A composite likelihood approach to the analysis of longitudinal clonal data on multitype cellular systems under an age-dependent branching process

RUI CHEN, OLLIVIER HYRIEN*

Department of Biostatistics and Computational Biology,
University of Rochester Medical Center, 601 Elmwood Avenue, Rochester, NY 14642, USA
ollivier_hyrien@urmc.rochester.edu

MARK NOBLE, MARGOT MAYER-PRÖSCHEL

Department of Biomedical Genetics, University of Rochester Medical Center,
601 Elmwood Avenue, Rochester, NY 14642, USA

SUMMARY
A recurrent statistical problem in cell biology is to draw inference about cell kinetics from observations collected at discrete time points. We investigate this problem when multiple cell clones are observed longitudinally over time. The theory of age-dependent branching processes provides an appealing framework for the quantitative analysis of such data. Likelihood inference being difficult in this context, we propose an alternative composite likelihood approach, where the estimation function is defined from the marginal or conditional distributions of the number of cells of each observable cell type. These distributions have generally no closed-form expressions but they can be approximated using simulations. We construct a bias-corrected version of the estimating function, which also offers computational advantages. Two algorithms are discussed to compute parameter estimates. Large sample properties of the estimator are presented. The performance of the proposed method in finite samples is investigated in simulation studies. An application to the analysis of the generation of oligodendrocytes from oligodendrocyte type-2 astrocyte progenitor cells cultured in vitro reveals the effect of neurotrophin-3 on these cells. Our work demonstrates also that the proposed approach outperforms the existing ones.

Keywords: Bias correction; Cell differentiation; Composite likelihood; Discrete data; Monte Carlo; Neurotrophin-3; Oligodendrocytes; Precursor cell; Stochastic model.

1. INTRODUCTION
In stem cell biology, there exists considerable interest in studying signals that may modulate or alter the processes that regulate the formation of tissues during development or repair. Understanding such processes has important clinical implications as it may offer potential means to restore or maintain tissue

*To whom correspondence should be addressed.

© The Author 2010. Published by Oxford University Press. All rights reserved. For permissions, please e-mail: journals.permissions@oup.com
and organ functions. One of the challenges in conducting detailed analysis of the effects of these signals, however, is the general lack of quantitative approaches that allow analysis of different aspects of progenitor or stem cell function.

The present paper is focused on oligodendrocyte type-2 astrocyte (O-2A) progenitor cells, which is one of the best studied precursor cell populations (e.g. Noble and others, 2004). These progenitor cells may either divide into 2 new progenitor cells or differentiate into terminally differentiated oligodendrocytes, which produce the myelin sheaths that enwrap axons in the central nervous system. The differentiation of progenitor cells into oligodendrocytes needs to be properly regulated to maintain normal function of the central nervous system. In addition to the function of O-2A progenitor cells in vivo, their ability to grow and differentiate in vitro has established these cells as an important tool in modern cell biology that allows the study of the most basic cellular functions (i.e. self-renewal by division, differentiation, and death) at the clonal level in tissue culture (e.g. Ibarrola and others (1996); Smith and others, 2000). The development of a typical clone of these cells cultured in vitro started from a single O-2A progenitor cell is displayed in Figure 1. It may be altered in multiple ways following exposure to agents, such as cytokines, drugs, or toxicants, and it is of great interest to biologists to determine and quantify these changes (Dietrich and others, 2006).

Over the past decade, a great deal of attention has been paid to the development of a statistical framework for gaining a quantitative understanding of the biological processes that govern the division of O-2A progenitor cells and their differentiation into oligodendrocytes (Yakovlev, Boucher, and others, 1998; Yakovlev, Mayer-Pröschel, and others, 1998; Yakovlev and others, 2000; von Collani and others, 1999; Boucher and others, 1999, 2001; Zorin and others, 2000; Hyrien and others, 2005a,b, 2006; Hyrien, 2007). These publications used multitype age-dependent branching processes as a means to model the temporal development of cell clones. The resulting methods provided quantitative insights into the generation of oligodendrocytes from cultured O-2A progenitor cells. For instance, Yakovlev, Boucher, and others (1998) observed that the generation of oligodendrocytes was regulated by a combination of cell-intrinsic factors and environmental signals that may modulate the probability of differentiation of O-2A progenitor cells. Other studies based on similar quantitative approaches (Yakovlev, Boucher, and others, 1998; Yakovlev, Mayer-Pröschel, and others, 1998; Yakovlev and others, 2000; von Collani and others,

![Fig. 1. Oligodendrocyte development under in vitro conditions.](https://academic.oup.com/biostatistics/article-abstract/12/1/173/392355)

When cultured under appropriate growing conditions, any O-2A progenitor cell will either divide into 2 new O-2A progenitor cells or it will differentiate into a single terminally differentiated oligodendrocyte. O-2A progenitor cells may also die. Oligodendrocytes are terminally differentiated cells that arise from O-2A progenitor cells. These cells are not able to divide and they ultimately die (i.e. disintegrate and disappear from the population). Cell death was not observed in our NT-3 experiment.
confirmed that thyroid hormone induces the differentiation of O-2A progenitor cells into oligodendrocytes (Barres and others, 1994; Ibarrola and others, 1996; Ahlgren and others, 1997; Rodríguez-Peña, 1999), and that the decision of O-2A progenitor cells to either divide or differentiate is made early in the cell cycle, if not before, rather than at the end (Hyrien and others, 2005a, 2006, 2010).

Most of these previous studies were concerned either with clonal experiments in which every cellular clone was observed only once in the course of the experiment thereby yielding independent observations (Yakovlev, Boucher, and others, 1998; Yakovlev, Mayer-Pröschel, and others, 1998; Yakovlev and others, 2000; von Collani and others, 1999; Boucher and others, 1999, 2001; Zorin and others, 2000; Hyrien and others, 2005a, b, 2006, 2010; Hyrien, 2007) or with time-lapse experiments (Hyrien and others, 2006), where the complete history of each clone was fully recorded. While time-lapse experiments provide complete information about the temporal development of each clone, clonal experiments are far less time-consuming. This is the reason why clonal analyses remain the experiment of choice to investigate the ability of agents to impact the regulation of precursor cell functions. There exists a third experimental setting in which each cell clone is observed longitudinally over time and which offers a compromise between the above 2 experiments. In this experimental setting, the family trees are not completely observed, but several snapshots of the temporal development of each individual clone are obtained at multiple time points thereby providing more information than independent clonal experiments. We shall refer to these alternate experiments as longitudinal clonal experiments (as opposed to “independent” clonal experiments where each cell clone is examined only once).

Longitudinal clonal data are not independent across time points, and their analysis presents new challenges. Some of the estimation methods that have been proposed for independent clonal data could also apply to the longitudinal setting. Since maximum likelihood estimation remains generally unattainable with age-dependent branching processes, simulation-based approaches have been adopted by some authors as a means to construct viable estimators. For instance, Hyrien and others (2005a, b) and Hyrien (2007) considered the utility of a simulated pseudolikelihood approach. This method relies solely on the expectation and variance–covariance functions of the process. It is attractive because it provides consistent and asymptotically Gaussian estimators, while remaining easy to implement, even for longitudinal clonal data. The disadvantage of this estimator is that it does not enjoy optimality properties, and, for instance, it is not as efficient as the method of maximum likelihood (Hyrien, 2007). The simulated maximum likelihood estimator proposed by Zorin and others, 2000 could be more efficient than the simulated pseudo maximum likelihood estimator but its implementation faces serious limitations that are amplified when cell clones are observed repeatedly over time. Specifically, it is very time-consuming, and it suffers from mismatches, as described by Zorin and others (2000). We shall discuss these limitations in greater details in Section 4 when motivating the construction of the proposed estimator.

This article proposes an alternative approach that provides a compromise between the statistical efficiency of the maximum likelihood estimator and the computational advantage of the simulated maximum pseudolikelihood estimator. The proposed method is based on a composite likelihood approach (sometimes also referred to as method of pseudolikelihood). The use of composite likelihood traces back to Besag (1974). It has been investigated in other settings by several authors, including Azzalini (1983), Lindsay (1988), Heagerty and Lele (1998), Cox and Reid (1987), Chandler and Bate (2007), Varin and Vidoni (2005), Hyrien and Zand (2008), to name a few. Composite likelihoods can be defined as estimating functions formed by combining together likelihood objects that remain tractable for the problem at hands. Our work shows that composite likelihood estimators provide a viable solution to the analysis of longitudinal clonal data.

The proposed method relies on a complex stochastic process. Simpler statistical approaches could be invoked to describe the time course of cell counts and assess whether specific treatment conditions alter the kinetics of the cell population. The level of sophistication of the proposed method is required for
gaining a quantitative insight into the processes of division and differentiation of precursor cells because the events of interest (division and differentiation) are not observable during clonal experiments.

The remainder of this article is organized as follows. Section 2 introduces the data structure. Section 3 presents a branching process model of the generation of oligodendrocytes from cultured O-2A progenitor cells. Section 4 describes simulated composite likelihood estimation as applied to longitudinal clonal data. The finite-sample properties of the proposed method are investigated via simulations in Section 5. Section 6 presents an application to the analysis of the proliferation of O-2A progenitor cells and their differentiation into oligodendrocytes exposed to neurotrophin-3 (NT-3) in tissue culture.

2. LONGITUDINAL CLONAL DATA

2.1 Clonal data on the generation of oligodendrocytes in vitro

A typical longitudinal clonal experiment conducted on oligodendrocytes begins by plating initiator O-2A progenitor cells in culture dishes at a density that is low enough so these cells (and the pools of subsequent progenies) will not interact with each other. Over time, these progenitor cells divide into O-2A progenitor cells and/or differentiate into oligodendrocytes to give rise to individual cell clones. These clones may contain O-2A progenitor cells or oligodendrocytes, or, as is more typically the case, a mixture of both cell types. O-2A progenitor cells and oligodendrocytes are morphologically distinguishable, which enables experimentalists to count separately the number of O-2A progenitor cells and the number of oligodendrocytes contained in any given clone through visual examination using a microscope. The location of each clone in the dish is identified for subsequent follow up, and the composition of each clone is examined repeatedly over time so one can observe how the numbers of O-2A progenitor cells and the number of oligodendrocytes change over time.

2.2 An example of longitudinal clonal data

We performed a longitudinal clonal experiment, as described above, to investigate the impact of NT-3 on the processes of division and differentiation of O-2A progenitor cells. We purified O-2A progenitor cells isolated from the corpus callosum tissue of 7-day old rats as described previously (Ibarrola and others, 1996) and plated cells at a clonal density in defined medium supplemented with NT-3 at 20 ng/ml. No platelet-derived growth factor (PDGF) was added to the culture medium at any time. In this experiment, \( n = 40 \) clones were followed for up to 6 days. Every clone was generated by a single O-2A progenitor cell. The composition of each clone was examined daily, so \( t_i = (1, 2, 3, 4, 5, 6) \) in this particular experiment (with time being expressed in days). The number of O-2A progenitor cells and the number of oligodendrocytes were reported at each time point for every clone. Half of these clones were cultured in the presence of NT-3, and the other half was cultured without NT-3.

Figure 2 shows the histograms for the number of progenitor cells (left panels) and for the number of oligodendrocytes (right panels) from day 0 to day 6 (from top to bottom). At the start of the experiment (day 0), all clones contained exactly one O-2A progenitor cell and zero oligodendrocyte. The composition of each clone evolved over time according to whether O-2A progenitor cells and their subsequent progenies divided or differentiated. In this particular experiment, all clones growing without NT-3 contained between 0 and 6 O-2A progenitor cells and between 0 and 4 oligodendrocytes. In the presence of NT-3, the number of O-2A progenitor cells ranged between 0 and 9, and the number of oligodendrocytes between 0 and 6.

The primary objectives of our experiment were to assess and investigate the effects of NT-3 on the proliferation of O2-A progenitor cells and their differentiation into oligodendrocytes. A visual comparison of the histograms of the number of O2-A progenitor cells and of the number of oligodendrocytes clearly
suggests that the numbers of O-2A progenitor cells were substantially smaller in clones treated with NT-3 starting from day 2. In contrast, the numbers of oligodendrocytes appeared similar among treated and untreated clones at all time points. The marginal distributions of the number of O-2A progenitor cells (respectively, oligodendrocytes) at any time point in clones treated with and in clones treated without NT-3 could be compared using (e.g.) Wilcoxon rank sum statistics. In order to assess the overall effect of NT-3, irrespective of time, and avoid issues associated with multiple testing, the resulting \( p \)-values could be combined using Fisher’s combination method. Since the individual \( p \)-values are dependent across time points, an overall \( p \)-value could be computed using a permutation testing approach, where clones (not just individual observations) are randomly assigned to one group or the other so the dependencies among observations are properly accounted for when performing the test.

We assessed the overall effect of NT-3 on the numbers of O-2A progenitor cells and on the numbers of oligodendrocytes separately using this approach. Our test suggested that the numbers of O-2A progenitor cells were significantly different depending upon whether they had been cultured with or without NT-3 \( (p = 0.02) \), but it did not detect any significant difference among the numbers of oligodendrocytes \( (p = 0.98) \). Thus, this analysis would suggest that NT-3 had a significant impact on the number of O-2A progenitor cells but not on the number of oligodendrocytes, at least for the first 6 days of culture.

The conclusion of these analyses is of interest in itself because it reveals the existence of a potential effect of NT-3 on the regulation of the processes of division and of differentiation of O-2A progenitor cells. From a biological standpoint, however, it remains limited in scope because it neither offers any mechanistic interpretation on how these processes might have actually been altered by NT-3 nor quantifies the effects of NT-3 on cellular functions. A number of biological reasons could be invoked to explain the observed effects of NT-3. For instance, O-2A progenitor cells exposed to NT-3 might have differentiated more frequently in the presence of NT-3, causing the number of O-2A progenitor cells to decrease...
prematurely. An alternative explanation is that O-2A progenitor cells divided more slowly and/or differen-
tiated faster when exposed to NT-3. A combination of these 2 scenarios is also quite plausible. In addition,
experimentalists would also be interested in evaluating the impact of such changes on the number of oligo-
dendrocytes produced by O-2A progenitor cells over an extended period of time. In order to gain such a
quantitative insight into cell proliferation kinetics, we propose to model the dynamics of cell clones using
an age-dependent branching process.

3. A BRANCHING PROCESS MODEL OF OLIGODENDROCYTE GENERATION

We model the temporal development of the population of O-2A progenitor cells and of terminally differen-
tiated oligodendrocytes using a multitype age-dependent branching process defined from the following set
of assumptions:

(A1) The process begins at \( t = 0 \) with a single cell of type 1 (an initiator O-2A progenitor cell).

(A2) For every \( k = 1, 2, \ldots \), upon completion of its lifespan, every type-\( k \) cell (an O-2A progenitor
cell of generation \( k \)) either divides into 2 new cells of age 0 and type \( k+1 \) with probability \( \pi_k \) or
it differentiates into a single type-0 cell (an oligodendrocyte) with probability \( 1 - \pi_k \). Following
Boucher and others (2001) and Hyrien and others (2005a,b), we assumed that the probability of
division is given by \( \pi_k = \min\{1, a + be^k\}, k \geq 1 \), where \( a \), \( b \), and \( c \) are unknown positive constants
with, for instance, \( a \) denoting the limiting probability of division of O-2A progenitor cells as the
number of division increases when \( c < 1 \).

(A3) Time to division of initiator O-2A progenitor cells: The time to division of any type-1 cell is
represented by a nonnegative random variable (r.v.) \( T_1 \) assumed to follow a gamma distribution
with cumulative distribution function (c.d.f.) \( G_1(t; \xi_1) \), where \( \xi_1 = (m_1, \sigma_1) \), and where \( m_1 \) and
\( \sigma_1 \) denote the expectation and standard deviation of \( T_1 \).

(A4) Time to division of O-2A progenitor cells of generation greater than or equal to 2: The time to
division of any O-2A progenitor cells of generation \( k, k \geq 2 \), is described by a nonnegative r.v. \( T_2 \)
assumed to follow a gamma distribution with c.d.f. \( G_2(t; \xi_2) \), where \( \xi_2 = (m_2, \sigma_2) \), and where \( m_2 \)
and \( \sigma_2 \) denote the expectation and standard deviation of \( T_2 \).

(A5) Time to differentiation of O-2A progenitor cells: The time to division of any O-2A progenitor
cells of generation \( k, k \geq 1 \), is described by a nonnegative r.v. \( T_0 \) assumed to follow a gamma
distribution with c.d.f. \( G_0(t; \xi_0) \), where \( \xi_0 = (m_0, \sigma_0) \), and where \( m_0 \) and \( \sigma_0 \) denote the expectation
and standard deviation of \( T_0 \). The distribution of the time to differentiation is thus independent of
the generation number.

(A6) Independence assumption: Cells evolve independently from each other.

The above process extends a model of the generation of oligodendrocytes previously proposed by
Hyrien and others (2005a) by allowing the distribution of the time to division of O-2A progenitor cells of
first generation to be dissimilar from that of O-2A progenitor cells of subsequent generations. This
extension was motivated by an analysis of time-lapse data on oligodendrocytes (Hyrien and others, 2006),
which indicated that first generation O-2A progenitor cells may have a shorter time to division than cells of
subsequent generations.

Let \( \theta = (m_0, \sigma_0, m_1, \sigma_1, m_2, \sigma_2, a, b, c)' \) denote the complete set of unknown model parameters. For
every \( k = 0, 1, \ldots \), let \( Z_k(t) \) denote the number of type-\( k \) cells at time \( t \geq 0 \), with \( Z_0(t) \) corresponding to
the number of oligodendrocytes, and, for every \( k \geq 1 \), \( Z_k(t) \) representing the number of O-2A progenitor
cells of generation \( k \). It is not possible to determine the generation of O-2A progenitor cells during clonal
experiments, and our observations consist of the number of O-2A progenitor cells \( Y_1(t) = \sum_{k=1}^{\infty} Z_k(t) \)
and of the number of oligodendrocytes \( Y_2(t) = Z_0(t) \). The observable process is therefore the vector

\[ (Y_1(t), Y_2(t)) \]

Downloaded from https://academic.oup.com/biostatistics/article-abstract/12/1/173/392355
by guest
on 07 May 2018
\( Y(t) = \{Y_1(t), Y_2(t)\} \). The distribution of the process \( Y(t) \) does not generally admit a closed-form expression, which is why we and others have resorted to simulations to approximate this distribution.

4. Statistical inference

4.1 A composite likelihood estimator

Let \( Y_{ij1} \) and \( Y_{ij2} \) denote the number of O-2A progenitor cells and the number of oligodendrocytes counted in the \( i \)th clone at time \( t_{ij} \), where \( j = 1, \ldots, n_i \) and \( i = 1, \ldots, n \). Write \( Y_{ij} = (Y_{ij1}, Y_{ij2}) \) and \( Y_i = (Y_{i1}, \ldots, Y_{in_i}) \) for the ordered time points \( t_i = (t_{i1}, \ldots, t_{in_i}) \). The log-likelihood function for the sample \( Y_1, \ldots, Y_n \) takes the form

\[
L(\theta) = \sum_{i=1}^{n} \sum_{j=1}^{n_i} \log p_{ij}(\theta),
\]

where \( p_{ij}(\theta) = \mathbb{P}_{\theta}(Y(t_{ij}) = (Y_{ij1}, Y_{ij2}) = (Y_{i,j-1}, \ldots, Y_{i1}) = Y_{i1}) \) denote the conditional distribution function of the r.v. \( Y(t_{ij}) \), given the trajectory of the process \( Y(t) \) is observed at all sampling times prior to \( t_{ij} \).

In the present setting, the \( p_{ij}(\theta) \)'s have usually no closed-form expressions, making (exact) maximum likelihood inference impossible. Zorin and others (2000) proposed a simulated maximum likelihood approach when a single observation is available per clone \((n_i = 1)\). The probabilities \( p_{i1}(\theta) \)'s are approximated by simulating independent clones from the branching process, and by computing the empirical proportions of simulated clones whose composition at time \( t_{i1} \) matches that of the \( i \)th experimental clone. Denoting the corresponding approximation by \( \hat{p}_{i1}(\theta) \), \( L(\theta) \) is replaced by \( \hat{L}(\theta) = \sum_{i=1}^{n} \log \hat{p}_{i1}(\theta) \), and a simulated maximum likelihood estimate is computed using a multidimensional version of the Kiefer–Wolfowitz algorithm (Kiefer and Wolfowitz, 1952; Blum, 1954).

There are 2 limitations to such an approach. The first one is that \( \hat{p}_{i1}(\theta) \) will be identically zero if the composition of the \( i \)th clone as observed during the actual experiment cannot be matched by any of the simulated clones. Consequently, \( \hat{L}(\theta) \) is undefined, making it difficult to compute a simulated maximum likelihood estimate. This problem was reported by Zorin and others (2000) in their analysis of oligodendrocytes generation. The second limitation has to do with the fact that the simulated log-likelihood function \( \hat{L}(\theta) \) is a biased estimator of \( L(\theta) \). The Kiefer–Wolfowitz algorithm considered by these authors is designed to converge to the parameter vector that maximizes the map \( \theta \mapsto g(\theta) = \mathbb{E}[\hat{L}(\theta)|Y_1, \ldots, Y_n] \), where the expectation is taken with respect to the simulations. Since the maximum of \( g(\theta) \) will be generally different than the maximum likelihood estimate, the simulated maximum likelihood estimator will be biased. The severity of the bias depends, of course, upon the noise of the simulated log-likelihood that is attributable to the simulations and upon the gradient of \( L(\theta) \) with respect to \( \theta \). One can argue that the bias will become small if the number of simulations \( S \) is large enough. Computing time, however, may become an issue with branching processes, especially when the process is supercritical (i.e. loosely speaking, when each cell produces on average more than one cell), which causes cell clones to grow large; see Hyrien and others (2010) for additional comments on this practical issue. Thus, the strategy that consists of letting \( S \) increase is not always applicable in this case. Furthermore, there exists a simple way to approximate \( L(\theta) \) that is able to mitigate both limitations simultaneously (see Section 4.2).

The simulated maximum likelihood estimator proposed by Zorin and others (2000) can be extended in an obvious way to the situation where clones are observed longitudinally over time \((n_i > 1)\). However, the method becomes less appealing than when \( n_i = 1 \) because now simulated clones need to match the trajectories of experimental clones observed at several time points, not just one. Since the probability of mismatching increases quickly with \( n_i \), the number of simulations would have to increase with \( n_i \) thereby making the approach even more time-consuming.
As an alternative to the method of maximum likelihood and simulated maximum likelihood, we propose to construct estimators of \( \theta \) using the estimating function

\[
P(\theta) = \sum_{i=1}^{n} \sum_{j=1}^{n_i} \log q_{ij}(\theta),
\]

where \( q_{ij}(\theta) = \mathbb{P}_\theta(Y(t_{ij}) = Y_{ij} | Y(t_{i,j-k}) = Y_{ij,k}, k \in I_{ij}) \), and where \( I_{ij} \subset \{1, \ldots, j-1\} \) for every \( i \) and \( j \). The probability \( q_{ij}(\theta) \) differs from \( p_{ij}(\theta) \) in that the conditioning event is restricted to a subset of the past observations. By selecting \( I_{ij} \) appropriately, we can thus reduce the curse of dimensionality encountered with the simulated log-likelihood. These probabilities \( q_{ij}(\theta) \) do not generally have explicit expressions either, but we shall leave aside this aspect of the method until Section 4.2. The estimating function \( P(\theta) \) is a composite likelihood, and the maximum composite likelihood estimator maximizes \( P(\theta) \). We shall denote it by \( \tilde{\theta} \).

The conditioning sets \( I_{ij} \) need not be identical for all \( i \) and \( j \) nor do they have to be collections of consecutive integers, although they will likely be so in most applications. Thus, \( I_{ij} = \{1, 3, 5\} \) would be allowed for instance. They are assumed to be independent of \( \theta \), however. Depending upon the choice of \( I_{ij} \), we obtain different estimating functions, the complexity of which (from a computational standpoint) increases as \( I_{ij} \) “approaches” \( \{1, \ldots, j-1\} \). The following cases are the most noteworthy:

- If \( I_{ij} = \{1, \ldots, j-1\} \), \( P(\theta) = L(\theta) \), and \( \tilde{\theta} \) is a maximum likelihood estimator.
- If \( I_{ij} = \emptyset \), no conditioning (other than upon the state of the process at the start of the experiment) is done. This is the simplest case. Hyrien and Zand (2008) used it in the context of CFSE-labeling data.
- If \( I_{ij} = \{1\} \), all probabilities are conditioned upon the most recent observation. The method leads to a maximum likelihood estimator when the process is Markovian and no efficiency is lost in this particular case. For non-Markovian models, this property does not hold true since the conditional distribution of \( Y(t_{ij}) \) given \( \{Y(t_{i,j-k}), k = 1, \ldots, j-1\} \), will still depend on \( Y(t_{i,1}), \ldots, Y(t_{i,j-2}) \).
- By taking subsets of the form \( I_{ij} = \{1, \ldots, k_{ij}\} \), for some \( k_{ij} \in \{1, \ldots, j-1\} \), we can modulate how far in the past the probability \( q_{ij}(\theta) \) is conditioned. Depending upon the memory of the process, \( P(\theta) \) will provide a substitute to \( L(\theta) \) that will yield more or less efficient estimators. For “near” Markovian processes, the proposed approach may not sacrifice much efficiency when using \( k_{ij} = 1 \) or 2 (e.g.) while greatly simplifying simulation-based inference.

In practice, the choice of \( I_{ij} \) will be driven by practical considerations (such as reducing the number of mismatches). The maximum composite likelihood estimator possesses desired asymptotic properties stated below.

**Proposition 1** Under classical regularity conditions, the maximum composite likelihood estimator is consistent as \( n \to \infty \). Furthermore, denoting the true value of \( \theta \) by \( \theta^* \), we have that

\[
\sqrt{n}(\tilde{\theta} - \theta^*) \to N(0, A(\theta^*)^{-1}B(\theta^*)A(\theta^*)^{-1}),
\]

where

\[
A(\theta) = \lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^{n} \sum_{j=1}^{n_i} \mathbb{E}_\theta \left\{ \frac{1}{q_{ij}(\theta)^2} \frac{\partial q_{ij}(\theta)}{\partial \theta} \frac{\partial q_{ij}(\theta)}{\partial \theta'} \right\}
\]

and

\[
B(\theta) = \lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^{n} \text{var}_\theta \left\{ \sum_{j=1}^{n_i} \frac{1}{q_{ij}(\theta)} \frac{\partial q_{ij}(\theta)}{\partial \theta} \right\}.
\]
Composite likelihood for branching processes

The proof of Proposition 1 rests on standard arguments used to establish asymptotic properties of \( M \)-estimators (for instance, see van der Vaart, 1998) and of multinomial likelihood estimators (Andersen, 1980). In particular, the consistency relies on the fact that the composite likelihood function satisfies the inequality \( \mathbb{E}_{\theta^*} \{ P(\theta) \} \leq \mathbb{E}_{\theta^*} \{ P(\theta^*) \} \) for every \( \theta \), with the equality holding true for \( \theta = \theta^* \) only if a condition on the identifiability of the model parameters applies. The inequality is easily established by noticing that each individual terms \( \log q_{ij}(\theta) \) is a log-likelihood object and therefore satisfies the classical inequality \( \mathbb{E}_{\theta^*} \{ \log q_{ij}(\theta) \} \leq \mathbb{E}_{\theta^*} \{ \log q_{ij}(\theta^*) \} \). The matrices \( A(\theta) \) and \( B(\theta) \) will generally be different from each other, and, in most cases, the composite likelihood estimator will not be fully efficient. Lindsay (1988) commented further on the properties of composite likelihood estimators.

4.2 Practical considerations

The probabilities \( q_{ij}(\theta) \) will generally have no closed-form expressions and estimation of \( \theta \) cannot proceed through direct maximization of \( P(\theta) \). We describe below 2 algorithms designed to compute a composite likelihood estimate for \( \theta \). Both algorithms use bias-corrected simulation-based approximations to \( \log q_{ij}(\theta) \) that are computed in the following way.

Let \( C_{s}(\theta, U_{s}), s = 1, \ldots, S \), be \( S \) independent clones simulated under the assumed model for the parameter value \( \theta \), where \( U_{s} \) denotes a set of r.v.’s uniformly distributed on the interval \([0, 1]\). These r.v.’s are used as seeds to generate the lifespans and the progeny vectors of all cells in the \( s \)th simulated clone. Write \( U^{(S)} = (U_{1}, \ldots, U_{S}) \). Let \( Y^{s}_{ij}(\theta, U_{s}) = Y^{s}_{ij,1}(\theta, U_{s}), \ldots, Y^{s}_{ij,D}(\theta, U_{s}) \) be a vector of cell counts for the \( D \) morphologically distinguishable cell types for the \( s \)th simulated clone at time \( t_{ij} \). To reduce the notational burden, we shall omit the notation \( U^{(S)} \) whenever unnecessary, and, for instance, write \( Y^{s}_{ij}(\theta, U_{s}) \) in place of \( Y^{s}_{ij}(\theta, U_{s}) \). Let \( N_{ij}(\theta; I_{ij}) = N_{ij}(\theta, U^{(S)}; I_{ij}) \) denote the number of simulated clones whose composition matches that of the \( i \)th experimental clone at the time points \( \{ t_{i,j-k}, \forall k \in I_{ij} \} \).

Formally, we have

\[
N_{ij}(\theta; I_{ij}) = \sum_{s=1}^{S} 1\{ Y^{s}_{i,j-k}(\theta) = Y_{i,j-k}, \forall k \in I_{ij} \},
\]

where \( 1\{ \cdot \} \) is the indicator function. The conditional probability \( q_{ij}(\theta) \) can be approximated by

\[
\hat{q}_{ij}^{S}(\theta) = \hat{q}_{ij}^{S}(\theta, U^{(S)}) = N_{ij}(\theta; I_{ij} \cup \{0\})/N_{ij}(\theta; I_{ij}),
\]

and a natural substitution in \( P(\theta) \) produces the simulated composite likelihood function

\[
P^{S}(\theta, U^{(S)}) = \sum_{i=1}^{n} \sum_{j=1}^{n_{i}} \log(\hat{q}_{ij}^{S}(\theta)).
\]

This simulated composite likelihood, however, suffers from the same limitations as those previously mentioned for the simulated log-likelihood function; that is, it is a biased estimator of \( P(\theta) \), and it is not defined when \( N_{ij}(\theta; I_{ij}) = 0 \). We mitigate both limitations simultaneously by considering a bias-corrected version of \( P(\theta) \) that is built as follows.

Conditional on \( \{ N_{ij}(\theta; I_{ij}) = s_{ij} \} \) and on \( \{ Y^{s}_{i,j-k}(\theta) = Y_{i,j-k}, \forall k \in I_{ij}, \forall s = 1, \ldots, s_{ij} \} \), we have that \( N_{ij}(\theta; I_{ij} \cup \{0\}) \sim \text{Bin}(s_{ij}, q_{ij}(\theta)) \), where \( \text{Bin}(s, q) \) denotes the binomial distribution with parameters \( s \) and \( q \). Following Pettigrew and others (1986), a bias-corrected estimator of \( \log q_{ij}(\theta) \) takes the form

\[
\hat{l}_{ij}^{S}(\theta) = \log \left( \frac{N_{ij}(\theta; I_{ij} \cup \{0\}) + 1/2}{s_{ij} + 1/2} \right).
\]
The bias of $\hat{l}_{ij}^S(\theta)$ is of smaller order—compared to that of $\log \tilde{q}_{ij}^S(\theta)$—when $s_{ij}q_{ij}(\theta) \to \infty$. When $s_{ij}q_{ij}(\theta) \to 0$, however, the bias of either approximation increases proportionally to $s_{ij}^{-2}q_{ij}(\theta)^{-2}$. Therefore, even with the biased-corrected estimator of $\log q_{ij}(\theta)$, it will remain important to reduce as much as possible the number of mismatches so the simulation-based approximation to the objective function will not be too biased. The performance of the estimator will otherwise be affected. The composite likelihood approach offers flexibility in choosing an estimating function that reduces the number of mismatches or even eliminates them. Note that $\hat{l}_{ij}^S(\theta)$ remains defined when there are mismatches.

This motivated us to consider the simulated composite likelihood defined as

$$P_{bc}^S(\theta) = P_{bc}^S(\theta, U^{(S)}) = \sum_{i=1}^{n} \sum_{j=1}^{n_i} \hat{l}_{ij}^S(\theta).$$

There exist multiple ways to compute an estimator of $\theta$ based on $P_{bc}^S(\theta)$. Two methods are discussed below.

**Method 1: sample path.** In order to estimate $\theta$, a first approach consists in maximizing the simulated composite likelihood $P_{bc}^S(\theta, U^{(S)})$ for a given vector $U^{(S)}$. In doing so, the simulated composite likelihood function is not subject to random fluctuations during the optimization. The resultant estimator can be proven to converge to $\tilde{\theta}$ as $S \to \infty$ (see Hyrien, 2007, for a proof of a similar result). Due to the discrete nature of branching processes, the function $P_{bc}^S(\theta, U^{(S)})$ is discontinuous, and optimizations algorithms that do not rest upon differentiability of the objective function (such as the simplex method of Nelder and Mead) can be used.

**Method 2: stochastic approximation.** Alternatively, $\theta$ can be estimated using a multidimensional version of stochastic procedure of Kiefer–Wolfowitz (Kiefer and Wolfowitz, 1952; Blum, 1954). In this alternate schema, an estimate of $\theta$ is computed through the sequence of r.v.'s $\{\theta(k)\}_{k \geq 1}$ defined as $\theta^{(k+1)} = \theta^{(k)} + a_k \Delta_k$, where $\Delta_k = (\Delta_{k,1}, \ldots, \Delta_{k,q})'$, with

$$\Delta_{k,l} = \{P_{bc}^S(\theta^{(k)} + b_k e_l, U_{k+}^{(S)}) - P_{bc}^S(\theta^{(k)} - b_k e_l, U_{k-}^{(S)})\}/2b_k,$$

where $U_{k-}^{(S)}$ and $U_{k+}^{(S)}$, $k = 1, 2, \ldots$, are i.i.d. copies of $U^{(S)}$, where $e_l, l = 1, \ldots, q$, are $q \times 1$ vectors with all elements equal to 0, except for 1 in the $l$th place, and where $(a_k)_{k \geq 1}$ and $(b_k)_{k \geq 1}$ are 2 given sequences if nonrandom positive numbers such that

$$\lim_{k \to \infty} b_k = 0, \quad \sum_{k=1}^{\infty} a_k = \infty, \quad \sum_{k=1}^{\infty} a_k b_k < \infty, \quad \text{and} \quad \sum_{k=1}^{\infty} (a_k/b_k)^2 < \infty.$$

Usual choices for $(a_k)_{k \geq 1}$ and $(b_k)_{k \geq 1}$ include the following: $a_k = a_0/k^{\alpha}$ and $b_k = b_0/k^{\beta}$, for some given positive constants $a_0$, $b_0$, $\alpha$, and $\beta$. In our analysis of clonal data on O-2A progenitor cells, we used $a_0 = 1$, $b_0 = 1$, $\alpha = 1$, and $\beta = 1/3$. Other choices are possible. If $P_{bc}^{(S)}(\theta)$ was an unbiased estimator for $P(\theta)$, $\theta^{(k)}$ would converge to $\tilde{\theta}$ as $k \to \infty$. This is not formally the case, but because $P_{bc}^S(\theta)$ is first-order bias-corrected, $\theta^{(k)}$ is expected converge to a value close to $\tilde{\theta}$.

The computation of composite likelihood estimates may take up to more than a day, depending on the number of simulations $S$, the number of model parameters, and the sample size. The number of mismatches increases substantially with the number of conditioning observations, and eliminating the mismatches may require increasing the number of simulations out of practical limits. The numbers of simulations being equal, the stochastic approximation procedure is expected to provide more accurate
estimates than the sample path approach since the noise arising from the simulations is averaged out as the sequence \( \{ \theta^{(k)} \}_{k \geq 1} \) progresses. This procedure is more time-consuming because it requires evaluation of the objective function at \( 2q \) points at every step \( k \), while the sample path approach needs only one such evaluation at every step. In practice, it is probably best to begin with the sample path approach and complete estimation using stochastic approximation, with \( a_0 \) and \( b_0 \) appropriately chosen to avoid large jumps of the sequence \( \{ \theta^{(k)} \}_{k \geq 1} \).

### 5. Simulation Studies

We performed simulation studies to evaluate the performance of the proposed simulated composite likelihood estimator in samples of finite sizes. We also compared it with the simulated pseudo maximum likelihood estimator proposed in past publications (Hyrien and others, 2005a,b; Hyrien, 2007). In order to keep the computing time reasonable (each estimate required between 1 h of computing time with \( S = 10 000 \) simulated clones, and up to 6 h with \( S = 50 000 \) simulated clones), the study was conducted using 2 simple branching processes that are special cases of our model of the generation of oligodendrocytes. The conclusions of these studies are therefore not meant to be exhaustive but they corroborate what can be intuitively expected. The 2 processes that we considered were defined as follows:

**Model 1 (a binary splitting process)**: This model assumes that every cell divides into 2 new cells upon completion of its lifespan, which follows a gamma distribution with mean \( m \) and variance \( \sigma^2 \). When \( m = \sigma \), the process reduces to the (Markov) linear birth process.

**Model 2 (a 2-type process)**: This model considers 2 cell types: every type-1 cell divides into 2 new type-0 cells with probability \( p \) or it turns into a single type-0 cell with probability \( 1 - p \). The lifespan of any type-1 cell follows a gamma distribution with mean \( m \) and variance \( \sigma^2 \). Type-0 cells have infinite lifespans.

Longitudinal clonal data were simulated from these 2 models, and cell clones were observed every day from day 1 to day 6 (same experimental design for both models). We considered various sample sizes: \( n = 20, 50, 100 \); and used \( S = 10 000 \) and \( S = 20 000 \) simulated clones to compute the simulated composite likelihoods for Models 1 and 2, respectively. For each simulated data set, we computed the simulated composite likelihood under various conditioning sets, all of the form \( I_{ij} = \emptyset \) or \( I_{ij} = \{1, \ldots, j \wedge j_0\} \), for \( j_0 = 1, \ldots, 5 \), where \( j \wedge j_0 \) denotes the smallest value of \( j \) and \( j_0 \). The particular case \( I_{ij} = \{1, \ldots, j \wedge 5\} \) corresponds to a simulated maximum likelihood estimator. The pseudo maximum likelihood estimator (Hyrien and others, 2005a,b; Hyrien, 2007), computed for comparison, is defined as the parameter vector that maximizes the pseudolikelihood function

\[
G(\theta) = - \sum_{i=1}^{n} \left[ \left( Y_i - m_i(\theta) \right)' V_i(\theta)^{-1} \left( Y_i - m_i(\theta) \right) \right] + \log \det V_i(\theta),
\]

where \( m_i(\theta) = E_{\theta}(Y_i) \) and \( V_i(\theta) = \text{Var}_{\theta}(Y_i) \), and where \( \det V \) denotes the determinant of the matrix \( V \). When these moments have no closed-form explicit expressions, they can be replaced by approximations computed either from simulations (Hyrien and others, 2005a,b; Hyrien, 2007) or using the method of saddlepoint (Hyrien and others, 2010). The pseudo maximum likelihood estimator is consistent and asymptotically Gaussian (Hyrien, 2007). A summary of the simulation results is displayed in Tables 1 and 2.

The biases of the estimators tended to decrease as the sample size increased. The standard errors depended on the method of estimation being used. Overall the pseudo maximum likelihood estimator appeared to be the least efficient of all tested estimators. This was particularly true for the smallest sample sizes.
Table 1. Results of the simulation study based on a binary splitting process (see text for explanation).

\( CL_k \) = simulated composite likelihood with \( I_{ij} = \{1, 2 \ldots, j \wedge k\} \); \( ML \) = simulated maximum likelihood

\( PL \) = simulated pseudolikelihood; s.e. = standard error

<table>
<thead>
<tr>
<th>Sample size</th>
<th>Parameter true value method</th>
<th>( m ) 40 h</th>
<th>Mean</th>
<th>s.e.</th>
<th>( \sigma ) 20 h</th>
<th>Mean</th>
<th>s. e.</th>
</tr>
</thead>
<tbody>
<tr>
<td>20</td>
<td>ML = CL_5</td>
<td>41.96</td>
<td>2.05</td>
<td>19.16</td>
<td>1.24</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>CL_4</td>
<td>42.31</td>
<td>1.65</td>
<td>19.26</td>
<td>1.11</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>CL_3</td>
<td>41.99</td>
<td>2.02</td>
<td>19.31</td>
<td>1.34</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>CL_2</td>
<td>41.38</td>
<td>1.93</td>
<td>18.93</td>
<td>1.67</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>CL_1</td>
<td>40.99</td>
<td>3.05</td>
<td>18.50</td>
<td>2.63</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Marginal CL</td>
<td>40.97</td>
<td>4.02</td>
<td>18.50</td>
<td>3.90</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>PL</td>
<td>42.31</td>
<td>4.70</td>
<td>16.98</td>
<td>3.51</td>
<td></td>
<td></td>
</tr>
<tr>
<td>50</td>
<td>ML = CL_5</td>
<td>40.66</td>
<td>1.50</td>
<td>20.04</td>
<td>1.04</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>CL_4</td>
<td>40.52</td>
<td>1.34</td>
<td>19.94</td>
<td>1.20</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>CL_3</td>
<td>40.41</td>
<td>1.37</td>
<td>19.97</td>
<td>1.11</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>CL_2</td>
<td>40.37</td>
<td>1.49</td>
<td>19.92</td>
<td>1.29</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>CL_1</td>
<td>40.03</td>
<td>1.59</td>
<td>19.59</td>
<td>1.98</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Marginal CL</td>
<td>39.67</td>
<td>2.30</td>
<td>19.44</td>
<td>2.74</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>PL</td>
<td>41.23</td>
<td>2.73</td>
<td>18.72</td>
<td>2.32</td>
<td></td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>ML = CL_5</td>
<td>40.14</td>
<td>1.17</td>
<td>20.12</td>
<td>1.13</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>CL_4</td>
<td>40.22</td>
<td>1.15</td>
<td>20.25</td>
<td>1.37</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>CL_3</td>
<td>40.18</td>
<td>1.19</td>
<td>20.21</td>
<td>1.18</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>CL_2</td>
<td>39.86</td>
<td>1.38</td>
<td>19.73</td>
<td>1.50</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>CL_1</td>
<td>39.67</td>
<td>1.61</td>
<td>19.49</td>
<td>1.84</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Marginal CL</td>
<td>39.69</td>
<td>1.63</td>
<td>19.39</td>
<td>2.03</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>PL</td>
<td>40.31</td>
<td>2.27</td>
<td>19.27</td>
<td>2.37</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Sizes. The simulated maximum likelihood estimator was not particularly more efficient than the simulated composite likelihood estimators. For instance, in the case of Model 2, the simulations indicated that the marginal composite likelihood estimator and the composite likelihood estimators defined by conditioning upon the most recent one or two observations had smaller standard errors. This was particularly true for the largest sample sizes. The most likely explanation of the fact that some estimators performed worse when we conditioned upon more observations is that the simulation-based approximations to the conditional probabilities that make up the expression for the estimating functions are more noisy in the latter case (due to the simulations) than when we condition on fewer observations.

We also found that the bias decreases generally quickly with increasing sample size but not always. For instance, the estimator of the probability of division (Model 2), say \( \hat{p} \), remained substantially biased even with \( n = 100 \) observed clones. When conditioning upon the most recent observation only, the bias of \( \hat{p} \) was about 0.06 with a sample size of \( n = 20 \) clones and close to 0.05 with a sample size of \( n = 50 \) clones. The bias became larger as the composite likelihood estimator was conditioning upon an increasing number of past observations. We performed additional simulations to assess whether this bias could be attributed (in part at least) to a value of \( S \) that was too small. Thus, we considered \( S = 50000 \) simulated clones (instead of \( S = 20000 \)). As a result of time constraints, we restricted our investigations to the above 2 cases and found that the bias of \( \hat{p} \) decreased to 0.03 both when \( n = 20 \) and when \( n = 50 \), and thus was reduced by almost a factor of 2. Also, the standard error of \( \hat{p} \) decreased to 0.06 and 0.04, respectively. The estimators of \( m \) and \( \sigma \) improved also in this last run of simulations, with standard errors smaller than those reported in Table 2.
Table 2. Results of the simulation study based on a 2-type process (see text for explanation). $CL_k = \text{simulated composite likelihood with } I_{ij} = [1, 2, \ldots, J \wedge k]$; $ML = \text{simulated maximum likelihood}; PL = \text{simulated pseudolikelihood}; s.e. = \text{standard error}$

<table>
<thead>
<tr>
<th>Sample size</th>
<th>Parameter true value method</th>
<th>$m_{35 \text{ h}}$ Mean</th>
<th>$m_{35 \text{ h}}$ s.e.</th>
<th>$\sigma_{10 \text{ h}}$ Mean</th>
<th>$\sigma_{10 \text{ h}}$ s.e.</th>
<th>$P_{0.35}$ Mean</th>
<th>$P_{0.35}$ s.e.</th>
</tr>
</thead>
<tbody>
<tr>
<td>20</td>
<td>ML = CL$_5$</td>
<td>35.83</td>
<td>2.08</td>
<td>10.31</td>
<td>2.65</td>
<td>0.27</td>
<td>0.09</td>
</tr>
<tr>
<td></td>
<td>CL$_4$</td>
<td>35.97</td>
<td>1.90</td>
<td>10.24</td>
<td>2.59</td>
<td>0.28</td>
<td>0.09</td>
</tr>
<tr>
<td></td>
<td>CL$_3$</td>
<td>36.05</td>
<td>2.07</td>
<td>10.01</td>
<td>2.57</td>
<td>0.27</td>
<td>0.09</td>
</tr>
<tr>
<td></td>
<td>CL$_2$</td>
<td>36.16</td>
<td>2.04</td>
<td>10.15</td>
<td>2.52</td>
<td>0.29</td>
<td>0.09</td>
</tr>
<tr>
<td></td>
<td>CL$_1$</td>
<td>35.89</td>
<td>2.02</td>
<td>10.34</td>
<td>2.75</td>
<td>0.29</td>
<td>0.09</td>
</tr>
<tr>
<td></td>
<td>Marginal CL</td>
<td>35.50</td>
<td>2.05</td>
<td>9.64</td>
<td>2.42</td>
<td>0.33</td>
<td>0.07</td>
</tr>
<tr>
<td></td>
<td>PL</td>
<td>36.30</td>
<td>2.32</td>
<td>12.34</td>
<td>2.98</td>
<td>0.30</td>
<td>0.09</td>
</tr>
<tr>
<td>50</td>
<td>ML = CL$_5$</td>
<td>35.61</td>
<td>1.58</td>
<td>9.73</td>
<td>1.42</td>
<td>0.29</td>
<td>0.06</td>
</tr>
<tr>
<td></td>
<td>CL$_4$</td>
<td>35.59</td>
<td>1.53</td>
<td>9.60</td>
<td>1.36</td>
<td>0.29</td>
<td>0.07</td>
</tr>
<tr>
<td></td>
<td>CL$_3$</td>
<td>35.41</td>
<td>1.56</td>
<td>9.63</td>
<td>1.47</td>
<td>0.30</td>
<td>0.07</td>
</tr>
<tr>
<td></td>
<td>CL$_2$</td>
<td>35.55</td>
<td>1.51</td>
<td>9.60</td>
<td>1.40</td>
<td>0.30</td>
<td>0.06</td>
</tr>
<tr>
<td></td>
<td>CL$_1$</td>
<td>35.54</td>
<td>1.67</td>
<td>9.59</td>
<td>1.23</td>
<td>0.30</td>
<td>0.06</td>
</tr>
<tr>
<td></td>
<td>Marginal CL</td>
<td>35.49</td>
<td>1.24</td>
<td>9.72</td>
<td>1.40</td>
<td>0.34</td>
<td>0.05</td>
</tr>
<tr>
<td></td>
<td>PL</td>
<td>35.30</td>
<td>1.63</td>
<td>10.56</td>
<td>1.82</td>
<td>0.33</td>
<td>0.05</td>
</tr>
<tr>
<td>100</td>
<td>ML = CL$_5$</td>
<td>35.78</td>
<td>1.26</td>
<td>9.82</td>
<td>1.30</td>
<td>0.27</td>
<td>0.06</td>
</tr>
<tr>
<td></td>
<td>CL$_4$</td>
<td>35.30</td>
<td>1.20</td>
<td>9.48</td>
<td>1.20</td>
<td>0.27</td>
<td>0.06</td>
</tr>
<tr>
<td></td>
<td>CL$_3$</td>
<td>35.74</td>
<td>1.24</td>
<td>10.07</td>
<td>1.43</td>
<td>0.28</td>
<td>0.05</td>
</tr>
<tr>
<td></td>
<td>CL$_2$</td>
<td>35.60</td>
<td>1.34</td>
<td>9.70</td>
<td>1.48</td>
<td>0.28</td>
<td>0.05</td>
</tr>
<tr>
<td></td>
<td>CL$_1$</td>
<td>35.60</td>
<td>1.12</td>
<td>9.75</td>
<td>1.04</td>
<td>0.31</td>
<td>0.05</td>
</tr>
<tr>
<td></td>
<td>Marginal CL</td>
<td>35.12</td>
<td>1.03</td>
<td>9.90</td>
<td>1.08</td>
<td>0.34</td>
<td>0.04</td>
</tr>
<tr>
<td></td>
<td>PL</td>
<td>35.52</td>
<td>1.16</td>
<td>11.34</td>
<td>1.45</td>
<td>0.34</td>
<td>0.04</td>
</tr>
</tbody>
</table>

6. AN APPLICATION TO OLIGODENDROCYTES GENERATION EXPOSED TO NT-3

Oligodendrocytes produce the myelin sheaths that enwraps axons in the central nervous system. They are involved in signal propagation along the nerves and are particularly essential for normal impulse conduction. Some demyelinating diseases, such as multiple sclerosis, have been shown to be associated with injury or dysfunction of the oligodendrocyte population (Blakemore, 2008; Korn, 2008). The biology of oligodendrocytes and their lineage is well documented in the literature (e.g. Franklin and Ffrench-Constant, 2008). Oligodendrocytes are assumed to be nonproliferating, terminally differentiated cells, and are generated from their immediate precursor cells, the O-2A progenitor cells. These O-2A progenitor cells have the ability to self-renew by dividing, to differentiate into oligodendrocytes, or to die. Due to the critical role of oligodendrocytes, the proper regulation of the differentiation of O-2A progenitor cells into oligodendrocytes is therefore particularly important for the normal functioning of the central nervous system. O-2A progenitor cells have been shown to respond to the presence of various environmental signals or factors that may modulate their ability to self-renew and differentiate, and one of the primary objectives of the research performed on the oligodendrocyte lineage is to identify factors that could restore the normal balance in the generation of oligodendrocytes.

NT-3 is one member of a family of growth factors known as neurotrophins. Althaus and others (2008) gave an overview on neurotrophins and their effects on oligodendroglial cells. It has been demonstrated that, in rodents and humans at least, NT-3 may be synthesized by oligodendrocytes. Although NT-3 has been shown to be involved in the growth, differentiation, and survival of neurons (Zhou and Rush, 1996;
Kalb, 2005), its role (and that of other neurotrophins as well) on the regulation of nonneuronal cells of the central nervous system is still unclear.

In an attempt to better understand the effect of NT-3 on the regulation of the processes of division and differentiation of O-2A progenitor cells, we conducted the longitudinal clonal experiments presented in Section 2.2. In particular, the objectives of this experiment were to assess the hypothesis that the presence of NT-3 alone (i.e. without coexposure to PDGF) in pure cultures has no effect on the proliferation and differentiation of O-2A progenitor cells, as suggested by Barres and others, 1994, and to gain some quantitative insight into how NT-3 may alter division and differentiation of O-2A progenitor cells.

Histograms for the number of O-2A progenitor cells and for the number of oligodendrocytes per clone with and without NT-3 suggested a striking difference between these 2 experimental conditions in respect to the number of O-2A progenitor cells. When cells were cultured without NT-3, the number of O-2A progenitor cells appeared to increase over time, suggesting that these cells maintained some capacity for self-renewal after 6 days of culture even in the absence of PDGF. By day 6, more than 60% of the clones still contained at least one O-2A progenitor cell, and the clones contained on average 1.9 O-2A progenitor cells and 2.1 oligodendrocytes. In contrast, in clones grown in the presence of NT-3, O-2A progenitor cells underwent division first, just as when they were cultured without NT-3, and then appeared to differentiate massively. By day 6, less than 20% of the clones still contained at least one O-2A progenitor cell. The mean numbers of oligodendrocytes per clone (\(\geq 1.9\)) was not much different than that without NT-3, but the number of O-2A progenitor cells was considerably lower (\(\leq 0.8\)).

We investigate further the effects of NT-3 on the processes of division and differentiation of O-2A progenitor cells to determine whether NT-3 affected the decision of O-2A progenitor cells to divide or differentiate the timing of cellular division or the timing of cellular differentiation, or all of the above. Changes in cell fate would be seen through either an increase or a decrease of the probability of division of O-2A progenitor cells, whereas a modification of the time it takes for O-2A progenitor cells to divide or to differentiate would be reflected by changes in the parameters of the distribution of the time to division and/or of the distribution of the time to differentiation of these cells.

We fitted our branching process model using the proposed simulated composite likelihood approach. The ultimate parameter estimates were obtained using 15,000 simulated clones to approximate the distribution of the number of cells of each type per clone under the above model. We started from multiple initial values and let optimization run for more than a day each time. We used the conditioning sets \(I_{ij} = \emptyset, \{1\}, \{1, j \wedge 2\}, \ldots, \{1, 2, \ldots, j \wedge 5\}\). In the last case, we obtain a simulated maximum likelihood estimator. The model was fitted separately to the data obtained with and without NT-3 because none of the model parameters were common to both experimental conditions.

In our experiments, the follow-up period was long enough so all initiator precursor cells had either divided or differentiated by the time the experiment ended. By determining the composition of a clone, it is possible to decipher whether the corresponding initiator cell divided or differentiated (in the latter case, the clone always contains a single oligodendrocyte). Thus, let \(N_{\text{div}}\) and \(N_{\text{diff}} = n - N_{\text{div}}\) denote the numbers of initiator cells having divided and differentiated by the end of the experiment, respectively. For every \(0 \leq j_1 < j_2\), let \(N_{\text{div}}(j_1, j_2)\) denote the number of initiator cells having divided between day \(j_1\) and \(j_2\). Define the vector \(N_{\text{div}} = \{N_{\text{div}}(t - 1, t); t = 1, \ldots, 6\}\). Conditional on \(N_{\text{div}} = N\), \(N_{\text{div}}\) follows a multinomial distribution with parameters \(N\) and \(\{p_1(m_1, \sigma_1), \ldots, p_6(m_1, \sigma_1)\}\), where \(p_t(m, \sigma) = G_1(t; m, \sigma) - G_1(t - 1; m, \sigma)\) denotes the probability that any initiator cell that will ultimately divide, completes its cycle between day \(t - 1\) and \(t\). The estimates for \(m_1\) and \(\sigma_1\) can therefore be computed by maximizing the multinomial likelihood function associated with the vector of counts \(N_{\text{div}}\). The mean and variance of the time to differentiation of O-2A progenitor cells can be estimated in a similar fashion for clones starting with an initiator cell that ended up differentiating. We used these maximum likelihood estimates as starting values in our composite likelihood analyses.
Table 3 reports parameter estimates resulting from these analyses. Also reported are the total numbers of mismatches for the final parameter estimates. In our analyses of cell clones cultured without NT-3, the number of mismatches remained low (0-1) until we started to condition upon the most recent 3 (or more) observations. For cell clones exposed to NT-3, the number of mismatches increased substantially when we started to condition upon the most recent observations. In either case, the largest number of mismatches we encountered ranged between 6 and 9, which is relatively high compared to the number of cell clones observed for each experimental condition (n = 20). Alternatively, we could have increased the number of simulations to reduce the number of mismatches, but each analysis required between 1 and 2 days. We therefore discuss results obtained with the marginal composite likelihood (no conditioning) for cells exposed to NT-3, and with the composite likelihood conditioning upon the most recent 2 observations for control (unexposed) cells. Standard errors for the corresponding parameter estimates were obtained using a parametric bootstrap approach. We compared parameter estimates across treatment groups using a Wald test. The fitted marginal distribution for the number of O-2A progenitor cells and for the number of oligodendrocytes per clone is displayed in Figure 2 as solid lines. The estimated probabilities of division of O-2A progenitor cells are presented in Figure 3 (panels A and B) as a function of the cell generation for each culture condition. These analyses pointed out the following conclusions:

(a) In each experiment, the probability of division of O-2A progenitor cells was found to decrease when the number of divisions increased. This is consistent with the results of previously analyzed data sets (Yakovlev, Boucher, and others, 1998; Yakovlev, Mayer-Pröschel, and others, 1998; Yakovlev and others, 2000; von Collani and others, 1999; Boucher and others, 1999, 2001; Zorin and others, 2000; Hyrien and others, 2005a,b, 2006).

(b) Exposure to NT-3 decreased the probability of division of O-2A progenitor cells. For instance, cells of generation 3 divided with a probability estimated as 0.35 (vs. 0.53) when cells were exposed (vs. unexposed) to NT-3. This may explain that the numbers of O-2A progenitor cells were

---

Table 3. Estimates of the model parameters (means and standard deviations expressed in hours). See text for explanations. CL<sub>k</sub> = simulated composite likelihood with I<sub>i j</sub> = \{1, 2, ..., j ∧ k\}; ML = simulated maximum likelihood; s.e. = standard errors

<table>
<thead>
<tr>
<th>Method</th>
<th>Mismatches</th>
<th>m&lt;sub&gt;11&lt;/sub&gt;</th>
<th>σ&lt;sub&gt;11&lt;/sub&gt;</th>
<th>m&lt;sub&gt;12&lt;/sub&gt;</th>
<th>σ&lt;sub&gt;12&lt;/sub&gt;</th>
<th>m&lt;sub&gt;2&lt;/sub&gt;</th>
<th>σ&lt;sub&gt;2&lt;/sub&gt;</th>
<th>α</th>
<th>b</th>
<th>c</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>With NT3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Marginal CL</td>
<td>1</td>
<td><strong>39.9</strong></td>
<td><strong>17.2</strong></td>
<td><strong>49.3</strong></td>
<td><strong>46.9</strong></td>
<td><strong>36.0</strong></td>
<td><strong>29.2</strong></td>
<td>0.12</td>
<td>0.67</td>
<td>0.70</td>
</tr>
<tr>
<td>(s.e.)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CL&lt;sub&gt;1&lt;/sub&gt;</td>
<td>5</td>
<td><strong>39.4</strong></td>
<td><strong>17.1</strong></td>
<td><strong>53.8</strong></td>
<td><strong>44.2</strong></td>
<td><strong>31.0</strong></td>
<td><strong>31.8</strong></td>
<td>0.09</td>
<td>0.70</td>
<td>0.59</td>
</tr>
<tr>
<td>CL&lt;sub&gt;2&lt;/sub&gt;</td>
<td>7</td>
<td><strong>40.4</strong></td>
<td><strong>15.4</strong></td>
<td><strong>51.3</strong></td>
<td><strong>45.2</strong></td>
<td><strong>31.3</strong></td>
<td><strong>31.7</strong></td>
<td>0.09</td>
<td>0.70</td>
<td>0.58</td>
</tr>
<tr>
<td>CL&lt;sub&gt;3&lt;/sub&gt;</td>
<td>7</td>
<td><strong>37.5</strong></td>
<td><strong>16.5</strong></td>
<td><strong>52.4</strong></td>
<td><strong>42.5</strong></td>
<td><strong>31.8</strong></td>
<td><strong>30.9</strong></td>
<td>0.09</td>
<td>0.68</td>
<td>0.55</td>
</tr>
<tr>
<td>CL&lt;sub&gt;4&lt;/sub&gt;</td>
<td>7</td>
<td><strong>39.4</strong></td>
<td><strong>16.4</strong></td>
<td><strong>54.3</strong></td>
<td><strong>45.4</strong></td>
<td><strong>30.1</strong></td>
<td><strong>32.6</strong></td>
<td>0.08</td>
<td>0.68</td>
<td>0.62</td>
</tr>
<tr>
<td>ML = CL&lt;sub&gt;5&lt;/sub&gt;</td>
<td>7</td>
<td><strong>38.3</strong></td>
<td><strong>16.7</strong></td>
<td><strong>53.3</strong></td>
<td><strong>43.9</strong></td>
<td><strong>29.8</strong></td>
<td><strong>30.8</strong></td>
<td>0.07</td>
<td>0.70</td>
<td>0.62</td>
</tr>
<tr>
<td>Without NT3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Marginal CL</td>
<td>0</td>
<td>23.7</td>
<td>22.9</td>
<td>51.0</td>
<td>24.2</td>
<td>40.2</td>
<td>28.7</td>
<td>0.18</td>
<td>0.90</td>
<td>0.74</td>
</tr>
<tr>
<td>CL&lt;sub&gt;1&lt;/sub&gt;</td>
<td>0</td>
<td>31.0</td>
<td>15.9</td>
<td>44.6</td>
<td>25.3</td>
<td>44.0</td>
<td>24.4</td>
<td>0.12</td>
<td>0.78</td>
<td>0.85</td>
</tr>
<tr>
<td>CL&lt;sub&gt;2&lt;/sub&gt;</td>
<td>1</td>
<td><strong>31.5</strong></td>
<td><strong>14.2</strong></td>
<td><strong>44.0</strong></td>
<td><strong>27.7</strong></td>
<td><strong>43.8</strong></td>
<td><strong>24.2</strong></td>
<td><strong>0.10</strong></td>
<td><strong>0.80</strong></td>
<td><strong>0.81</strong></td>
</tr>
<tr>
<td>(s.e)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CL&lt;sub&gt;3&lt;/sub&gt;</td>
<td>7</td>
<td>29.2</td>
<td>15.8</td>
<td>42.5</td>
<td>22.6</td>
<td>43.8</td>
<td>25.7</td>
<td>0.11</td>
<td>0.73</td>
<td>0.80</td>
</tr>
<tr>
<td>CL&lt;sub&gt;4&lt;/sub&gt;</td>
<td>6</td>
<td>29.2</td>
<td>15.8</td>
<td>42.5</td>
<td>22.6</td>
<td>43.8</td>
<td>25.7</td>
<td>0.11</td>
<td>0.73</td>
<td>0.80</td>
</tr>
<tr>
<td>ML = CL&lt;sub&gt;5&lt;/sub&gt;</td>
<td>9</td>
<td>26.8</td>
<td>15.2</td>
<td>45.9</td>
<td>26.2</td>
<td>44.9</td>
<td>23.9</td>
<td>0.08</td>
<td>0.72</td>
<td>0.79</td>
</tr>
</tbody>
</table>

Analysis results were discussed based on the bolded parameter estimates. Bolded values were selected for the discussion.
much smaller in the presence of NT-3. The difference between the probabilities of division across treatment groups was almost significant ($p = 0.07$).

(c) In both culture conditions, the mean time to division was shorter for initiator O-2A progenitor cells than for O-2A progenitor cells of subsequent generations (40 vs. 50 h with NT3, and 31.5 vs. 44 h without NT3). This is consistent with results from previous studies based on time-lapse experiments (Hyrien and others, 2006).

(d) Exposure to NT-3 increased the mean time to division of O-2A progenitor cells (e.g. 50 vs. 44 h for cells of generation greater than or equal to 2) and decreased their mean time to differentiation (36 vs. 44 h). When comparing the distribution of the time to division across treatment groups using the Wald test, we found that the difference was significant ($p < 0.0001$). The effect of NT-3 on the time to differentiation was almost significant ($p = 0.05$).
The overall conclusion of our analysis is that NT-3 may modulate the proliferation of O-2A progenitor cells and their differentiation into oligodendrocytes, and that it does not act only in concert with PDGF. In particular, we found that NT-3 induces differentiation of O-2A progenitor cells into oligodendrocytes by decreasing the probability of division, by increasing the mean time to division, and by decreasing the mean time to differentiation. The results of our analysis are tentative and should be confirmed in other experiments.

The effect of NT-3 may be perceived as subtle when expressed as a change of the probability of division, of the distribution of the time to division, and of the distribution of the time to differentiation. To better appreciate the impact of such changes on the composition of cell clones over time, we simulated the fitted branching process over 20 days and calculated the mean number of O-2A progenitor cells and the mean number of oligodendrocytes cultured with and without NT-3 as a function of time (Figure 3). While the number of O-2A progenitor cells declines under either culture conditions (as expected from the theory of branching processes since \( \lim_{k \to \infty} p_k < 0.5 \)), the number of oligodendrocytes differed substantially. For instance, by day 20, the number of oligodendrocytes was 2-fold larger in untreated cells. This seems to contradict our conclusion that NT-3 induces differentiation, but it does not actually. Indeed the best strategy to increase the number of oligodendrocytes over a medium-term period is to first increase the pool of progenitor cells by self-renewal, and then have some of these progenitor cells differentiate into oligodendrocytes. This seems to be what happened when cells were not exposed to NT-3 compared to when they were.

The mean numbers of oligodendrocytes in clones exposed and unexposed to NT-3 drifted apart from each other after day 6 only. This would explain why our permutation test did not detect any significant difference between day 0 and day 6 in counts of oligodendrocytes.

7. DISCUSSION

The proposed simulated composite likelihood method provides a good alternative to existing methods for the analysis of clonal data using age-dependent branching processes. It possesses good properties both asymptotically and in samples of finite sizes. It appeared to outperform the pseudo maximum likelihood estimator proposed in past publications (Hyrien and others, 2005a,b; Hyrien, 2007), and it mitigates the computational issues encountered with the simulated likelihood approach (Zorin and others, 2000). Another advantage of the method is that it is adaptable to the complexity of the situation. It allows the selection of an estimating function that reduces the number of mismatches. The proposed method was framed within the context of a particular multitype age-dependent branching process of the generation of oligodendrocytes in vitro. The structure of this model could be modified, if needed, using the more general process described by Hyrien and others (2005a).

ACKNOWLEDGMENTS

The authors acknowledge the reviewers and an associate editor for their careful reading of the manuscript and for their suggestions. Conflict of Interest: None declared.

FUNDING

National Institutes of Health (1R01 CA134839-01, 2R01 NS039511, 1R01AI069351, R01 HD39702, 2P30 ES001247, T32 ES007271).

REFERENCES


[Received August 26, 2008; revised June 3, 2010; accepted for publication June 29, 2010]