [8,10] tackled the issue of drug resistance using the probability to spontaneous mutation(s) towards a resistant phenotype (i.e. with a single parameter).

Of course all these authors were aware that tumour growth and response to a drug are very complex phenomena, with many interacting factors, from molecules to environmental constraints. For instance, the dynamics of the response to a drug challenge was recently modelled using no fewer than six time-dependent parameters, associated with cell cycle block and cell killing in each phase of the cell cycle [2]. In that work, considering a single population of tumour cells growing in vitro in the best environmental conditions, a less rich model was unable to fit all available data (a dataset including measures with different techniques). Thus, in a certain sense, the success of simple models at the higher in vivo level is probably a consequence of the scant datasets currently available. On the other hand, using complex models with many parameters to describe tumour growth seems a purely theoretical exercise, not giving any useful additional clinical insights if – as is often the case - those same experimental time courses are satisfactorily fitted (taking into account the poor precision of the data) by simple polynomial or exponential functions.

In this paper we propose an intermediate approach. We adopt a richer data set from patients with breast cancer receiving pre-operative chemotherapy, including a measure of proliferation (Ki67 staining), and use an interpretative (phenomenological) model to fit the data. This new model is basically different from the previously reported model [17,18] of breast cancer response to therapy. The "old" model attempted to estimate the values of cell kill and resistance parameters that were consistent with the overall decrease in volume of individual tumours, simplifying with a single growth equation the net result of the interplay between cell cycle, quiescence and loss.

Differently from models based on simple descriptive functions (like exponential and Gompertz), the new model was based on that underlying interplay. Because of the nature of breast cancer, where only a fraction of the cells are actually cycling (the fraction estimated by the measure of Ki67), an essential feature in the new model is the consideration of quiescence [9,19,20], together with proliferation and cell loss, exploiting the results of the mathematical theory of age-structured cell populations with a quiescent compartment [21,22]. As a consequence, the response to treatment was modelled taking into account of the different effects against cycling and quiescent cells. In addition, variants of the basic model with increasing complexity were considered, starting from a model neglecting the resistance up to a model including resistance and the shift of treatment efficacy during the course of treatment. We adopted the principle of parsimony, accepting the simplest model when the inclusion of additional parameters did not significantly improve the fit. We investigated not only how well this model fitted the data of several patients, but also the precision of the estimates of parameters. Taking that precision into account, the values of parameters characterising treatment efficacy (killing of cycling and of quiescent cells, drug-resistant fraction) were categorized and further analysed for association and correlation with common biological markers. Eventually, we will discuss which additional data would be useful to improve the precision of parameters estimate.

**Methods**

**Patients and histology**

Thirty-five patients with large operable breast cancers (stages T2 or T3) received preoperative chemotherapy based on cyclophosphamide 1 g m⁻² and doxorubicin 50 mg m⁻² every three weeks. Eleven patients already had hormone therapy for up to three months which had not induced any tumour shrinkage.

Histological confirmation of invasive breast cancer was obtained by wedge biopsy of the primary tumour or a palpable axillary node, from which material was also made available for determination of the estrogen receptor (ER) concentration and Ki67 as a measure of tumour cell proliferation. ER concentration was measured by the dextran-coated charcoal (DCC) method. Ki67 antigen was revealed by immunohistochemical staining with antibody MIB1 (Europath Ltd, Cornwall, UK) diluted x50. Reactivity was detected by an ABC-peroxidase-antiperoxidase (PAP) method providing the percentage of Ki67-positive.
cells (%Ki67+) at the start of treatment. Mean %Ki67+ measured in this set of tumours was 37% (standard deviation: 24%).

Tumour volume was estimated weekly for at least six weeks, by measuring two orthogonal diameters with callipers and applying the ellipsoid volume formula, assuming the third dimension as the average of the other two. The tumour volumes measured on day 0 ranged from 22 cm³ to 204 cm³, mean 70 cm³, median 58 cm³; the detection limit was 0.1 cm³. The number of positive axillary nodes (NOD) and the clinical and pathological response were assessed at the time of surgery.

Clinical response was classified according to the UICC criteria [23]. Clinical complete remission (CR) was defined as the disappearance of all palpable tumour deposits and partial remission (PR) as a more than 50% reduction of tumour volume. Tumour reduction less than 50% or an increase up to 25% was scored as stable disease (SD). An increase of more than 25% was designated as progressive disease (PD).

Pathological complete remission (pCR) was defined as a tumour with no residual microscopic disease in either the breast or ipsilateral axillary lymph nodes.

All patients in this study were treated within the context of a clinical trial which had been approved by the regional ethics committee, and for which patients gave informed consent. Subsequently further ethics approval was obtained from the same committee to cover a broad range of research projects, including that reported herein on the tumour tissues, which did not require patients to be explicitly re-consented.

Model
The program combined the three models of tumour growth, treatment and resistance. The underlying mathematics is reported in Appendix 1.

Tumour growth model
The model included a compartment of quiescent cells, and a compartment of cycling cells with age structure (Figure 1). Cycling cells spent a time Tc in the cell cycle before division. After division, newborn cells either re-start the cycle, with probability θ, or enter the quiescent compartment, with probability 1-θ. Quiescent cells may die (with rate μq, so that μqdt is the probability to die in the time interval dt) or re-enter the cycle, with rate γ. We assumed that most cells become quiescent before dying, and the probability of "natural" death (in the sense of "not induced by a treatment") was applied only to quiescent cells.

Within this model, in order to identify univocally the tumour growth characteristics we need to specify four measurable (at least in principle) macroscopic quantities: the doubling time (T doubling), Tc (or Tpot, the potential doubling time), the growth fraction (GF) and γ. The other parameters (θ and μq) are dependent on these four quantities (see Appendix 2).

Treatment model
The main feature of the model is the drug’s different effects on cycling and quiescent cells. The basic parameters are:

Sp: the fraction of cycling cells surviving a single drug administration.

Sq: the fraction of quiescent cells surviving a single drug administration.

To reproduce the (slow) process of disruption/elimination of killed cells, we adopted the method described by Simeoni et al. [11], successfully applied to experimental tumours. In brief, dying cells stop proliferating and pass through three stages, with progressive degrees of damage, before they are definitively lost. The passage from one stage to the next is mathematically described by a rate constant k, which thus adds to the parameters of the model (see Appendix 1).

Resistance model
Two distinct modalities of resistance were compared: 1) "Initial Resistance", when a fraction of the cells was already resistant to a drug at the start of treatment; 2) "Induced Resistance", when a fraction of sensitive cells surviving treatment became resistant. Each modality was modelled by a single parameter, either:
\textit{IniR}: the fraction of cells initially resistant to the drugs, or
\textit{Rind}: the fraction of sensitive cells becoming resistant after a single treatment.

**Fitting procedure and sensitivity analysis**

Eventually, merging the three models of tumour growth, treatment and resistance and including the tumour volume at the beginning of treatment ($V_{0}$), we have a model with nine independent parameters: $V_{0}$, $Td$, $Tc$, $GF$, $\gamma$, $Sp$, $Sq$, $k$, $IniR$ (or $Rind$).

In order to reduce the number of parameters for fitting, we assumed a specific type of tumour growth, exploiting the measure of Ki67 as an estimate of the growth fraction, and found the parameters of cell kill and resistance by an optimisation procedure. A constrained, non-linear fitting procedure was used, maximising the likelihood function of the logs of tumour volumes, with Gaussian distribution of data errors, taking their standard deviation as a parameter. The following constraints were used:

\[ 0.7 \cdot V_{\text{measured}} < V_{0} < 1.3 \cdot V_{\text{measured}}; \] this assumes that the measure of volume at $t = 0$ had 30% precision;

\[ 0 \leq IniR \text{ (or } Rind) < 1; \] these are probabilities;

\[ Sp < Sq; \] this assumes that cycling cells are more sensitive to treatment than quiescent cells;

\[ Sp > 0.001; \] this assumes that no single treatment will kill more than 99.9% of sensitive cycling cells;

\[ Sq \leq 0.98; \] this assumes that no single treatment will kill less than 2% of sensitive cells (otherwise cells are defined as "resistant");

\[ 0.5 \leq k < 1; \] the minimum value of $k$ (0.5) produced a delay in the loss of dead cell such that 14 days are needed to lose 99% of the dying cells. When $k$ approaches 1 almost all dead cells are lost in three days.

This procedure was repeated assuming other types of tumour growth, covering all possible combinations of doubling time (either short – 30 days, intermediate – 150 days, or long – 10000 days), cell cycle durations (2, 5 or 8 days) and recycling rate (0 or 0.01), consistent with the value of the growth fraction (see Appendix 2). The maximum-likelihood of the fits obtained with all growth types were compared and the one with highest likelihood ($L_{\text{best}}$) was assumed as best fit for the patient’s tumour time-course. The likelihood ratio test statistics (LRTS) was used to compare a fit with a given set of parameters "X" with the best one (LRTS = $2(\log(L_{\text{best}}) - \log(L_{X}))$. As LRTS follows a chi-square distribution, X was considered equivalent to the best when LRTS < $\chi^{2}_{0.05,1}$. Similarly, likelihood-based 95% confidence intervals for each parameter were obtained by raising or lowering its value until L was reduced to the value of $\log(L) = \log(L_{\text{best}}) - \chi^{2}_{0.05,1}/2$ [24, 25]. The overall range of variability of a parameter was calculated by joining up the confidence intervals obtained with all tumour growth models equivalent to the best.

The whole procedure was repeated using three models of treatment and resistance, at increasing levels of complexity. The lowest level (Level I) considered no resistant cells and the same cell survival at each drug administration, with a difference in sensitivity between cycling and quiescent cells (parameters: $Sp$ and $Sq$). Level II included a parameter measuring drug resistance, in two variants describing either initial or drug-induced resistance (parameters: $Sp$, $Sq$, $IniR$ ($Rind$). Level III included a sensitivity shift after some chemotherapy cycles (parameters $Sp$, $Sq$, $Sp2$, $Sq2$, describing the fractions of surviving cells respectively in the first and second period).

**Implementation**

All analyses were done with a computer program using Microsoft Excel with its standard features (Visual Basic and Solver). With a user-friendly interface, an automatic graphic output of the simulation curve with data is provided at each change of parameter values. The program (“PAOTHERAPY”) is available for non-commercial purposes in Additional file 1.

**Results**

**Adequacy of the models**

Data were fitted using the principle of parsimony, adopting a lower-level model when a higher level did not significantly improve the fit. About half of time-courses had a regular decreasing trend and were fitted simply by applying the same fraction of surviving cells ($Sp$, $Sq$) at each therapy cycle (Level I). In 25% of cases the efficacy declined during therapy, and the data were fitted assuming a fraction of cells resistant to treatment (Level II). In another 25% of cases the efficacy appeared to increase, after 20 (5 patients) or 60 days (4 patients), and the data were fitted with a sensitivity shift, with two phases characterised by different surviving fractions (Table 1).

Representative data and fits are shown in Figure 2. Panel A shows a time course of data (dots) together with the fitting using a level I model (magenta line). The arrows indicate the treatment times. The outlines of cycling and quiescent cells predicted by the model are also reported. In this case the model predicted almost 30% cycling and few percent quiescent cells killed at each treatment cycle, regularly up to the last one, with no onset of resistance and no increase of sensitivity.
Table 1: Summary of the models required for fitting.

<table>
<thead>
<tr>
<th>Model type</th>
<th>Resistance</th>
<th>Sensitivity shift</th>
<th>Parameters</th>
<th>Cases</th>
</tr>
</thead>
<tbody>
<tr>
<td>Level I</td>
<td>NO</td>
<td>NO</td>
<td>Sp, Sq (1)</td>
<td>17/35</td>
</tr>
<tr>
<td>Level II</td>
<td>YES</td>
<td>NO</td>
<td>Sp, Sq, IniR (2)</td>
<td>9/35</td>
</tr>
<tr>
<td>Level III</td>
<td>NO</td>
<td>YES</td>
<td>Sp, Sq, Sp2, Sq2 (3)</td>
<td>9/35</td>
</tr>
</tbody>
</table>

(1) fraction of cycling (Sp) and quiescent (Sq) cells surviving a single cycle of therapy.
(2) fraction of cells resistant to treatment (calculated at time zero)
(3) fraction of cells surviving a single cycle of therapy in the first (Sp, Sq) and second (Sp2, Sq2) phase sensitivity.

Panel B shows an example of fitting with a level II model, including a subpopulation of cells resistant to treatment (dashed line). The model predicted about 95% cycling and 10% quiescent cells killed at each treatment, but this efficacy cannot be maintained after the fourth cycle. In order to fit the last points of the time course, the model predicted that 1% of the cells were initially resistant to the

Figure 2
Examples of fitting. Representative examples of fitting (continuous magenta line) of experimental time courses (dots). The arrows indicate treatment times. Outlines of cycling (green line) and quiescent (light blue line) cells predicted by the model are also reported. Panel A: fitting with a level I model, with Sp = 0.68 and Sq = 0.98 (k = 0.50). Panel B: fitting with a level II model, including a subpopulation of cells resistant to treatment (dashed line). The fit was obtained with Sp = 0.04, Sq = 0.90 and IniR = 0.01 (k = 0.83). Panel C: fitting with a level III model, with an increase of efficacy from the fourth cycle. The fit was obtained with Sp = 0.43 and Sq = 0.98 up to day 60 and Sp = 0.004 and Sq = 0.98 thereafter (k = 0.50). Panel D: an example of unsatisfactory fit, with Sp = 0.34 and Sq = 0.98 (k = 0.50).
therapy and this subpopulation of cells eventually became prevalent (more than 80% resistant cells by day 100).

Panel C shows an example of fitting with a level III model, with an increase of efficacy from the fourth cycle. The model predicted that about 60% of cycling cells were killed in each of the first three cycles, but the fourth killed more than 99%. The efficacy of treatment against quiescent cells remained low.

Best fits for the data of 31/35 patients had a discrepancy between simulated tumour volume and datum below 30% (calculated as the average of all time points), and 26/35 were below 20%. Panel D shows one of the four instances where a less satisfactory fit was obtained. The presence of one or two (probable) outliers in the time course explains the poor score in these cases. They were not excluded from subsequent analyses, because the simulation caught the general pattern of the outline in these cases too.

The performance improves on that obtained with the previously published model [18], with no distinction between cycling and quiescent cells. With the latter simplified model the fittings were always worse, and only in 19/35 instances was the average distance from data below 30%.

The simulation program permits extrapolation of the time course of the tumour volume up to a fixed time, in order to compare patients at the same time-point. Assuming day 100 as end-point, the model suggested that the volume was reduced at least 50% in all 26 non-resistant cases, and by more than 90% in 16/26 instances.

**Estimate of cell kill**

Figure 3 shows the best fit estimates of drug efficacy on cycling non-resistant cells of all tumours, in terms of the surviving fraction Sp, with their 95% confidence intervals. Although in some instances the confidence intervals were very wide (in 6/35 more than 0.5), in most cases the data were informative enough to indicate that more than 50% of cycling cells were killed at each chemotherapy cycle. A score was assigned to cell killing, from 1 to 5, indicating "very low" (Sp > 0.5), "low" (0.25 < Sp ≤ 0.5), "medium" (0.10 < Sp ≤ 0.25), "high" (0.01 < Sp ≤ 0.1) and "very high" (Sp ≤ 0.01) drug efficacy. Taking confidence intervals into account, the score was reduced by 1 when the range exceeded the limit of the previous category. When the confidence range was 0.001 – 0.98, Sp was considered not detectable (ND) and no score was assigned. Sp was not detectable in only one case.

Table 2 shows the cases in the five categories. Cell killing was high or very high in 31.3% detectable patients, medium in 35.3%, low or very low in the remaining 32.3%.

The survival of quiescent cells (Sq) was estimated with lower precision than Sp (Figure 4) and in ten cases (28.6%) confidence intervals were > 0.5. Three cases were not detectable (confidence range: 0.001 – 0.98).

Although best fits suggested values above 0.9 (90% quiescent cells surviving a single treatment) in 18/32 detectable patients, Sq was below 0.5 in six cases and in another six cases the best fit was above but the lower limit of the confidence interval was below 0.5. Thus, the results suggest that killing of quiescent cells was important at least in some – not rare-instances.

Comparing the model results with the clinical/pathological response, 8/10 cases identified as responders (pCR or CR) were coherently classified using the cycling cell killing scores 4 or 5, the other two with score 3. Three cases scored by the model with high cell killing were only PR: two had a late sensitivity shift and the third experienced a 2-log mass reduction.

**Table 2: Frequency of the scores of killing of cycling cells per treatment in the dataset.**

<table>
<thead>
<tr>
<th>Score Sp</th>
<th>freq</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>5</td>
<td>14.7</td>
</tr>
<tr>
<td>2</td>
<td>6</td>
<td>17.6</td>
</tr>
<tr>
<td>3</td>
<td>12</td>
<td>35.3</td>
</tr>
<tr>
<td>4</td>
<td>6</td>
<td>17.6</td>
</tr>
<tr>
<td>5</td>
<td>5</td>
<td>14.7</td>
</tr>
<tr>
<td>ND</td>
<td>1</td>
<td></td>
</tr>
</tbody>
</table>
group 1, 18 cases where a resistant subpopulation was excluded (formally: best fit with \( IniR = 0 \) (model type I) and maximum \( IniR < 0.1 \) at the 95% confidence interval);
group 2, "undetermined resistance", comprising the eight ambiguous cases above;
group 3, nine cases where the presence of resistant cells was demonstrated by a significantly better fit.

The nine cases in group 3 were further analysed with an alternative model where no resistant cells were initially present but resistance was induced by treatment. This enabled us to fit the data with similar precision (not shown). Table 3 shows the range of values of the resistance parameters \( IniR \) (the fraction of cells initially resistant, used in the standard model) and \( Rind \) (the fraction of sensitive cells becoming resistant after a single treatment, used in the alternative model). With the first assumption, resistant cells amounted, on average, to 19% of the tumour at the beginning of therapy. However, assuming resistance induced by treatment, 33% of surviving cells, on average, must become resistant as a consequence of a single drug challenge, which is probably unrealistically high.

**Resistance and cell kill related to estrogen receptor content and number of positive nodes**

Figure 6 shows ER values in the three groups of patients based on the resistance parameter. Average ER value in group 1 (no detectable resistance) was 4 fmol/mg (maximum 15), compared to 80 in group 3 (detectable resistance), with ER > 15 in 6 out of 9 instances. In group 2, i.e. the 8 patients with "undetermined" resistance, average ER was 14 fmol/mg (≤ 15 fmol/mg in 7/8 patients).

Figure 7 reports ER values in relation to treatment efficacy on non-resistant cells, as measured by the score of the parameter \( Sp \). ER was unrelated to the strength of cell killing. Dividing the cases according to the type of model required for fitting showed that the cases fitted with each model clustered in different regions of the plot. Cases fitted with model type III, including biphasic efficacy, clustered in the region of highest cell kill (scores 4–5) and low (<15) ER; model type II, including resistance, clustered in

**Table 3: Resistance parameters.**

<table>
<thead>
<tr>
<th></th>
<th>( IniR )</th>
<th>( Rind )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average</td>
<td>0.19</td>
<td>0.33</td>
</tr>
<tr>
<td>Minimum</td>
<td>0.01</td>
<td>0.02</td>
</tr>
<tr>
<td>Maximum</td>
<td>0.53</td>
<td>0.85</td>
</tr>
</tbody>
</table>

Statistic of best fit resistance parameters in a model where cells are initially resistant to therapy (\( IniR \)) and in an alternative model where resistance is induced by treatment (\( Rind \)), in the subset of tumours where the presence of resistant cells was demonstrated.
Figure 6
ER content in the three resistance groups. Ordinate: ER content (fmol/ml); Abscissa: group 1: no resistance; group 2: resistance undetermined (data were fitted without resistance, but models including more than 10% resistant cells gave equivalent fits); group 3: resistance required (models including resistance significantly improved the fit). Short horizontal lines represent the average value in each resistance group.

a region of intermediate efficacy (scores 2–3) and high ER values; model type I without resistance was present in the region of intermediate or high efficacy and low ER, while model type I but with undetermined resistance (with the criteria specified above) grouped in the low sensitivity region (score 1), with a spread of ER values.

Among patients which were given chemotherapy after not responding to conventional hormone therapy, those with ER > 20 fmol/mg at the start of chemotherapy were classified as “resistant” (four) or “undetermined” resistance (one), while the six patients with ER < 20 fmol/mg all responded to chemotherapy (not shown).

Table 4 shows the relationship between resistance and the number of positive nodes. Ten of the 18 patients with no detectable drug resistance were node-negative and only three had more than two positive nodes, while in group 3 two out of nine were node-negative and six had more than two positive nodes.

Discussion
Despite the increasing knowledge of the molecular biology of tumours and their interactions with anticancer drugs, less is known about measuring the benefits of treatments in patients in terms of percentages of tumour cells killed and the proportion that is resistant. This study addresses these two questions during classical neoadjuvant chemotherapy in early breast cancer by interpreting a series of time-course data of tumour volumes, plus the measure of Ki67, using basic models of tumour growth, treatment and resistance. By fitting the tumour outlines, with sensitivity analysis, the range of cell killing and resistance parameters were estimated.

The importance of proliferation markers as predictors of treatment outcome in breast cancer management has been previously investigated, with controversial results [26]. A single marker, like Ki67 or apoptotic index, provides only a partial view of the underlying tumour kinetics, which is based on the interplay between cell cycling, quiescence and cell death (by apoptosis or other mechanisms). Furthermore, the different effects of treatment on quiescent and cycling cells, besides the presence of drug-resistant subpopulations, although theoretically recognised, is seldom taken into account in the analysis of out-

Table 4: Resistance vs. Number of positive nodes. Patients’ subsets on the basis of the presence of resistant cells and of the number of positive nodes.

<table>
<thead>
<tr>
<th>Presence of resistant cells</th>
<th>Nr. of nodes</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0</td>
</tr>
<tr>
<td>No (group 1)</td>
<td>10</td>
</tr>
<tr>
<td>Uncertain (group 2)</td>
<td>3</td>
</tr>
<tr>
<td>Yes (group 3)</td>
<td>2</td>
</tr>
<tr>
<td>Total pts:</td>
<td>15</td>
</tr>
</tbody>
</table>
come or in the design of chemotherapy regimens. This paper takes a step towards including all these variables in a comprehensive reconstruction of the dynamics of the response to treatment. This mathematical reconstruction extends the conventional assessment of clinical and pathological responses, which cannot recognise the details of a partial remission, to provide clues on a patient's sensitivity which are potentially useful to decide extension or changes of treatment.

The mathematical model used in this study performs better in reproducing the volume decrease during treatment than previous approaches [18] and explicitly includes cell cycle, quiescence and loss. This made the model more complex but also more realistic and allowed connections with the kinetic quantities that are – at least in principle – measurable in vivo, namely the doubling time, the potential doubling time and the growth fraction. We could then incorporate the measure of Ki67, as an estimate of the individual tumour growth fraction.

Because in most breast tumours only a minority of cancer cells are actually Ki67-positive and cycling [27], it seemed mandatory to include quiescent cells in the model [28]. As a consequence, two parameters of killing, one for cycling and the other for quiescent cells, were used in the fitting. However, because the data, including the Ki67 measure, were still not sufficient for an unequivocal assessment of tumour growth in the absence of treatment, we adopted a conservative strategy, examining a range of possibilities for each patient from fast to slow tumours. We repeated each fitting with 18 different growth models, combining representative values for each kinetic parameter. That choice was pursued with a conservative sensitivity analysis, including all growth models, to assess the uncertainty of estimates of parameters describing cell kill and resistance.

Our analysis identified different patterns of response during chemotherapy. A subset of patients was characterized by unequivocal presence of resistance, as the goodness of fit was significantly improved with the inclusion of a subpopulation of cells unaffected by the treatment. For this group of "resistant" patients, the average ER level was more than ten times higher than in the remaining patients. Conversely, six out of the seven patients with ER > 15, were in the "resistant" group, and the seventh had "undetermined" resistance. Instead the association between resistance and number of positive nodes was less striking. The association of the response to treatment with negative ER status has been reported by many others [29,30], but the association ER with the presence of resistant cells could not be directly demonstrated before. Instead, the cell kill parameter was not associated to ER status in our analysis, suggesting that drug efficacy on sensitive cells was unrelated to ER. One possible mechanism explaining the relationship between resistance and ER status is the co-expression of ER and the Bcl-2 proto-oncogene [31], which would be expected to reduce the rate of apoptosis.

A second subset of patients objectively identified by the model was characterized by a sensitivity shift, tumours becoming more drug-sensitive after the first cycle or even later. In this subset of patients the tumour volume at the beginning of treatment was not particularly high and the sensitivity shift occurred after a moderate debulking (not shown). Possibly in some instances the true cell number reduction was not reflected in a volume decrease in the first weeks so the drug effect was initially underestimated. Sensitivity in this group was therefore scored on the basis of the killing parameter estimated in the second, more effective, phase. This was a very sensitive subset of patients: eight out of nine had the highest cell killing scores (4 or 5 for cycling cells), reached in only three of the 25 patients without sensitivity shift (see Figure 7). The observation that sensitivity can increase after up to two months of low treatment efficacy, raises questions about the practice of automatically changing therapy after a poor initial response, and about the optimal number of cycles in some patients.

Considering the subset of patients characterised by high cell killing among cycling cells (score 4 or 5), treatment was also effective against quiescent cells in half cases, indicating the drugs were not exclusively active on cycling cells. However in the other half quiescence limited drug efficacy. These patients might have benefited from strategies aimed at mobilising quiescent cells, or suitably spaced cycles in order to avoid repeatedly hitting tumours with few target cycling cells, or from some entirely different strategy.

This study found that, by combining the time course of tumour volumes with a measure of Ki67, the estimates of parameter values were precise enough to permit a potentially useful and informative classification of the responses. However, that precision would be improved by using richer datasets, including additional measures of proliferation and cell death, not available for the present group of patients. For instance a direct estimate of Td or Tpot would have limited the uncertainty of the growth model. However, Tpot, is nowadays rarely measured in clinical routine and a measure of Td is probably even less feasible, requiring a second volume measurement before the start of treatment, e.g. at diagnosis, or during discontinuation of treatment, when planned in particular situations.
For what concerns cell death, apoptosis is usually measured on histopathological samples, using morphology, caspase activation or the terminal transferase dUTP nick end labelling (TUNEL) technique. This gives at least a qualitative assessment of whether some cells are killed, but unfortunately the relationship between the measured percentage of apoptotic cells and the percentage of killed cells remains unclear, as we do not know how long apoptotic cells remain visible. Thus, such measures of apoptosis are not valid quantitative measures of cell kill.

Another way to improve the precision is with a second evaluation of the growth fraction at the end of therapy, by Ki67 measures on the surgical specimen, or during therapy by fine-needle aspirates or core-cut biopsies. These measures are currently taken in trials of new neoadjuvant therapies, to detect changes in Ki67 as a marker of the response, in the absence of other strong short-term endpoints [32,33]. The mathematical approach would allow to interpret these data within a formal framework including all actors in play, i.e. cycling, quiescent and resistant cell subpopulations. In this respect, the effort made here to assess an objective method of parameter estimate could be exploited, and model’s kill rate categories, as well as detection of resistance, could become useful markers of outcome, at an intermediate level between the detection of inhibition of a target and the evaluation of survival.

Finally, new imaging techniques – CT, NMR, PET or even optical – are potential sources of data that can be interpreted by the model, as these techniques give more precise estimates of the overall tumour mass, and permit time-course evaluations for otherwise undetectable or unmeasurable tumours. However, additional research and modelling will be required to connect specific functional measures to the basic “phenomena” of proliferation, quiescence and loss which are the core of our representation of the tumour.

Conclusion
We presented here a new approach to the evaluation of chemotherapy using a mathematical model to interpret the time courses of tumour mass during treatment, including measures of the proliferative activity by Ki67. The model gives as output objective estimates of the fraction of cells killed by each cycle of treatment and of the fraction of resistant cells. The approach was proven feasible, providing items of information not otherwise available in pre-operative breast cancer chemotherapy. Additional data, particularly sequential Ki67 measures, could be added to the system, potentially reducing uncertainty in estimates of parameter values.

Competing interests
The authors declare that they have no competing interests.

Authors’ contributions
PU performed modelling, data analysis and drafted the manuscript; DC was involved in study design, performed clinical database management, contributed to data interpretation and manuscript preparation. All authors have read and approved the final manuscript.

Appendix

Appendix 1. The models of tumour growth, cell killing and resistance

The equations of the model of tumour growth shown in Figure 1 were thoroughly described by Bertuzzi, Gandolfi et al [21], thereafter referred as BG theory, assuming balanced exponential growth. From BG theory, neglecting for simplicity the distinction between G1, S and G2M phases and assuming “natural” cell loss only from quiescent cells, the density of proliferating cells of age “a” at time “t” is given by the formula

\[ n_p(a, t) = C e^{bt} e^{-\beta t} \]

where C is a suitable constant and b is the growth rate constant (\( b = \ln(2)/Tc \)).

By integration on age over the cell cycle time Tc, we obtain the number of proliferating cells:

\[ Np(t) = C e^{bt} (1-e^{-\beta Tc})/b \]

At a given time t, cells of age Tc end their cycle by division. Considering eq 1, their density is:

\[ g(t) = C e^{bTc} e^{bt} \]

It is useful to define the quantity \( u = g(t)/Np(t) \). In this way, using eq. 2 and eq. 3 we have:

\[ u = g(t)/Np(t) = b/(e^{bTc} - 1) \]

Notice that in the BG theory the following more general version of eq. 4 holds:

\[ u = g(t)/Np(t) = \beta/(e^{bTc} - 1) \]

where \( \beta = b + \mu (\beta = \alpha + \mu \) in the original notation used in [21]) and \( \mu \) is the cell loss rate of proliferating cells (assuming \( \mu_{G1} = \mu_{S} = \mu_{G2M} = \mu \) and thus \( \beta_{G1} = \beta_{S} = \beta_{G2M} = \beta \)). Eq. 4 bis can be readily demonstrated by combining the BG equation giving \( g(t) \) (eq. 13 in [21]) with BG equations giving \( Np(t) \) (eqs. 20 and 21 in [21], remembering that \( Np(t) = N_{G1}(t) + N_{S}(t) + N_{G2M}(t) \)).

Eq. 4 bis reduces to eq. 4 when neglecting “natural” cell loss in proliferating cells (\( \mu = 0 \)).
Summarizing the theory, we can say that in a small time and age interval \((dt = da)\):

i) \(u \cdot N_p(t) \cdot dt\) cells divide originating \(2 \cdot u \cdot N_p(t) \cdot dt\) newborn cells, where \(u\) is given by eq. 4;

ii) \(\theta \cdot 2 \cdot u \cdot N_p(t) \cdot dt\) newborn cell enter the proliferating status with zero age and

iii) \((1 - \theta) \cdot 2 \cdot u \cdot N_p(t) \cdot dt\) newborn cell enter the quiescent compartment.

Noticeably, the age-dependence of the theory is no more explicit in these relationships, being conveyed by the quantity "\(u\)" which depends only on \(T_c\) and \(T_d\).

Considering non-infinitesimal time intervals, \(\Delta\), the number of dividing cells in the interval \((t-\Delta, t)\) is given by:

\[
\int_{t-\Delta}^{t} u \cdot N_p(\tau) \, d\tau = u \cdot N_p(t-\Delta) \cdot \int_{0}^{\Delta} e^{b \tau} \, d\tau = u \cdot N_p(t-\Delta) \cdot z
\]

where:

\[
z = \frac{(e^{b \Delta} - 1)}{b} \tag{5}
\]

Similarly, the number of quiescent cells becoming proliferating or dying in the interval \((t-\Delta, t)\) is \(\gamma \cdot N_q(t-\Delta) \cdot z\) and \(\mu \cdot N_q(t-\Delta) \cdot z\) respectively.

The above definitions and equations allowed to simulate tumour growth by finite differences with time step \(\Delta\) (\(\Delta = 1\) day in the simulations presented in this paper), calculating the numbers of cycling \((N_p(t))\) and quiescent \((N_q(t))\) cells at time \(t\) from those at time \(t-\Delta\), in the absence of treatment. In this way it was possible to save a huge amount of computational time and to implement the model in a flexible and interactive spreadsheet program.

The resulting balance equations were the following:

\[N_p(t) = (\text{number of proliferating cells at time } t-\Delta) + (\text{newborn proliferating cells}) - (\text{cells which had divided}) + (\text{quiescent cells entered in the proliferative status})\]

\[= N_p(t-\Delta) + 2 \cdot \theta \cdot u \cdot N_p(t-\Delta) \cdot z - u \cdot N_p(t-\Delta) \cdot z + \gamma \cdot N_q(t-\Delta) \cdot z\]

\[N_q(t) = (\text{number of quiescent cells at time } t-\Delta) + (\text{newborn quiescent cells}) - (\text{dead quiescent cells}) - (\text{quiescent cells entered in the proliferative status})\]

\[= N_q(t-\Delta) + 2 \cdot (1 - \theta) \cdot u \cdot N_p(t-\Delta) \cdot z - \mu_q \cdot N_q(t-\Delta) \cdot z - \gamma \cdot N_q(t-\Delta) \cdot z\]

where \(u\) is given by eq. 4 and \(z\) by eq. 5.

\(z\) is close to \(\Delta = 1\) day, as \(b\Delta = \ln(2) \cdot \Delta/T_d \ll 1\), and allows to match exactly Td of the simulation with the theoretical Td during unperturbed balanced growth. After a treatment with differential efficacy \((S_p \neq S_q)\) the age distribution of proliferating cells will be unbalanced by quiescent cell entering the cycle if \(\gamma \neq 0\). In this case both \(u\) and \(z\) were approximated values, and some discrepancy of the simulation respect to a full age-dependent model is expected, for a short time after treatment. Because the interval between subsequent data points was seven days or more, this approximation can give only a small contribute to the errors of the estimate of the parameters.

The same growth equations were applied also to resistant cells \((N_{rp}(t)\) and \(N_{rq}(t))\).

Dying cells enter and exit three stages \((d_1, d_2, d_3)\) of death before being lost as follows:

\[N_{d_1}(t) = N_{d_1}(t-\Delta) - k \cdot N_{d_1}(t-\Delta)\]

\[N_{d_2}(t) = N_{d_2}(t-\Delta) + k \cdot N_{d_1}(t-\Delta) - k \cdot N_{d_2}(t-\Delta)\]

\[N_{d_3}(t) = N_{d_3}(t-\Delta) + k \cdot N_{d_2}(t-\Delta) - k \cdot N_{d_3}(t-\Delta)\]

The overall number of dying-not-yet-lost cells is given by the sum of the cells in the three stages:

\[N_d(t) = N_{d_1}(t) + N_{d_2}(t) + N_{d_3}(t)\]

The overall number of tumour cell at a time "\(t\)" is the sum of sensitive cycling, sensitive quiescent, resistant cycling, resistant quiescent and dying cells, namely:

\[N(t) = N_p(t) + N_q(t) + N_{rp}(t) + N_{rq}(t) + N_d(t)\]

\(N(t)\) is the quantity compared with measured tumour volumes, via a proportionality constant. At the beginning of the treatment we have:

\[N_p(0) = N(0) \cdot GF \cdot (1 - IniR)\]

\[N_q(0) = N(0) \cdot (1 - GF) \cdot (1 - IniR)\]

\[N_{rp}(0) = N(0) \cdot GF \cdot IniR\]

\[N_{rq}(0) = N(0) \cdot (1 - GF) \cdot IniR\]

where \(GF\) is the growth fraction, estimated by %Ki67+, and \(IniR\) represents the fraction of cells initially resistant to the drugs.
At the times of treatment, the situation immediately before \((t-)\) is considered separately from that immediately after \((t+)\) the treatment and the number of -surviving- cycling and quiescent cells is reduced as follows:

\[ Np(t+) = Np(t-) \cdot Sq \]
\[ Nq(t+) = Nq(t-) \cdot Sp \]

where \(Sp\) and \(Sq\) are the fraction of cells surviving the treatment, while non-surviving cells enter the first stage of dying cells:

\[ Nd_{1}(t+) = Nd(t-) + Np(t-) \cdot (1-Sp) + Nq(t-) \cdot (1-Sq) \]

When considering drug-induced resistance, the equations of surviving cells become:

\[ Np(t+) = Np(t-) \cdot Sp \cdot (1-Rind) \]
\[ Nq(t+) = Nq(t+) \cdot Sq \cdot (1-Rind) \]
\[ Np(t+) = Np(t-) \cdot Sp \cdot (1-Rind) \]
\[ Nq(t+) = Np(t-) \cdot (1-Sp) \cdot Sp \cdot Rind + Np(t-) \cdot (1-Sq) \cdot Sq \cdot Rind \]

where \(Rind\) represents the fraction of – surviving – cells which become resistant as a consequence of the treatment.

The contribution of spontaneous mutations to a resistant phenotype during the 100 days of treatment was considered negligible [8].

Because the drugs were given contemporaneously, the effect of each of them cannot be evaluated separately. Thus \(Sp\) and \(Sq\) measure the effect of the combined treatment. Similarly, cells resistant to single drugs could be identified, and a single subpopulation of cells “resistant to treatment” was considered.

Because the same dosage was given each time to the patients, the same \(Sp\) and \(Sq\) were repeatedly applied on days 0, 2, 3, 4, 5, 8, 15, 30, reproducibly the true schedule of this study. In few instances, a more complex model was needed, shifting of the values of \(Sp\) and \(Sq\) to new \(Sp2\) and \(Sq2\) values in the course of the treatment.

The model is about numbers of tumour cells (\(N\)), while the data are tumour mass (\(V\), volume), including non-cancerous cells and tissues. Nevertheless, in the absence of specific information about non tumour cells (at each time and for each patient) we assumed proportionality between \(N\) and \(V\), through the equivalence \(1 \text{ cm}^3 = 10^9\) tumour cells. The specific value of the proportionality constant does not affect the results.

\section*{Appendix 2. Selection of tumour growth types}

In order to simplify the optimisation procedure, we fixed \(Td, Tc, \gamma\), to representative values. This choice was justified by a preliminary study on our dataset, indicating that wide changes of these parameters only slightly modified the fits (not shown).

Representative values of the growth parameters were chosen as follows:

\(Td\). Reports of doubling time of breast cancers indicate an average between 100 and 200 days, increasing with the age of the patient [34]. In the statistics of Spratt [35] only 1% have \(Td < 30\) gg. Thus we considered the values \(Td = 30, 150\) and 10000 days, as representative of fast, average, slow tumour, respectively.

\(Tc\). The parameter represents the average length of the non-G0 part of the cycle (not to be confused with estimates of other reports [36] where quiescent cells were not considered a part). Thus \(Tc\) values usually found in cell lines \textit{in vitro} (1–2 days) are a reasonable lower boundary. However such short \(Tc\) are not consistent in tumours with moderately high \(GF\), unless accepting very high natural cell loss. We considered the values \(Tc = 2, 5\) and 8 days, as representative of short, average, long cell cycle.

\(\gamma\). The value of \(\gamma\) is in part automatically constrained by the values of the other parameters of the model of tumour growth. It is also the reciprocal of the mean residence time in the quiescent status. We consider two extreme values: 0, as representative of a tumour with negligible recycling from quiescence into the cycling stage, and 0.01, i.e. 1% quiescent cells becoming cycling per day, corresponding to an average residence time in G0 of 100 days.

Combining the values of \(Td = 30, 150, 10000\) days, \(Tc = 2, 5, 8\) days and \(\gamma = 0, 0.01\) we obtained eighteen different types (type1: \(Td = 30, Tc = 2, \gamma = 0\); type2: \(Td = 150, Tc = 2, \gamma = 0\); etc.) representative of tumour breast cancer growth.

For each tumour growth type, given the value of \(GF\) provided by \%Ki67+, the theory [21], with cell loss only within quiescent cells, allowed to calculate additional kinetics characteristics of the tumour, namely the potential doubling time and the rate of natural cell loss, using the following formulae:

\[ T_{pot} = (Td/GF) \cdot (e^{\ln(2)Tc/Td} - 1) \quad \text{(derived from eq. 30 in [21])} \]
\[ \mu_{n} = \ln(2) \cdot (1/T_{pot} - 1/Td)/(1 \cdot GF) \quad \text{(derived from eqs. 19 and 29 in [21])} \]
Some combinations of $T_d$, $T_c$, $G_F$, $\gamma$ were not biologically consistent because mathematically they would require a negative cell loss. For what concerns $T_{pot}$ we referred to the $T_{pot}$ estimates obtained with BrdU \textit{in vivo} in breast cancer patients by Rew and Wilson [37]. Because in that database the highest $T_{pot}$ value was 50 days, we conservatively accepted a combination of parameters as biologically consistent if $T_{pot} < 75$ days.

Thus, for each patient, only a subset of the eighteen types was considered for fitting, those with $T_{pot} > 75$ days or $\mu_i < 0$ (if any) being excluded as biologically not consistent.

The data were in general poorly sensitive to the values adopted for the tumour growth parameters. In $19/35$ cases, we found (Table 5) a fit statistically equivalent to the best with $T_d = 30$, $150$ and $10000$ days. In six instances only fast growing models were compatible with the data, while in another nine fast growth was excluded.

As concerns $T_c$ (Table 6), only in a minority of cases do the data indicate that two days or eight days should be preferred. In all the other cases the value remained uncertain.

The recycling rate remained undetermined in $22/35$ instances (not shown), while for the reminder the fit indicated $\gamma = 0$.

### Table 5: Doubling time compatible with data

<table>
<thead>
<tr>
<th>Category</th>
<th>Range</th>
<th>n</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>ND</td>
<td>30–150–10000</td>
<td>19</td>
<td>54.3</td>
</tr>
<tr>
<td>Slow tumour</td>
<td>150–10000</td>
<td>9</td>
<td>25.7</td>
</tr>
<tr>
<td>Fast/Intermediate</td>
<td>30–150</td>
<td>1</td>
<td>2.9</td>
</tr>
<tr>
<td>Fast tumour</td>
<td>30</td>
<td>6</td>
<td>17.1</td>
</tr>
</tbody>
</table>

$$
\theta = 0.5 \cdot \frac{e^{\ln(2)T_d/T_c} - 0.5 \cdot \gamma \cdot (2 - e^{\ln(2)T_d/T_c})}{(\mu_i + \ln(2)/T_d)}
$$

(derived from eq. 15 in [21])

### Table 6: Cell cycle time compatible with data

<table>
<thead>
<tr>
<th>Category</th>
<th>Range</th>
<th>n</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>ND</td>
<td>2-5-8</td>
<td>4</td>
<td>11.4</td>
</tr>
<tr>
<td>Long Tc</td>
<td>8</td>
<td>9</td>
<td>25.7</td>
</tr>
<tr>
<td>Intermediate/Long Tc</td>
<td>5–8</td>
<td>14</td>
<td>40.0</td>
</tr>
<tr>
<td>Short/Intermediate Tc</td>
<td>2–5</td>
<td>2</td>
<td>5.7</td>
</tr>
<tr>
<td>Short Tc</td>
<td>2</td>
<td>6</td>
<td>17.1</td>
</tr>
</tbody>
</table>

### Additional material

**Additional file 1**

PAOtherapyX. Microsoft Excel program used for the simulation and fitting of tumour reduction during chemotherapy

Click here for file [http://www.biomedcentral.com/content-supplementary/1471-2407-8-201-S1.xls]

### Acknowledgements

The generous contribution of the Italian Association for Cancer Research and the Nerina and Mario Mattioli Foundation is gratefully acknowledged.

### References


Pre-publication history
The pre-publication history for this paper can be accessed here:
http://www.biomedcentral.com/1471-2407/8/201/prepub

Publish with BioMed Central and every scientist can read your work free of charge
"BioMed Central will be the most significant development for disseminating the results of biomedical research in our lifetime."
Sir Paul Nurse, Cancer Research UK

Your research papers will be:
- available free of charge to the entire biomedical community
- peer reviewed and published immediately upon acceptance
- cited in PubMed and archived on PubMed Central
- yours — you keep the copyright

Submit your manuscript here:
http://www.biomedcentral.com/info/publishing_adv.asp