Kinetic Modelling of Mesenchymal Stem Cells

Sebastian Rudden

Abstract

Mesenchymal stem cells (MSCs) are adult stem cells that can be isolated relatively easily and have shown promising results in the treatment of various ailments. Kinetic modelling increases process understanding and has potential to replace sacrificial cell sampling, which is of importance in small cell cultures. Improved process understanding leads to more optimised MSC growth, which will ultimately lead to a reduction in manufacturing costs. To study MSC growth kinetics, immortalized adipose tissue-derived (hTERT-AT-MSCs from now on) were expanded in monolayer for 8 days in batch mode. Cell growth was modelled using 3 kinetic models: (1) First-order (2) Logistic and (3) Weibull. The results showed that all 3 models were able to express viable cell concentration in the exponential phase well ($R^2 > 0.9$), and the Weibull model exhibited strong predictive capacity for the final 2 days, after being trained with 5 days' data (MAPE < 2%). In conclusion, this work is a step toward being able to reliably predict growth for MSCs under these conditions.

Introduction

Literature Review

Cell therapy is currently a major branch of the healthcare industry - from 2016-2019 there have been an average of 20,000 stem cell publications, per year, on PubMed [1]. Mesenchymal stem cells (MSCs) can be isolated from adult bone marrow, fat tissue, and other sources, and are of particular interest since they provide the major advantages of stem cell therapy - large proliferative capacity and differentiation potential – without the ethical issues incurred by embryonic stem cells [2]. Further, MSCs exhibit immunomodulatory effects, allowing them to be used in the delicate treatment of autoimmune diseases, like Graft vs. Host Disease [3]. MSCs have shown promising results in small-scale *in vivo* models for the treatment of various conditions, such as Alzheimer's disease and liver cirrhosis [4][5], and from 2007 to 2017 there were 178 clinical trials using MSCs isolated from cord tissue [6]. However, stem cell treatments have a high cost of manufacture, and therefore a sizeable price tag, compared to other biopharmaceuticals [7]. This is partially because MSCs are anchorage dependent, and there are therefore various challenges involved in their cultivation, including spatial constraints, aggregation of microcarriers and a lack of scalability in the downstream process [8]. These challenges in culturing MSCs suggest a need for more efficient growth methods.

Kinetic modelling has been used increasingly to achieve the goal of more efficient cell culture processes [9], offering several key advantages, including culture simulation, which can reduce the number and cost of experiments, and optimisation of growth parameters. Also, in-depth models can aid in the understanding of underlying cell biology, such as the relationship between antibody concentration and production [10]. A practical use of a simple model, such as those presented here, is to reduce sacrificial sampling, and predict final product concentration with a certain degree of confidence [9]. Models for cell culture can be classified based on homogeneity of population, and homogeneity of the cell itself, as shown in Table 1. Models can also be deterministic, with model output completely dependent on input, or stochastic, which incorporates an element of randomness. The most realistic model would allow for heterogeneity at both the population level (e.g. cells in different phases) and the individual cell level (e.g. intracellular pathways); it would also be stochastic to account for the randomness of life processes [11]. This level of detail can result in overwhelming complexity, and acceptable results can be achieved with more generalised models [12][13].

Table 1. Classification of Cell Culture Models (adapted from Tziampazis and Sambanis, 1994 [11])

	Unsegregated	Segregated	References	
Unstructured	Homogeneous cell population; single-component cells	Heterogeneous cell population; single component cell	[12]	[14]
Structured	Homogeneous cell population; multi-component cells	Heterogeneous cell population; multi-component cells	[15]	[10]

To evaluate the performance of 3 different kinetic models (described below), MSCs were expanded in monolayer for 8 days in batch mode. Then, three unstructured, unsegregated, and deterministic models were fit to the viable cell data using a non-linear least-squares method. The models were analysed and interpreted, and an attempt was made to validate them. The specific growth rate and specific death rate were assumed constant for models (1) and (2), as done in other MSC modelling attempts [16][12]. The models used did not examine differentiation of MSCs, since this only happens upon application of strong chemicals, however this has been attempted [17].

Method

Expansion of MSCs

hTERT-AT-MSCs were cultured according to a standard tissue culture protocol. Briefly, the immortalized MSCs (hTERT-ATMSC, ATCC® SCRC-4000 - ATCC, UK) were thawed from liquid nitrogen tanks, transferred to vials with pre-warmed ATCC proprietary medium, ATCC® PCS-500-040 (ATTC, UK) supplemented with a MSC growth kit ATCC-PCS-500-040 (ATTC, UK) (expansion medium from now on), in a 1:9 ratio, gently mixed and then centrifuged (5 min, 400xg) to remove the cryoprotectant (CS10, Biolife solutions, US), which was added to the cells before long term storage in liquid nitrogen tanks. The cell pellet obtained after the centrifugation cycle was resuspended in expansion medium, counted using the NC-3000 (Chemometec, Denmark), and inoculated in t-flasks (25 cm^2) at 6000 cells/cm^2. The cultures were run in triplicates and in batch mode for 8 days. For every day of the culture, sacrificial cell counting measures and viability assessment were performed; the culture supernatant was stored at -20 C°.

Models

The most basic model for explaining exponential cell growth is a first-order kinetic model, meaning that the model assumes that growth of the cells only depends upon the concentration of one reactant, in this case the number of viable cells [18]. X_v is the viable cell number, μ is the specific growth rate, and k_d is the specific death rate, and X_0 is the viable cell number at the end of the beginning of the exponential phase.

$$\frac{dX_V}{dt} = X_v(\mu - k_d) \rightarrow X_v = X_0 e^{(\mu - k_d)t} \quad (1)$$

The next model builds on the first by adding a term to account for resistance to additional growth, based on the maximum population size; this is known as a logistic model. This term was coined by Verhulst in 1838 [19], and the model has been used to successfully express MSC growth before [12][20]. This model uses the same terms as equation (1), but with the addition of X_{max}, which represents maximum viable cell number. The analytical solution shown on the right ignores the k_d term.

$$\frac{dX_v}{dt} = \mu X_v \left(\frac{X_{max} - X_v}{X_{max}}\right) - k_d X_v \rightarrow X_v = \frac{X_{max} X_0 e^{\mu t}}{X_{max} - X_0 + X_0 e^{\mu t}} \quad (2)$$

Finally, the Weibull distribution is a continuous probability distribution with two parameters to account for shape and scale. Its cumulative distribution function can take a sigmoid shape, making it a plausible model for predicting cell growth. Santhagunam et al. used the following form of the model, taking into account a term for cell confluency [13]. Unlike the previous models, the estimated parameters have no biological significance, and were just used for fitting the data, where β is a scale factor and γ is a shape factor. However, the instantaneous specific growth rate can be calculated from equation (4) [13]. X_{max} was estimated as 1.4 x 10^6, for both equations (2) and (3).

$$X_v = X_{max} - (X_{max} - X_0) e^{(-\beta t^{\gamma})} \quad (3)$$

$$\mu = \frac{1}{X_v} \frac{dX_v}{dt} \quad (4)$$

Fitting Data

Parameters were estimated by fitting the models to the viable cell (X_v) data using the MATLAB (The MathWorks, MA, USA) function *lsqcurvefit,* a non-linear least-squares solver. The models were first fit using all 7 data points (P7), and then on only the first 5 data points (P5). The fitted parameters from the P5 data were used to predict the next two points.

Using the P5 fit, the square of the correlation coefficient (R^2) for the pairs of data points (actual and predicted) was calculated using the Excel (Microsoft, WA, USA) function *correl*; this goodness of fit was measured for the first 5 points, R^2(P5.5), and for all 7 points, R^2(P5.7).

Finally, the Mean Absolute Percentage Error (MAPE) was calculated for the last two pairs, using the P5 fit, as described elsewhere [21].

Results and Discussion

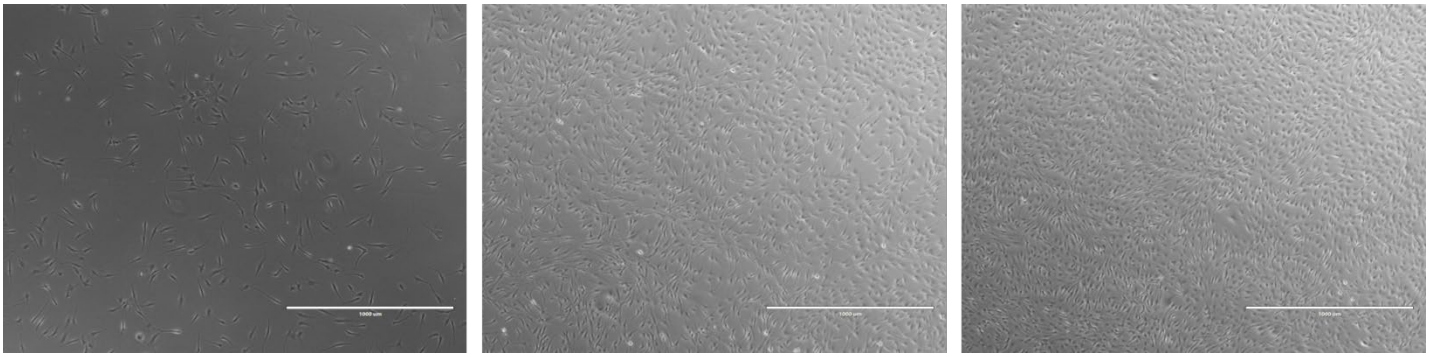

Figure 1. hTERT cells on day 2 (left), day 5 (middle), and day 7 (right) at 4x magnification

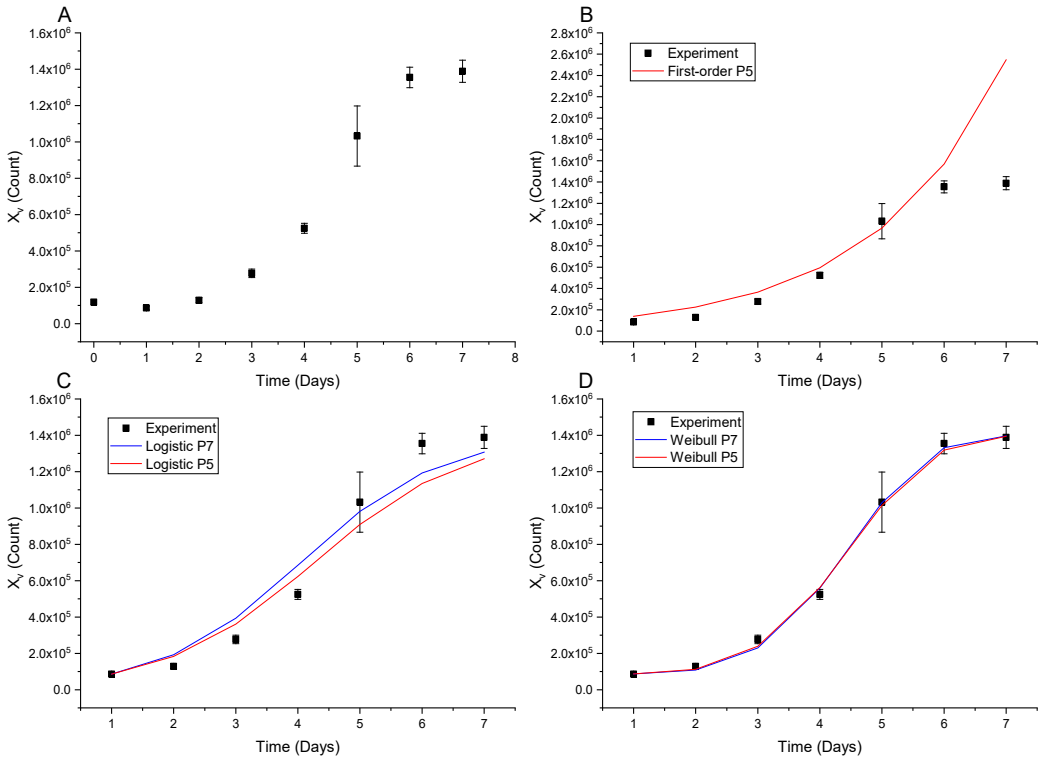

Figure 2. Viable cell count (X_v): Experiment only (A), First-order (B), Logistic (C), and Weibull (D)

Table 2. Parameter estimation, R^2 and MAPE values for models (1), (2), and (3)

Model	Specific Growth Rate (day^{-1})	Specific Death Rate (day^{-1})	R^2 (P5.5)	R^2(P5.7)	MAPE
(1)	0.50	0.02	0.990	0.845	50%
(2)	0.84	0.00	0.959	0.981	12%
(3)	0.75	N/A	0.994	0.998	2%

Figure 1 can be interpreted in conjunction with Figure 2A, which shows the hTERT cells at the start and middle of the exponential phase on days 2 and 5, respectively, and at the plateau phase on day 7. Based on the microscope images in Figure 1, the cells did not exhibit an abnormal phenotype, and appeared to be near confluence, in terms of availability of free space, on day 7.

Figure 2A shows a drop in viable cell number from inoculation (day 0) to day 1. This is presumably because not all cells successfully attached to the microcarriers. As such, the initial viable cell value was set as day 1, instead of the inoculation concentration, to optimise the model fitting. Further, a semi-log plot of Fig 2A (not shown) revealed that the exponential phase started on day 2 and ended on day 5. Day 1 appeared to be in a transition phase between lag and exponential, but was treated as part of the exponential phase nonetheless.

Figure 2B-D shows the results of the models being trained on 5 points (P5) of data, and tested with the last two, as well as being fit to the full 7 days (P7), for comparison. No attempt was made to fit equation (1) to all 7 points of data since, by examination, it is a simple exponential curve with no mechanism to account for a reduction in growth rate.

Constant specific growth rate is generally an acceptable assumption for 'balanced growth of microorganisms' [18] i.e. cells are in the exponential phase without significant growth impedance. As mentioned, cells were in the exponential phase for days 1-5, but various other factors affect cell growth, including nutrient availability (glucose), presence of toxins (ammonia, lactate) and oxygen level [8][22]. The data for these quantities were unavailable, but the assumption is commented on below.

In terms of goodness of fit, the simple exponential model (1) performed extremely well ($R^2 = 0.990$) when compared to the results for days 1 through 5 (P5). This suggests that a constant specific growth rate was a fair assumption for the first 5 days. As expected, this model does not predict in any way the plateau phase on days 6 and 7, which is evidenced by a high MAPE of 50%. Also, it was noted that the R^2(P5.7) was more than 0.8, which can represent a "good fit" [9], despite the prediction on day 7 being off by over 1 million cells. This suggests that R^2 should not be used as a stand-alone measure of a model's predictive capability; the MAPE value gave a more useful result.

Fitting the analytical solution of the logistic model, while it included a k_d term, proved computationally challenging. The model was therefore was solved numerically using the MATLAB function *ode45*, which uses a 5th order Runge-Kutte method. Interestingly, predicted k_d was negligible, but the numerical solution produced a better fit than the analytical one (not shown), so it was utilised nonetheless. Both R^2 values for this model were high, yet it presented the worst explanation of the variance in the first 5 data point and had a MAPE that was above 10%. Upon examination of the analytical solution of the model, μ can be considered a scale factor, which essentially dictates how quickly X_v approaches X_{max}. Therefore, regardless of how μ is adjusted, the shape of the predicted curve remains the same, limiting the model's ability to explain variation in data. However, it must be noted that adherent cells, such as MSCs, are particularly affected by the amount of free space left to grow [11], so this model may produce better results than (1) in other cases. A limitation of the logistic model is that X_{max} must be known in advance, which hinders its predictive capability.

The Weibull model exhibited the best performance of this study. The values of β and γ for the P5 fit were 8.9466×10^{-4} and 4.4881, respectively. Remarkably, this model's predictions for viable cell number on days 6 and 7 were almost the same whether it was trained on 5 days' or 7 days' data. A plausible reason for

this is that the Weibull model, unlike the logistic, can accommodate changes in shape as well as scale. A downfall of this model is that it was likely overfitted to this data, which would hinder the its ability to account for randomness in other sets of data, while using the same parameter values.

Continuing, Table 2 shows three significantly different estimations for the specific growth rate. Note that these values were calculated based on just the exponential phase (P5), for which there was a supposedly constant μ; the value for the Weibull model was obtained at day 4 using equation (4). Given that all the models show a good fit for this data ($R^2 > 0.95$) it can be concluded that while a model may explain viable cell data well, using it to predict specific growth rate should be done with caution. An alternative approach for modelling both the growth rate and death rate is to calculate them based on substrate inhibition kinetics, and feed this into the viable cell equation, rather than predicting parameters based on fitting of cell data [10]. This approach, however, requires additional data, like glucose concentration, and knowledge of pre-determined parameters, which can vary based on cell line and growth environment [18].

Finally, an ideal model should be able to predict cell growth reliably (within a certain confidence interval) if the same cells are grown under the same settings. In this work the models were only fit to one set of data, and therefore said confidence intervals could not be established. As such, even the Weibull model, which showed a superb fit on this set of data, cannot be deemed reliable (in this sense) without further experimentation. Given that a model can only replace sacrificial testing after this verification, the fact that X_{max} must be known for models (2) and (3) is not a major downfall, as it can be estimated from previous growth attempts.

Conclusion

In conclusion, the models evaluated can be used to express MSC growth with excellent R^2 and MAPE values for the conditions examined. Notably, these models should be used with caution to estimate kinetic parameters, such as the specific growth rate. To assess the models' ability to account for the variability in life, they must be fitted to the data from various runs involving the same cell line and growth conditions. A widely accessible database, containing viable cell count data for a range of conditions, durations, and types of MSCs would significantly aid in this process. Ultimately this would allow for the elimination of sacrificial testing, aiding the thrust to improve MSC growth processes, which should reduce the manufacturing cost of MSC treatment.

Acknowledgements

The author would like to thank *immensely* Dr. Rana Khalife and Mr. Pedro da Silva Couto Azevedo for sharing their stem cell expertise, their critical feedback on report iteration, and most importantly for their willingness and patience as they worked with him.

Skills Developed

This project was 'extra' experimental in the sense that there was no clear path from start to finish since the plan was adjusted significantly because of COVID-19 precautions. The major skills I learnt are as follows:

1. Cell culturing: Growing MSCs under standard protocol (in theory)
2. Data analysis in MATLAB: fitting non-linear data using the *lsqcurve* function and numerical integration using the *ode45* function, which uses the powerful Dormand-Prince (5,4) Runge-Kutte method [23]. This is useful when analytical solving is impossible, and for linked systems of differential equations.
3. General research: Graphing in OriginPro (Origin Lab), which is used by papers in professional journals, parameter optimisation in Microsoft Excel (not used in final project), and a bit about managing one's independent research, especially in the face of uncontrollable change (which may not seem as concrete as the other skills, but was possibly the most complex).

References

[1] PubMed n.d. https://pubmed.ncbi.nlm.nih.gov/ (accessed August 6, 2020).

[2] Wang S, Qu X, Zhao RC. Clinical applications of mesenchymal stem cells. J Hematol Oncol 2012;5:19. https://doi.org/10.1186/1756-8722-5-19.

[3] Le Blanc K, Rasmusson I, Sundberg B, Götherström C, Hassan M, Uzunel M, et al. Treatment of severe acute graft-versus-host disease with third party haploidentical mesenchymal stem cells. Lancet 2004;363:1439–41.

[4] Shin JY, Park HJ, Kim HN, Oh SH, Bae JS, Ha HJ, et al. Mesenchymal stem cells enhance autophagy and increase β-amyloid clearance in Alzheimer disease models. Autophagy 2014;10:32–44. https://doi.org/10.4161/auto.26508.

[5] Watanabe Y, Tsuchiya A, Seino S, Kawata Y, Kojima Y, Ikarashi S, et al. Mesenchymal Stem Cells and Induced Bone Marrow-Derived Macrophages Synergistically Improve Liver Fibrosis in Mice. Stem Cells Transl Med 2019;8:271–84. https://doi.org/10.1002/sctm.18-0105.

[6] Couto PS, Shatirishvili G, Bersenev A, Verter F. First decade of clinical trials and published studies with mesenchymal stromal cells from umbilical cord tissue. Regen Med 2019;14:309–19. https://doi.org/10.2217/rme-2018-0171.

[7] Malik NN, Durdy MB. Chapter 7 - Cell Therapy Landscape: Autologous and Allogeneic Approaches. In: Atala A, Allickson JG, editors. Transl. Regen. Med., Boston: Academic Press; 2015, p. 87–106. https://doi.org/https://doi.org/10.1016/B978-0-12-410396-2.00007-4.

[8] Couto PS, Bersenev A, Rafiq QA. Process development and manufacturing approaches for mesenchymal stem cell therapies. In: Fernandes TG, Diogo MM, Cabral JMS, editors. Eng. Strateg. Regen. Med., Elsevier Inc.; 2020, p. 33–71. https://doi.org/10.1016/b978-0-12-816221-7.00002-1.

[9] Schmidberger T, Posch C, Sasse A, Gülch C, Huber R. Progress toward forecasting product quality and quantity of mammalian cell culture processes by performance-based modeling. Biotechnol Prog 2015;31:1119–27. https://doi.org/10.1002/btpr.2105.

[10] Kontoravdi C, Asprey SP, Pistikopoulos EN, Mantalaris A. Application of global sensitivity analysis to determine goals for design of experiments: An example study on antibody-producing cell cultures. Biotechnol Prog 2005;21:1128–35. https://doi.org/10.1021/bp050028k.

[11] Tziampazis E, Sambanis A. Modeling of cell culture processes. Cytotechnology 1994;14:191–204.

[12] Dos Santos F, Andrade PZ, Boura JS, Abecasis MM, Da Silva CL, Cabral JMS. Ex vivo expansion of human mesenchymal stem cells: A more effective cell proliferation kinetics and metabolism under hypoxia. J Cell Physiol 2009;223:27–35. https://doi.org/10.1002/jcp.21987.

[13] Santhagunam A, Santos F dos, Madeira C, Salgueiro JB, Cabral JMS. Isolation and ex vivo expansion of synovial mesenchymal stromal cells for cartilage repair. Cytotherapy 2014;16:440–53. https://doi.org/10.1016/j.jcyt.2013.10.010.

[14] Zheng Z-Y, Yao S-J, Lin D-Q. Using a kinetic model that considers cell segregation to optimize hEGF expression in fed-batch cultures of recombinant Escherichia coli. Bioprocess Biosyst Eng 2005;27:143–52. https://doi.org/10.1007/s00449-004-0376-y.

[15] Çelik E, Çalık P, Oliver SG. A structured kinetic model for recombinant protein production by Mut + strain of Pichia pastoris. Chem Eng Sci 2009;64:5028–35. https://doi.org/10.1016/j.ces.2009.08.009.

[16] Higuera G, Schop D, Janssen F, van Dijkhuizen-Radersma R, van Boxtel T, van Blitterswijk CA. Quantifying in vitro growth and metabolism kinetics of human mesenchymal stem cells using a mathematical model. Tissue Eng Part A 2009;15:2653–63. https://doi.org/10.1089/ten.tea.2008.0328.

[17] Lemon G, Waters SL, Rose FRAJ, King JR. Mathematical modelling of human mesenchymal stem cell proliferation and differentiation inside artificial porous scaffolds. J Theor Biol 2007;249:543–53. https://doi.org/10.1016/j.jtbi.2007.08.015.

[18] Doran PM. Bioprocess Engineering Principles. 2nd ed. Oxford: Academic Press; 2013.

[19] Cramer JS. The Origins of Logistic Regression. Tinbergern Inst Work Pap 2002. https://doi.org/10.2139/ssrn.360300.

[20] Amodu OS, Ojumu T V., Ntwampe SKO. Kinetic modelling of cell growth, substrate utilization, and biosurfactant production from solid agrowaste (Beta vulgaris) by Bacillus licheniformis STK 01. Can J Chem Eng 2016;94:2268–75. https://doi.org/10.1002/cjce.22631.

[21] de Myttenaere A, Golden B, Le Grand B, Rossi F. Mean Absolute Percentage Error for regression models. Neurocomputing 2016;192:38–48. https://doi.org/10.1016/j.neucom.2015.12.114.

[22] Jang JD, Barford JP. An unstructured kinetic model of macromolecular metabolism in batch and fed-batch cultures of hybridoma cells producing monoclonal antibody. Biochem Eng J 2000;4:153–68. https://doi.org/10.1016/S1369-703X(99)00041-8.

[23] Solve nonstiff differential equations — medium order method - MATLAB ode45 n.d. https://www.mathworks.com/help/matlab/ref/ode45.html (accessed August 10, 2020).