Combining Population Modeling and Bayesian Inference for Tumor Growth Prediction
Cristina Vaghi

To cite this version:
Cristina Vaghi. Combining Population Modeling and Bayesian Inference for Tumor Growth Prediction. SMB 2019 annual meeting, Jul 2019, Montreal, Canada. hal-02424592

HAL Id: hal-02424592
https://hal.science/hal-02424592
Submitted on 27 Dec 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
Combining Population Modeling and Bayesian Inference for Tumor Growth Prediction

Cristina Vaghi
Ph.D. Candidate at Inria Sud-Ouest-Bordeaux, MONC team

SMB 2019
Montreal, July 22nd 2019
Motivations

• Few studies have modeled tumor growth kinetics with a population approach and across tumor types.

• The Gompertz model is a widely accepted model of tumor growth. Several studies have reported a strong correlation between the two parameters of the model [1].

• Prediction of the time from cancer initiation would have important clinical implications, such as the determination of invisible metastasis at diagnosis.

Objectives

- Test the descriptive power of different tumor growth models within a population.

- Study the correlation between the parameters of the Gompertz model within a population and define a novel, simplified model: the reduced Gompertz model.

- Use the estimated population parameters to perform individual predictions of tumor initiation using Bayesian inference.
• Material and methods
  ▸ Preclinical data
  ▸ Tumor growth models
  ▸ Nonlinear mixed effects modeling: population approach
  ▸ Bayesian inference
• Results
  ▸ Population analysis
  ▸ The reduced Gompertz model
  ▸ Backward prediction of the initiation time of the tumor
• Conclusions
Data

- Preclinical data
- Longitudinal measurements of the primary tumor
- Three data sets:
  - Breast cancer (volume data): 66 individuals (J. Ebos et al., Roswell Cancer Park)
  - Breast cancer (fluorescence data): 8 individuals (A. Rodallec et al., SMARTc team)
  - Lung cancer (volume data): 20 individuals (S. Benzekry, C. Lamont, Center of Cancer and System Biology)
Material and methods

Tumor growth models

Exponential

\[
\begin{align*}
\frac{dV}{dt} &= \alpha^i V \\
V(t_0) &= V_I \\
V(t_f) &= V_I
\end{align*}
\]

Logistic

\[
\begin{align*}
\frac{dV}{dt} &= \alpha^i V \left(1 - \frac{V}{K^i}\right) \\
V(t_0) &= V_I \\
V(t_f) &= V_I
\end{align*}
\]

Gompertz

\[
\begin{align*}
\frac{dV}{dt} &= \left(\alpha^i - \beta^i \log \left(\frac{V}{V_{inj}}\right)\right) V \\
V(t_0) &= V_I \\
V(t_f) &= V_I
\end{align*}
\]

Material and methods

Nonlinear mixed effects modeling

- Statistical framework to study longitudinal data
- The goal is to understand the intra-subject process of tumor growth and its variability across the individuals

\[ y_j^i = f(t_j^i; \theta^i) + e_j^i \]

- \( y_j^i \) is the observation of the \( j \)-th measurement of the \( i \)-th individual at time \( t_j^i \)
- \( f(t_j^i; \theta^i) \) is the structural model
- \( e_j^i \) is the residual error
- \( \theta^i = \mu \exp(\eta^i) \)
- \( \eta^i \sim \mathcal{N}(0, \Omega) \)

Model for the individual parameters

- Error model

\[ e_j^i = (\sigma_1 + \sigma_2 f(t_j^i, \theta^i)) \varepsilon_j^i \]
- \( \varepsilon_j^i \sim \mathcal{N}(0, 1) \)


Bayesian inference

\[ p(\theta^j | y^j; \phi) = \frac{p(y^j | \theta^j; \phi)}{p(\theta^j; \phi)} \]

**posterior distribution**  **prior distribution**  **likelihood**

Algorithm:

- Draw a realization from the posterior distribution
- Compute \( a_{\text{pred}}^{(l)i} = f^{-1}\left( V_{\text{inj}}; \theta^{(l)i}\right), l = 1, \ldots, L \)

\[ \text{err}^i = \frac{a_{\text{pred}}^i - a^i}{a^i} \]

\( k \)-fold cross validation

Results

Population analysis

Figure 1. Population analysis of experimental tumor growth kinetics. A) Visual predictive checks assess goodness-of-fit for both structural dynamics and inter-animal variability by reporting model-predicted percentiles (together with confidence prediction intervals (P.I) in comparison to empirical ones. B) Prediction distributions. C) Individual weighted residuals (IWRES) with respect to time. D) Observations vs predictions Left: Exponential, Center: Logistic, Right: Gompertz models.
Results

Population analysis

Figure 1. Population analysis of experimental tumor growth kinetics. A) Visual predictive checks assess goodness-of-fit for both structural dynamics and inter-animal variability by reporting model-predicted percentiles (together with confidence prediction intervals (P.I) in comparison to empirical ones. B) Prediction distributions. C) Individual weighted residuals (IWRES) with respect to time. D) Observations vs predictions Left: Exponential, Center: Logistic, Right: Gompertz models.
The reduced Gompertz model

\[ R^2 = 0.97 \]
\[ p\text{-value} < 10^{-5} \]

\[ \alpha^i = k \beta^i \]

\[ \frac{dV}{dt} = \left(k \beta^i - \beta^i \log \left( \frac{V}{V_{\text{inj}}} \right) \right) V \]

\[ V(t_f) = V_I \]
Results

The reduced Gompertz model: biological interpretation

\[ K^i = V_{\text{inj}} e^{\frac{\alpha^i}{\beta^i}} \approx V_{\text{inj}} e^k \quad \forall i. \]

**Constant maximal tumor size** within a tumor type in a given species
## Results

### Population analysis

<table>
<thead>
<tr>
<th>Model</th>
<th>-2LL</th>
<th>AIC</th>
<th>BIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gompertz</td>
<td>7128</td>
<td>7142</td>
<td>7157</td>
</tr>
<tr>
<td>Reduced Gompertz</td>
<td>7259</td>
<td>7269</td>
<td>7280</td>
</tr>
<tr>
<td>Logistic</td>
<td>7584</td>
<td>7596</td>
<td>7609</td>
</tr>
<tr>
<td>Exponential</td>
<td>8652</td>
<td>8660</td>
<td>8669</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Model</th>
<th>Parameter</th>
<th>Unit</th>
<th>Fixed effects</th>
<th>CV (%)</th>
<th>R.S.E. (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gompertz</td>
<td>(\alpha)</td>
<td>day(^{-1})</td>
<td>0.573</td>
<td>34.73</td>
<td>2.56</td>
</tr>
<tr>
<td></td>
<td>(\beta)</td>
<td>day(^{-1})</td>
<td>0.0705</td>
<td>391.49</td>
<td>3.61 [18.3, 7.36]</td>
</tr>
<tr>
<td></td>
<td>(\sigma)</td>
<td>-</td>
<td>[19.1, 0.12]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reduced Gompertz</td>
<td>(\beta)</td>
<td>day(^{-1})</td>
<td>0.0725</td>
<td>180.69</td>
<td>1.91</td>
</tr>
<tr>
<td></td>
<td>(k)</td>
<td>-</td>
<td>7.98</td>
<td>0</td>
<td>0.363 [22.3, 5.17]</td>
</tr>
<tr>
<td></td>
<td>(\sigma)</td>
<td>-</td>
<td>[13.9, 0.183]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Logistic</td>
<td>(\alpha)</td>
<td>day(^{-1})</td>
<td>0.324</td>
<td>42.90</td>
<td>1.88</td>
</tr>
<tr>
<td></td>
<td>(K)</td>
<td>mm(^3)</td>
<td>1332</td>
<td>0.02</td>
<td>4.39 [9.8, 8.74]</td>
</tr>
<tr>
<td></td>
<td>(\sigma)</td>
<td>-</td>
<td>[57.2, 0.136]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Exponential</td>
<td>(\alpha)</td>
<td>day(^{-1})</td>
<td>0.229</td>
<td>34.98</td>
<td>1.35 [6.06, 14.3]</td>
</tr>
<tr>
<td></td>
<td>(\sigma)</td>
<td>-</td>
<td>[283, 0.254]</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Results

Population analysis: example of individual fits

Exponential

Logistic

Gompertz

Reduced Gompertz
Results

Backward predictions: posterior distribution

![Graph showing backward predictions with posterior distribution]

- **Gompertz**
- **Reduced Gompertz**

![Prior distribution and Posterior distribution plots]
Results

Backward predictions: posterior distribution

Figure 4. Backward predictions computed with likelihood maximization and with Bayesian inference.

Three examples of backward predictions of individuals A, B and C computed with likelihood maximization (LM) and Bayesian inference: Gompertz model with likelihood maximization (first row); reduced Gompertz with likelihood maximization (second row); Gompertz with Bayesian inference (third row) and reduced Gompertz with Bayesian inference (fourth row). Only the last three points are considered to estimate the parameters. The grey area is the 90% prediction interval (P.I) and the dotted blue line is the median of the posterior predictive distribution. The red line is the predicted initiation time and the black vertical line the actual initiation time.
Results

Backward predictions

Figure 4. Backward predictions computed with likelihood maximization and with Bayesian inference. Three examples of backward predictions of individuals A, B and C computed with likelihood maximization (LM) and Bayesian inference: Gompertz model with likelihood maximization (first row); reduced Gompertz with likelihood maximization (second row); Gompertz with Bayesian inference (third row) and reduced Gompertz with Bayesian inference (fourth row). Only the last three points are considered to estimate the parameters. The grey area is the 90% prediction interval (P.I) and the dotted blue line is the median of the posterior predictive distribution. The red line is the predicted initiation time and the black vertical line the actual initiation time.
Backward predictions

- Reduced Gompertz >> Gompertz
- Bayesian inference >> likelihood maximization (LM)

<table>
<thead>
<tr>
<th>Model</th>
<th>Method</th>
<th>Accuracy (%)</th>
<th>Precision (days)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reduced Gompertz</td>
<td>Bayesian</td>
<td>12.1 (1.02)</td>
<td>15.2 (0.503)</td>
</tr>
<tr>
<td>Reduced Gompertz</td>
<td>LM</td>
<td>74.1 (11.6)</td>
<td>186 (52.8)</td>
</tr>
<tr>
<td>Gompertz</td>
<td>Bayesian</td>
<td>19.6 (1.77)</td>
<td>40.1 (1.94)</td>
</tr>
<tr>
<td>Gompertz</td>
<td>LM</td>
<td>205 (55.4)</td>
<td>-</td>
</tr>
</tbody>
</table>

Accuracy was defined as the absolute value of the relative error (in percent). Precision was defined as the width of the 95% prediction interval (in days). Reported are the means and standard errors (in parenthesis). LM = likelihood maximization.
Conclusions

- The Gompertz model described well tumor growth kinetics while the Exponential and the Logistic showed inferior predictive power.

- We proposed a novel reduced Gompertz model with only one individual parameter.

- The combination of nonlinear mixed effects modelling and Bayesian inference allowed to have reliable predictions of individual tumor age.

- The method proposed herein remains to be extended to clinical data: it would yield important epidemiological insights and could also be informative in routine clinical practice for prediction of metastatic extent.
Thank you for your attention!

Preprint available!

A reduced Gompertz model for predicting tumor age using a population approach


doi: https://doi.org/10.1101/670869