Preprints of the 12th IFAC Symposium on Computer Applications in Biotechnology
The International Federation of Automatic Control
16-18, 2013, December. Mumbai, India

Quantifying Heterogeneity of Cell Death ⋆

Monica Schliemann * **Samuel Livingstone** §
Mark C. Readman ‡ **Dimitrios Kalamatianos** †
Eric Bullinger *,¶

* *Department of Electrical Engineering and Computer Science (Institut Montefiore) and GIGA (Interdisciplinary Cluster for Applied Geno-proteomics), Université de Liège, Liège, Belgium (monica.schliemann@systems-biology.eu, E.Bullinger@ulg.ac.be)*
§ *Off The Fence, Bristol, United Kingdom (samuelilivingstone@gmail.com)*
‡ *Stockport College, Stockport, United Kingdom and School of Mathematics, The University of Manchester, Manchester, United Kingdom (mark.readman@gmail.com)*
† *Division of Developmental Biology, Biomedical Research Foundation of the Academy of Athens, Athens, Greece (dkalam@bioacademy.gr)*
¶ *Hamilton Institute, National University of Ireland Maynooth, Maynooth Co. Kildare, Ireland*

Abstract: Heterogeneity is a common property of biological signalling systems. In most cases, heterogeneity is described qualitatively, only for some classes of responses have qualitative measures been proposed. For cell death signalling, this paper is the first to propose a quantification of heterogeneity. The challenge hereby is the dual aspect of heterogeneity. First, only part of the cell population may die while the others survive a specific death stimulus. Second, the time of death can vary from cell to cell.
The proposed heterogeneity measure is based on an $\mathcal{L}_1$ measure of the deviation between a homogeneous response and the population cumulative density function, on a nonlinearly scaled time. This measure allows for a quantitative study of the dependency of the heterogeneity of the responses to different stimulus doses or parametric variations. This will for example enable sensitivity analyses of heterogeneity.
The heterogeneity measure is illustrated by applying it to two different published cell ensemble models of apoptosis signalling, each having approximately 50 states and over 100 kinetic parameters. This analysis reveals that heterogeneity is more pronounced at an intermediate range of doses. In other words, high doses or low ones yield more homogeneous responses in the cell population.

Keywords: systems biology, polynomial models, robustness, perturbation analysis, biochemical systems.

1. INTRODUCTION

An important property of biological systems is that the underlying cells are heterogeneous. Heterogeneity of apoptotic signalling system models is usually specified by varying protein concentrations across the cell population, [Spencer et al., 2009; Schliemann et al., 2011], as have been measured experimentally. These distributions affect both the proportion of cells dying in response to a specific death stimulus and the distribution of when dying cells actually die.

This paper proposes a quantitative measure of heterogeneity applicable to both experimental and simulation data of heterogeneous cell population responses.

⋆ This work was supported in part by the Francqui Foundation and the Belgian Network DYSCO (Dynamical Systems, Control, and Optimisation), funded by the Interuniversity Attraction Poles Programme, initiated by the Belgian State, Science Policy Office.

The manuscript is organised as follows. First, Section 2 gives a brief introduction to the biological implication of the present study as well as of the two mathematical models used as test cases. Section 3 presents the proposed measure of heterogeneity, while Section 4 applies it to both mathematical models. The paper finishes with the Conclusions in Section 5.

2. BACKGROUND

This section gives a brief description of the modelling of apoptosis signalling in heterogeneous cell populations as well as of the two mathematical models the proposed method is applied to.

2.1 Heterogeneity in biological systems

The advancement of single cell experimental techniques such as flow cytometry and fluorescent microscopy has

revealed that cell-to-cell variability is essential for a number of biological phenomena. Examples include adaptivity [Gatenby et al., 2009], immune response [Hawkins et al., 2007], stress response [Yeyati and Heyningen, 2008], differentiation and development [Tomlin and Axelrod, 2007], cancer cell drug resistance [Cohen et al., 2008; Gatenby et al., 2009], *Salmonella* infection [Restif et al., 2013] and programmed cell death [D'Herde et al., 2000; J. Schmid et al., 2012; Spencer et al., 2009; Schliemann et al., 2011].

Several recent papers quantify heterogeneity in biology systems, using standard deviation or the width of e.g. the central 90% of the distribution [Brooks and Grigsby, 2013; Cicone et al., 2013; Maruoka et al., 2013; Cicone et al., 2013].

However, heterogeneity in signalling leading to programmed cell death can have two outcomes. First, only a sub-population may die. Second, the time at which death occurs varies from cell to cell. Therefore, heterogeneity of apoptosis cannot be quantified by the standard deviation of the distribution of time of death. Currently, no quantitative measure of heterogeneity of cell death exists.

2.2 Biological modelling

To correctly describe biological systems with cell-to-cell variability and all-or-none or oscillatory behaviour requires an appropriate modelling framework, such as cell ensemble modelling, which combines a large number of individual models that differ in some key parameters, see Figure 1. The cell ensemble modelling approach has been already successfully applied to heterogeneous cultures of *E. coli* growth [Domach et al., 1984], yeast metabolic oscillations [Henson, Müller and Reuss, 2002], apoptosis in a mammalian cell batch culture [Meshram et al., 2012], as well as Fas- [Toivonen et al., 2011] or TRAIL-induced apoptosis [Spencer et al., 2009].

Applying cell ensemble modelling to systems of biochemical reactions modelled by the law of mass action and neglecting spatial and stochastic effects, leads to a systems of ordinary differential equations. We will use here the notation Horn and Jackson, 1972; Feinberg, 1995; Gatermann and Huber, 2002. A model of a biochemical signalling network consists of a set of r reactions of the following form

$$\alpha_1 \mathrm{S}_1 + \ldots + \alpha_{n_s} \mathrm{S}_{n_s} \to \beta_1 \mathrm{P}_1 + \ldots + \beta_{n_p} \mathrm{P}_{n_p},$$

where S_i denotes substrates that are transformed into the products P_i. The factors α_i and β_i denote the stoichiometric coefficients of the reactants. All substrate and product concentrations are collected intro a vector $c \in \mathbb{R}^n$. Let us denote by $x \in \mathbb{R}^m$ the vector of complexes, i.e. all left and right hand sides of the reactions and by $B \in \mathbb{R}^{m \times n}$ the corresponding incidence matrix, such that

$$c = Bx.$$

By the definition of the complexes, any reaction of the model maps exactly one complex to another one. The corresponding incidence matrix is I_a. Then, the model can be described by

$$\frac{d}{dt} c = B I_a \operatorname{diag}(k) \theta_B(c)$$

where

$$\theta_{Bi}(c) = c_1^{B_{i1}} \cdot c_2^{B_{i2}} \cdot \ldots c_n^{B_{in}}$$

and k_i denotes the reaction rate of the i-th reaction.

2.3 Apoptosis

Apoptosis is an important form of programmed cell death utilised to remove damaged or unneeded cells within an organism. Therefore, its regulation is very important and misregulation of apoptosis can lead to severe pathologies. Reduced apoptosis is for example present in cancer, autoimmune diseases or viral infections, while enhanced cell death is observable in neurodegenerative diseases or AIDS.

The precise timing of the response of cell populations to an apoptotic stimulus is usually heterogenenous. Below we present two models incorporating this population heterogeneity [Aldridge et al., 2011; Schliemann et al., 2011] In both, the heterogeneity is quantified in Section 4.

2.4 Aldridge model

Aldridge et al. [2011] propose a model of TRAIL-induced apoptotic signalling in HeLa cells. The model code is available as supplementary material of [Aldridge et al., 2011].

The model consists of 57 species in 2 compartments, cytoplasma and mitochondrion. These species participate in 100 reactions, of which 46 are reversible and 54 irreversible. Of the 146 flows, 25 are quadratic, i.e. involve two species, 103 are linear and 18 constant. There is a permanent turnover of the 17 proteins of the cell. In the cell population, their production rates are varied across the population to follow lognormal distributions and simulated using a cell ensemble model. Each cell is initialised in the steady state, adjusted to cope with the modified protein production rates.

Cell death corresponds to a sufficiently large activation of Caspase-3, and is defined as the time at which the Caspase-3 substrate PARP drops below a threshold of 50% of its initial value.

2.5 Schliemann model

Schliemann et al. [2011] proposed a model of TNF-induced pro- and anti-apoptotic signalling in KYM-1 cells, a human rhabdomyosarcoma derived cell line. The model code is available as supplementary material of Schliemann et al. [2011], as well as on `biomodels.org`.

The model consists of 47 species, 89 complexes and 106 kinetic parameters (70 irreversible and 18 reversible reactions). Five flows involve three species, 38 are quadratic and 45 are linear in the species. The remaining ones are constant flows.

Again, cell death corresponds to a sufficiently large activation of Caspase-3, i.e. the time point at which the Caspase-3 substrate PARP drops below half its initial value.

The Schliemann model has a permanent turnover of the 19 proteins of the cell. Their production rates are varied

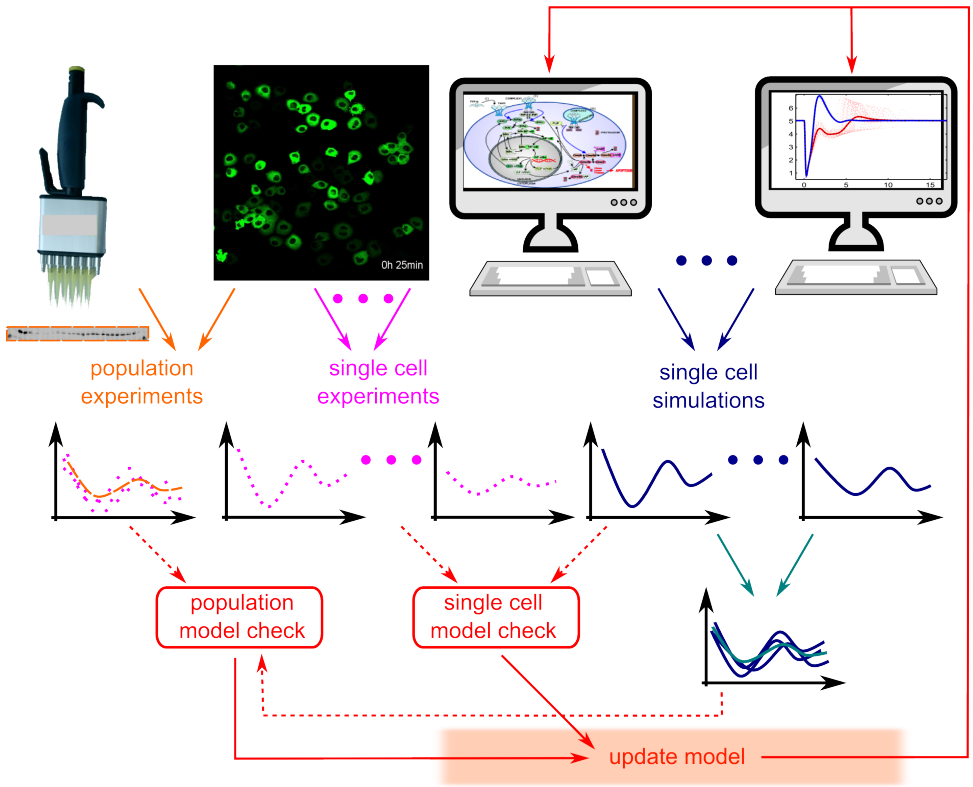

Figure 1. Cell Ensemble Model Development. Based on prior knowledge, a single cell model is designed. Its simulations are compared to single cell experimental data (top). If they do not agree, the model needs to be adjusted by varying kinetic parameters or modifying the model structure. A cell population model is obtained by simulating an ensemble of single cell models. The individual cell models are identical, except for the numerical values of some parameters, which vary from one cell to the other according to some distribution. As for the single cell level, a comparison of cell ensemble simulations with experimental data (bottom) might require a recalibration of the model, which can be achieved on the single cell level or on the cell ensemble level. The model development is finished when simulations and experiments agree on both single cell and cell population level.
One aspect of the population model check is the degree of heterogeneity. The here proposed criterion gives a quantitative assessment of heterogeneity for model and data analysis.

across the population to follow lognormal distributions and simulated using a cell ensemble model. Each cell is initialised in the steady state without caspases, i.e. adjusted to cope with the modified protein production rates.

3. MEASURE OF HETEROGENEITY

Analysing heterogeneous cell population models such as those presented in the previous section is challenging as the heterogeneity depends on both a Boolean output (life or death for a certain stimulus) and a continuous one (time of death). To better quantify heterogeneity of the cell population response, we needed a measure accounting for heterogeneity in the sense of survival rate versus death rate combined with the timing of death.

In economics, Lorenz curve and Gini coefficient are commonly used to quantify homogeneity of the wealth distribution [Schader and F. Schmid, 1994]. This is illustrated in Figure 2. The Gini coefficient quantifies heterogeneity, it is defined as the area between the Lorenz curve $L(\cdot)$ and the bisecting line, scaled by the maximal area, i.e.

$$\mathrm{GC} = 2\int_0^{x_{100\%}} \big(x - L(x)\big)dx.$$

Thus, a Lorenz curve on the bisecting line yields a Gini coefficient of 0%. This corresponds to equal wealth distribution. A Gini coefficient of 100% is obtained by a Lorenz curve that follows the horizontal axis between 0 and 100% and then jumps to 100%.

For programmed cell death, a homogeneous population has all deaths occurring simultaneously. Thus, if we plot the cumulative survival rate of the population, also denoted the cell viability $v(\cdot)$, versus time, a homogeneous population has a step like curve. Any deviation from this is due to inhomogeneity of the population. To obtain a finite time axis, the x-axis is compressed using the nonlinear mapping

$$t \mapsto \tanh \frac{t}{\tau}$$

with a positive scaling time τ. Then, find the time of death of the median cell, $t_{50\%}$. For a homogeneous population, the cell viability $v_{\mathrm{hom}}(\cdot)$ is 100% for $t < t_{50\%}$ and 0% for $t > t_{50\%}$. As measure of heterogeneity, we therefore propose the L1-norm of the difference between $v(\cdot)$ and $v_{\mathrm{hom}}(\cdot)$, i.e. the area above the cell viability in the interval $(0\ t_{50\%})$ plus the one below in the interval $(t_{50\%}\ \infty)$, normalised by the maximal area, which is equal to $\frac{1}{2}$:

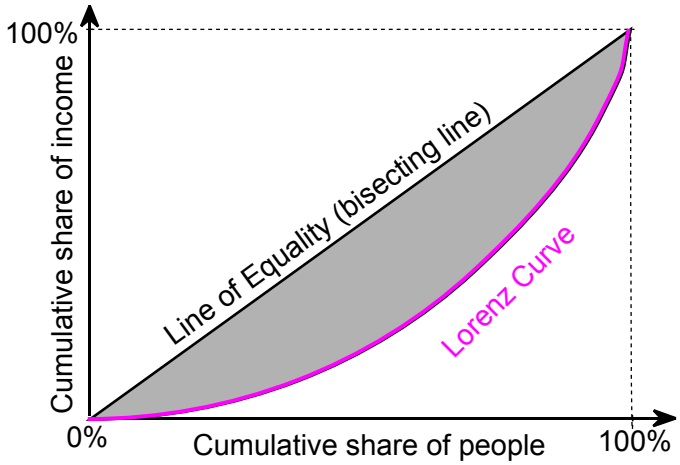

Figure 2. Illustration of Lorenz curve and Gini coefficient. The Lorenz curve is the cumulative share of income as a function of the cumulative share of people. The closer the Lorenz curve is to the diagonal, the more the wealth is evenly distributed.
The Gini coefficient is a number between 0 and 1 quantifying the homogeneity: It is twice the grey area, i.e. the difference between Lorenz curve and diagonal. Therefore, a Gini coefficient of 0 corresponds to an even income distribution, while a 1 stands for maximal inhomogeneity.

$$\mathcal{M}_{\mathrm{H}} = 2\int_0^{\frac{1}{2}} \big(100\% - v(\xi)\big)dx + 2\tau\int_{\frac{1}{2}}^{1} v(\xi)dx$$

with $\xi = \tau\,\mathrm{atanh}(x)$, or, in a more compact notation,

$$\mathcal{M}_{\mathrm{H}} = 2\left\| \min\left\{1 - v(\xi),\ v(\xi)\right\}\right\|_{\mathcal{L}^1([0\ 1])}.$$

In the original time coordinates

$$t = \tau\tanh(x),$$

the equivalent definition is

$$\mathcal{M}_{\mathrm{H}} = 2\int_0^{\tau} \frac{\mathrm{sech}^2(\frac{t}{\tau})}{\tau}\big(100\% - v(t)\big)dt + 2\int_{\tau}^{\infty} \frac{\mathrm{sech}^2(\frac{t}{\tau})}{\tau} v(t)dt$$

This definition ensures that the measure of heterogeneity lies in the closed interval [0, 1] and that a homogeneous population has a measure of 0. The parameter τ allows for a weighting of survival rate versus timing of death for dying cells in the quantification of heterogeneity. For populations where at least half the population ultimately dies, a useful choice for τ is $\tau = t_{50\%}$, the time of death of the median cell.

Figure 3 illustrates graphically the heterogeneity measure. The plot shows the cell viability $v(\cdot)$ (magenta line) as a function of time, on a time axis scaled by $t \mapsto \tanh\frac{t}{t_{50\%}}$ where $t_{50\%}$ is the time of death of the median cell.

In a homogeneous population, all cells die simultaneously. Thus, the cell viability curve of a homogeneous population is constantly 100% until the common time of death, where the curve drops directly to 0%. Any deviation from this

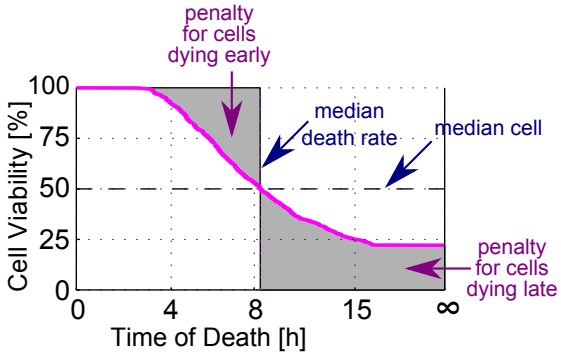

Figure 3. Illustration of the Heterogeneity Measure. The viability of the median cell is denoted by the solid black line: before its time of death the cell lives (100% viable), afterwards it is dead (0% viable). The percentage of living cells within the cell population is the magenta line. Obviously, it crosses 50% at the time of death of the median cell $t_{50\%}$. The abscissa is a normalised time axis, such that 0 h is a 0, $t_{50\%}$ at $\frac{1}{2}$ and ∞ at 1. The difference between median cell and cell population is highlighted in grey. The sum of the grey surfaces, normalised by $\frac{1}{2}$, is the measure of heterogeneity. The scaling of $\frac{1}{2}$ is used as this is exactly the maximal area of the grey surfaces.

step is due to heterogeneity. In Figure 3, these deviations are highlighted by the grey background. The heterogeneity measure HM combines a penalty for cells dying earlier than the median cell as well as for those dying later.

If the majority of cells in the population survives, the above definition for $t_{50\%}$ is not possible. Then, it is necessary to scale the time axis by a user-defined value, for example half the time interval of interest.

4. RESULTS

The heterogeneity measure proposed above is now applied to the two models of ligand-induced apoptosis. First, the measure is applied to the Aldridge model, then to the Schliemann model.

4.1 Aldridge model

The heterogeneity of the Aldridge model was analysed for nine different stimulus doses, spanning five orders of magnitude. Figure 4 shows the heterogeneity surfaces graphically. Increasing the dose moves the time of death of the median cell forward in time. The resulting heterogeneity measure is however very similar in all cases, see Table 1.

Table 1. Heterogeneity measures of the Aldridge model for nine different doses of stimuli. The unit mpc stands for molecules per cell.

TRAIL [mpc]	HM	TRAIL [mpc]	HM
10	0.24	1000	0.27
30	0.49	3000	0.27
100	0.33	10 000	0.27
300	0.28	30 000	0.28
		100 000	0.28

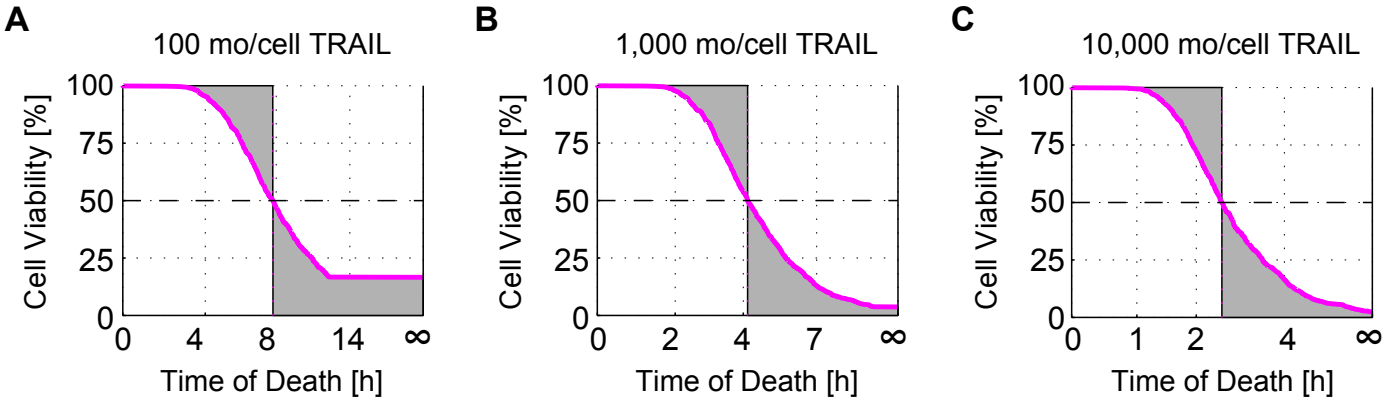

Figure 4. Heterogeneity analysis of the Aldridge model stimulated with three different TRAIL concentrations. A-C: The cell viability of the cell population is plotted versus time as magenta line. The deviation between the population cell viability and that of the median cell is highlighted by grey areas. The normalised area of grey is the corresponding heterogeneity measure. A: 100, B: 1000 and C: 10 000 molecules per cell of TRAIL

.

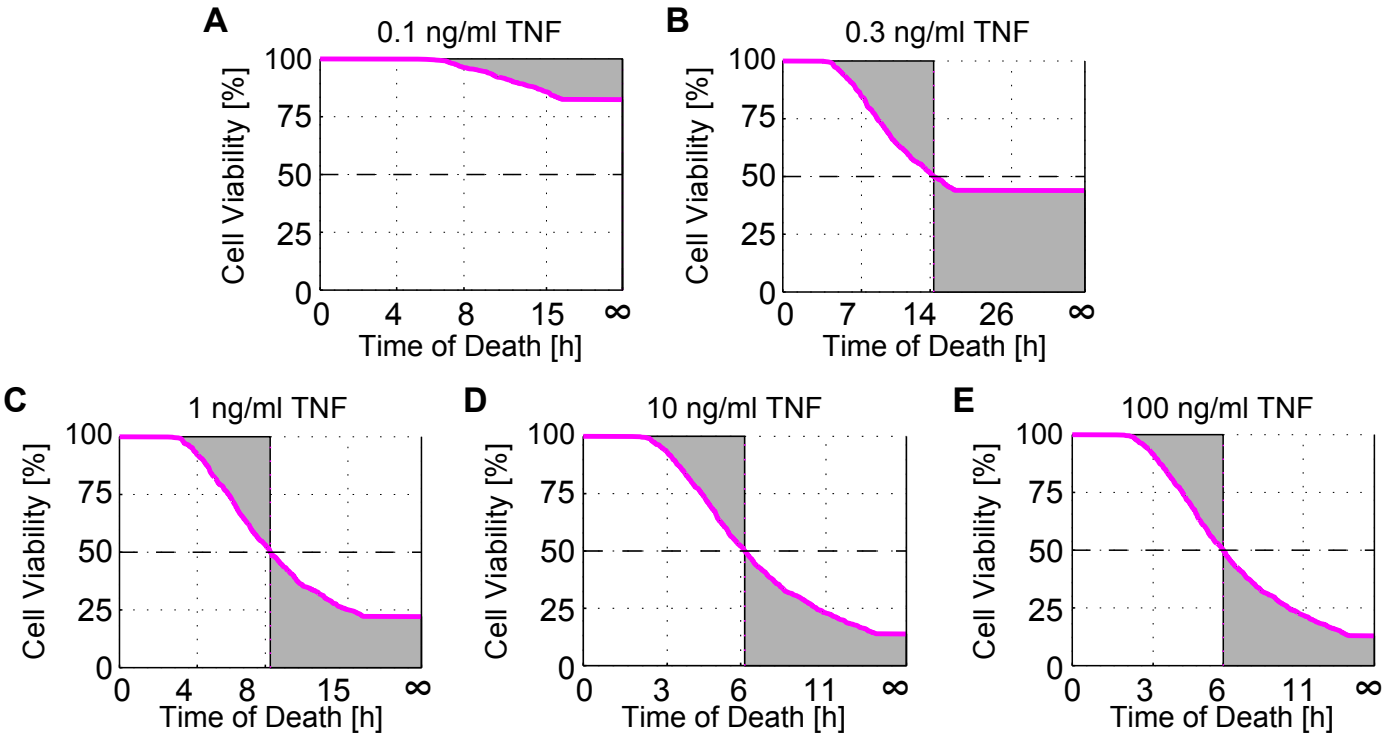

Figure 5. Heterogeneity analysis of the Schliemann model stimulated with different concentrations of TNF. A-F: The cell viability of the cell population is plotted versus time as magenta line. The deviation between the population cell viability and that of the median cell is highlighted by grey areas. The normalised area of grey is the corresponding heterogeneity measure. A: 0.10, B: 0.30, C: 1, D: 3, E: 10 and F: 100 ng/ml TNF.

4.2 Schliemann model

The heterogeneity of the Schliemann model is now quantified for TNF stimuli ranging five orders of magnitude, as depicted in Figure 5. The numerical values of the heterogeneity measures can be found in Table 2. For high doses of TNF, the responses are similarly heterogeneous, with HM around 0.4. Interesting is a TNF dose around 0.30 ng/ml, where the heterogeneity is very high. The corresponding plot (Figure 5B) reveals this is partly due to the high survival rate in combination with some cells dying at approximatively half the time of death of the median cell.

For low stimuli such as 0.10 ng/ml, most cells of the Schliemann model survive, see Figure 5 A, or the star-marked HM in Table 2. In these cases, the automatic choice of $\tau == t_{50\%}$ is not possible. The obtained heterogeneity measure thus depends on the choice of τ. For very low τ, the heterogeneity measure is close to 0. Increasing τ results in larger measures. In the limit of very large τ, the heterogeneity measure approaches $2 * (100\% - v_\infty)$, where v_∞ is the cell viability for very large times. This shows that for populations with most cells surviving the heterogeneity measure can take arbitrary values between

Table 2. Heterogeneity measures of the Schliemann model for different doses of stimuli. HM marked by a star * indicate that some, but less than 50% of the cell population dies. There, $\tau = t_{50\%}$ is not possible and has been replaced by $\tau = 8.50\,\mathrm{h}$, half the maximal simulation time. A larger τ increases these HM up to the value in bracket.

TNF[ng/ml]	HM	TNF[ng/ml]	HM
0.01	0	1	0.44
0.03	0.0060* (≤ 0.018)	3	0.40
0.10	0.14* (≤ 0.35)	10	0.40
0.30	0.62	30	0.39
		100	0.39

0 and a maximal one. For $v_\infty = 50\%$ and very large τ, HM approaches one.

4.3 Summary

As could be expected from the cell ensemble modelling, both analysed models show heterogeneous responses. For relatively large doses of stimulation ligand, the heterogeneity of the response does not depend very much on the specific dose. This is in contrast to the time of death of the median cell, which decreases with increasing dose.

While the Aldridge model shows a relative low dosage dependence of the heterogeneity, the Schliemann model is more heterogeneous for low doses than for high ones. This is in part due to more cells surviving, but also to a higher variability of the time of death among dying cells.

5. CONCLUSIONS AND OUTLOOK

This paper proposes a quantitative measure of heterogeneity that is applicable to models of apoptosis. By scaling the time axis onto a finite interval, heterogeneity can be defined as via integration of cell viability curves. The heterogeneity can also be calculated based on experimental cell viability data, thus no mathematical model is required for calculating heterogeneity. The measure's free parameter τ can be automatically fixed if at least half the population dies by chosing τ as the time of death of the median cell.

The proposed measure allows for the quantitative analysis of models and the study of the sensitivity of heterogeneity in models of apoptosis as well as the impact of perturbations such as knock-outs on the heterogeneity of the system response.

REFERENCES

B. B. Aldridge, S. Gaudet, D. A. Lauffenburger and P. K. Sorger (2011). Lyapunov exponents and phase diagrams reveal multi-factorial control over TRAIL-induced apoptosis. *Mol Syst Biol* 7, 553–553.

F. J. Brooks and P. W. Grigsby (2013). Quantification of heterogeneity observed in medical images. *BMC Med Imaging* 13.1, 7.

F. Cicone et al. (2013). Quantification of dose nonuniformities by voxel-based dosimetry in patients receiving 90Y-ibritumomab-tiuxetan. *Cancer Biother Radiopharm* 28.2, 98–107.

A. A. Cohen et al. (2008). Dynamic proteomics of individual cancer cells in response to a drug. *Science* 322.5907, 1511–1516.

K. D'Herde et al. (2000). Ultrastructural localization of cytochrome c in apoptosis demonstrates mitochondrial heterogeneity. *Cell Death Differ* 7.4, 331–337.

M. M. Domach, S. K. Leung, R. E. Cahn, G. G. Cocks and M. L. Shuler (1984). Computer model for glucose-limited growth of a single cell of *Escherichia coli* B/r-A. *Biotechnol Bioeng* 26.9, 1140–1140.

M. Feinberg (1995). The existence and uniqueness of steady states for a class of chemical reaction networks. *Archive for Rational Mechanics and Analysis* 132.4, 311–370.

R. A. Gatenby, A. S. Silva, R. J. Gillies and B. R. Frieden (2009). Adaptive therapy. *Cancer Res* 69.11, 4894–4903.

K. Gatermann and B. Huber (2002). A family of sparse polynomial systems arising in chemical reaction systems. *Journal of Symbolic Computation* 33.3, 275–305.

E. D. Hawkins, M. L. Turner, M. R. Dowling, C. van Gend and P. D. Hodgkin (2007). A model of immune regulation as a consequence of randomized lymphocyte division and death times. *Proc Natl Acad Sci U S A* 104.12, 5032–5037.

M. A. Henson, D. Müller and M. Reuss (2002). Cell population modelling of yeast glycolytic oscillations. *Biochem J* 368.Pt 2, 433–446.

F. J. M. Horn and R. Jackson (1972). General mass action kinetics. *Archive for Rational Mechanics and Analysis* 47.2, 81–116.

Y. Maruoka et al. (2013). Quantification of heterogeneity on 201Tl gated SPECT: evaluation of coronary artery disease. *Clin Nucl Med* 38.1, 7–12.

M. Meshram et al. (2012). Population-based modeling of the progression of apoptosis in mammalian cell culture. *Biotechnol Bioeng* 109.5, 1193–1204.

O. Restif et al. (2013). Quantification of the effects of antibodies on the extra- and intracellular dynamics of *salmonella enterica*. *J R Soc Interface* 10.79, 20120866.

M. Schader and F. Schmid (1994). Fitting parametric Lorenz curves to grouped income distributions — a critical note. *Empirical Economics* 19.3, 361–370.

M. Schliemann et al. (2011). Heterogeneity reduces sensitivity of cell death for TNF-stimuli. *BMC Syst Biol* 5, 204.

J. Schmid et al. (2012). Systems analysis of cancer cell heterogeneity in caspase-dependent apoptosis subsequent to mitochondrial outer membrane permeabilization. *J Biol Chem* 287.49, 41546–41559.

S. L. Spencer, S. Gaudet, J. G. Albeck, J. M. Burke and P. K. Sorger (2009). Non-genetic origins of cell-to-cell variability in TRAIL-induced apoptosis. *Nature* 459.7245, 428–432.

H. T. Toivonen et al. (2011). Modeling reveals that dynamic regulation of c-FLIP levels determines cell-to-cell distribution of CD95-mediated apoptosis. *J Biol Chem* 286.21, 18375–18382.

C. J. Tomlin and J. D. Axelrod (2007). Biology by numbers: mathematical modelling in developmental biology. *Nat Rev Genet* 8.5, 331–340.

P. L. Yeyati and V. van Heyningen (2008). Incapacitating the evolutionary capacitor: Hsp90 modulation of disease. *Curr Opin Genet Dev* 18.3, 264–272.