Periodic chronic myelogenous leukaemia: spectral analysis of blood cell counts and aetiological implications

PASCAL FORTIN AND MICHAEL C. MACKEY* Departments of Physiology and Physics and *Departments of Physiology, Physics and Mathematics, Centre for Nonlinear Dynamics, McGill University, Montreal, Canada

Received 23 June 1998; accepted for publication 23 October 1998

Summary. Of 24 published clinical reports of periodic chronic myelogenous leukaemia (PCML), 21 had sufficient data to analyse for periodicity, and 12 showed significant periodicity (p < 0·05) using the Lomb periodogram. Leucocyte oscillations had periods $T$ ranging from 37 to 83 d. When data were also reported for platelets and reticulocytes there was no significant difference between their periods and those of the leucocytes. These data and their analysis provide strong circumstantial evidence for a haemopoietic stem cell origin of PCML. Namely, the known chromosomal changes in CML patients may, on occasion, also be accompanied by a destabilization resulting in an oscillatory efflux into the leucocyte, platelet and erythrocytic pathways.

Keywords: chronic myelogenous leukaemia, periodicity, spectral analysis, stem cell, apoptosis.

In 1967 four patients with CML whose leucocyte counts showed gross and apparently spontaneous oscillations with periods from 30 to 100 d were reported by Morley et al (1967). Two of these patients also showed oscillations in the platelet counts with the same periods as in the leucocyte counts. Other cases of putative cyclic leucocytosis have now been reported in CML (Chikkappa et al, 1976; Delobel et al, 1973; Gatti et al, 1973; German et al, 1976; Iizuka et al, 1984; Inbal et al, 1983; Kennedy, 1970; Mastrangelo et al, 1974, 1976; Mehta & Agarwal, 1980; Rodriguez & Lutcher, 1976; Shadduck et al, 1972; Umemura et al, 1986; Vodopick et al, 1972; Yamauchi & Ide, 1992). In most patients, oscillations in the leucocyte counts were accompanied by oscillations in the platelet and reticulocyte counts with the same periodicity. The presence of oscillations in more than one cell population strongly suggests that the oscillations originate in the stem cell compartment.

We report here an analysis of previously published serial blood counts from 24 PCML patients. Periodogram analysis was used to test for the presence of statistically significant periodic oscillations, and to find the phase relationship between the different cell populations.

DATA AND METHODS

Cases of PCML have been reported in 24 different clinical studies. Of these, the studies by German et al (1976) and Mehta & Agarwal (1980) were rejected because the serial blood counts were not included in the reports. Another was rejected because the serial blood counts were truncated (Shadduck et al, 1972). In the study by Kennedy (1970) the patients were treated for CML with hydroxyurea, and in that by Rodriguez & Lutcher (1976) they were treated with busulphan. In Kennedy (1970) no data were available to us before the initiation of hydroxyurea therapy. However, the data that we analysed were all collected during constant hydroxyurea dosage. In Rodriguez & Lutcher (1976) there was clear evidence for cycling before initiation of busulphan therapy which only decreased the amplitude of the cycling and left the period unchanged. The other patients were not treated for CML, which implied that the oscillations were occurring naturally and not induced by treatment.

The published serial blood counts were initially scanned and converted to PostScript files with Photoshop 4.0. GhostView 2.2 was used to extract the time and value of each data point, so a digitized version of the serial blood counts could be used for statistical analysis.

The most commonly used technique for detecting periodicity in a time series is the Fourier power spectrum,
or periodogram. This technique is applicable when the data are evenly sampled but can give erroneous results when the data are unevenly sampled, which is usually the case for serial blood counts. While studying celestial phenomena, astrophysicists also encountered the problem of unevenly sampled data, and they developed an extension of the Fourier power spectrum (the Lomb periodogram) for evenly or unevenly sampled data (Lomb, 1976). The statistical significance (p value) of any peak can also be determined (Scargle, 1982).

Specifically, let \( x_j \) be the number of a particular type of cell as measured at times \( t_j \), where \( j = 1, \ldots, N \) and \( N \) is the number of data points. As usual, the mean and variance of the data values are given by

\[
\bar{x} = \frac{1}{N} \sum_{j=1}^{N} x_j \quad \sigma^2 = \frac{1}{N-1} \sum_{j=1}^{N} (x_j - \bar{x})^2
\]

(1)

Then the Lomb normalized periodogram \( P(T) \) at a period \( T \) is defined by

\[
P(T) = \frac{1}{\sigma^2} \left\{ \frac{\sum_{j=1}^{N} (x_j - \bar{x}) \cos \frac{2\pi(t_j - \tau)}{T}}{\sigma^2 \sum_{j=1}^{N} \cos^2 \frac{2\pi(t_j - \tau)}{T}} + \frac{\sum_{j=1}^{N} (x_j - \bar{x}) \sin \frac{2\pi(t_j - \tau)}{T}}{\sigma^2 \sum_{j=1}^{N} \sin^2 \frac{2\pi(t_j - \tau)}{T}} \right\}
\]

(2)

where the constant \( \tau \) is defined implicitly by

\[
\tan \left( \frac{4\pi\tau}{T} \right) = \frac{\sum_{j=1}^{N} \sin \left( 4\pi t_j / T \right)}{\sum_{j=1}^{N} \cos \left( 4\pi t_j / T \right)}
\]

(3)

The value of \( P(T) \) indicates the likelihood of a periodicity with period \( T \) in the data set. Given that the null hypothesis is that the values \( x_j \) are independent Gaussian random noise, and that \( P(T) \) has an exponential probability distribution with unit mean, the significance level (p value) of any peak is given by

\[
p = 1 - (1 - e^{-p(T)})^{M}
\]

(4)

where \( M = N \) (Press et al, 1992). When the significance level \( p \) is small (significant), Equation 4 can be approximated to give

\[
p = Me^{-p(T)}.
\]

(5)

We implemented Equation 2 for a series of different periods \( T \). The estimation of the significance level of \( P(T) \) is straightforward as long as some criteria are satisfied for the choice of the range and the number of periods that are scanned (Scargle, 1982; Press et al, 1992). A data set was considered periodic if the significance level \( p \) of the principal peak in the periodogram satisfied \( p < 0.05 \) (5%). An adaptation of the procedure proposed in Press et al (1992), using Matlab, was used to analyse the digitized data. Free copies of this program are available from the authors for the analysis of analogous data for non-commercial purposes.

Once a significant periodicity of period \( T \) has been detected through periodogram analysis, the estimation of the phase (\( \phi \)) and amplitude (\( A \)) of the sine wave that fits best the data can be calculated by a least squares fitting procedure (Lomb, 1976; Scargle, 1982) of the data \( X_j \) using

\[
X_j = A \sin \left( \frac{2\pi t_j}{T} + \phi \right), \quad j = 1, \ldots, N
\]

(6)

RESULTS

Periodogram analysis was performed on the serial blood counts of 21 patients. Nine had apparent oscillations in the leucocyte counts, but the analysis indicated that the oscillations did not have a statistically significant period. Three of these patients were reported by Morley et al (1967), three were reported by Kennedy (1970) and one patient each was reported by Inbal et al (1983), Mastrangelo et al (1974) and Mastrangelo et al (1976). Significant periodic oscillations were detected in all the other patients. Table I summarizes the results of the determination of the periods \( T \), their significance levels \( p \), and their phase shifts \( \phi \) relative to the leucocyte oscillation. Figs 1–4 show the serial blood counts and the periodogram analysis for each of these 12 patients.

Periodicities in the leucocyte, platelet, and reticulocyte counts

Fig 1 shows the serial blood counts and corresponding Lomb periodogram of three patients with marked oscillations in the leucocyte population. Typically, leucocyte counts oscillated from normal to high values. Although no data were published for the platelet and reticulocyte counts, Kennedy (1970) and Gatti et al (1973) noted that their patients also showed platelet oscillations with the same period as the leucocytes. Dolobel et al (1973) did not mention if oscillations were present in other cell lineages.

The serial blood counts and the corresponding Lomb periodograms for six patients with both leucocyte and platelet oscillations are shown in Figs 2 and 3. The two cases reported by Vodopick et al (1972) also showed variation in the reticulocyte counts, but with a much smaller amplitude than in the leucocytes and platelets. (The periodicities could not be established due to lack of data.) The leucocyte and platelet counts of these two patients were oscillating, with the same period, between normal and elevated levels. Yamauchi & Ide (1992) reported that the reticulocyte counts of their patient were stable, but the leucocytes oscillated between normal and higher levels, and the platelets oscillated around normal levels with the same period as the leucocytes. The leucocyte counts of the patient reported by Kennedy (1970) oscillated between normal and high levels, but the platelets oscillated from normal to low levels. Cyclic variations were not noted in the reticulocyte counts. Morley et al (1967) reported oscillations between normal and high levels in the leucocyte and platelet counts of their patient, but the authors did not mention if the reticulocyte counts were oscillating. The leucocyte and platelet counts of the patient reported by Umemura et al
Fig 1. Data and analysis for patients in which only leucocyte data were published. The left column contains serial blood counts of three patients with oscillations in the leucocyte population, and the right column contains the corresponding Lomb periodogram $P(T)$ [power $P$ versus period $T$ in days]. The dashed line in the left column represents the normal level of cells, and the horizontal lines in the Lomb periodogram give the $p = 0.05$, $p = 0.01$ and $p = 0.001$ significance levels.
Fig 2. Analysis of published leucocyte and platelet data for three patients. All other notation as in Fig 1.
Fig 3. Analysis of published leucocyte and platelet data for three patients. All other notation as in Fig 1.
Fig 4. Analysis of serial blood counts from three patients with oscillations in the leucocyte, platelet, and reticulocyte populations. All other notation as in Fig 1.
Kennedy (1970) 58

Reference Leucocytes Platelets Reticulocytes

Period (days) and p value

Platelets Reticulocytes

<table>
<thead>
<tr>
<th>Reference</th>
<th>Leucocytes</th>
<th>Platelets</th>
<th>Reticulocytes</th>
<th>Platelets</th>
<th>Reticulocytes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kennedy (1970)</td>
<td>58 ± 1 (1 × 10−5)</td>
<td>-*</td>
<td>-†</td>
<td>-§</td>
<td>-§</td>
</tr>
<tr>
<td>Delobel et al (1973)</td>
<td>73 ± 5 (2 × 10−6)</td>
<td>-‡</td>
<td>-‡</td>
<td>-‡</td>
<td>-‡</td>
</tr>
<tr>
<td>Gatti et al (1973)</td>
<td>71 ± 2 (5 × 10−10)</td>
<td>-*</td>
<td>-†</td>
<td>-§</td>
<td>-§</td>
</tr>
<tr>
<td>Vodopick et al (1972)</td>
<td>70 ± 4 (7 × 10−9)</td>
<td>64 ± 4 (3 × 10−3)</td>
<td>-*</td>
<td>9:1 ± 2:5</td>
<td>-§</td>
</tr>
<tr>
<td>Vodopick et al (1972)</td>
<td>54 ± 3 (1 × 10−9)</td>
<td>54 ± 3 (2 × 10−3)</td>
<td>-*</td>
<td>10:2 ± 1:4</td>
<td>-§</td>
</tr>
<tr>
<td>Yamazaki &amp; Ide (1992)</td>
<td>70 ± 2 (1 × 10−1)</td>
<td>70 ± 2 (3 × 10−3)</td>
<td>- †</td>
<td>10:7 ± 2:7</td>
<td>-§</td>
</tr>
<tr>
<td>Kennedy (1970)</td>
<td>43 ± 1 (2 × 10−5)</td>
<td>42 ± 1 (3 × 10−6)</td>
<td>- †</td>
<td>5:9 ± 2:4</td>
<td>-§</td>
</tr>
<tr>
<td>Morley et al (1967)</td>
<td>54 ± 14 (3 × 10−1)</td>
<td>53 ± 14 (1 × 10−1)</td>
<td>-‡</td>
<td>9:5 ± 1:6</td>
<td>-§</td>
</tr>
<tr>
<td>Unemura et al (1986)</td>
<td>83 ± 12 (2 × 10−1)</td>
<td>83 ± 11 (2 × 10−4)</td>
<td>- †</td>
<td>13:2 ± 2:2</td>
<td>-§</td>
</tr>
<tr>
<td>Iizuka et al (1984)</td>
<td>57 ± 2 (2 × 10−6)</td>
<td>55 ± 2 (5 × 10−4)</td>
<td>55 ± 2 (5 × 10−3)</td>
<td>5:9 ± 2:1</td>
<td>24:3 ± 2:4</td>
</tr>
<tr>
<td>Chikkappa et al (1976)</td>
<td>63 ± 3 (2 × 10−2)</td>
<td>59 ± 1 (9 × 10−1)</td>
<td>62 ± 3 (7 × 10−1)</td>
<td>9:7 ± 3:5</td>
<td>31:9 ± 5:6</td>
</tr>
<tr>
<td>Rodriguez &amp; Latcher (1976)</td>
<td>37 ± 3 (2 × 10−2)</td>
<td>36 ± 3 (2 × 10−1)</td>
<td>37 ± 3 (3 × 10−1)</td>
<td>8:6 ± 2:7</td>
<td>16:7 ± 1:9</td>
</tr>
</tbody>
</table>

* Oscillating, but data not published. † Not apparently oscillating. ‡ Data not published and no comments by author concerning periodicity.
§ Since data were not available for platelets and/or reticulocytes, the phase shift could not be calculated.

---

**Discussion**

Our analysis shows that PCML is a periodic blood disorder with some characteristics similar to those observed in cyclical neutropenia (CN) (Haurie et al. 1998a). In both diseases, oscillations in one or more cell populations are present, and usually the populations oscillate with the same period. A phase shift between the oscillating populations has also been noticed in CN, but it was not as consistent as that observed in PCML (Haurie et al. 1998a). It is unclear how the oscillations of CN arise, but there is strong circumstantial evidence that an instability in the HSC is responsible for the induction of these cycles (Haurie et al. 1998a; Jones et al. 1975a; b; Weiden et al. 1974; Krance et al. 1982; Puit et al. 1973) and that this instability is due to an elevated rate of HSC apoptosis (Haurie et al. 1998b; Mackey, 1978, 1996). Further, a connection has been established between cyclical neutropenia and idiopathic and congenital neutropenia (in which there is an elevated rate of apoptosis) (Haurie et al. 1998b).

The observations of this paper strongly suggest that the origin of the oscillations in PCML, as in CN, is located in the HSC such that there is cyclical differentiation efflux to the more mature cells. Given the derangements of apoptosis implicated in CML (Amarante-Mendes et al. 1998; Clarkson et al. 1997; Gisslinger et al. 1997; Maguer-Satta et al. 1998; Sattler & Salgia, 1997; Seong et al. 1998; Wright et al. 1998), we speculate that in PCML there is a HSC instability induced by decreased apoptotic rates such that an oscillatory efflux from the HSC gives rise to the oscillations in the peripheral blood cell counts analyzed here.

Further evidence for a stem cell origin comes from theoretical investigations of the stem cell dynamics. A mathematical model of the stem cell population, which takes into account the presence of a delay due to a finite proliferation time, was shown to produce oscillations with periods in the range reported here (Mackey, 1978, 1979, 1996).

It is completely unclear as to what extent PCML may occur...
in the entire population of CML patients. This lack of knowledge concerning the frequency with which PCML occurs is largely related to the paucity of the serial clinical data usually available on CML patients. Though it is difficult to obtain an analytic rule of thumb connecting the number \( N \) of data points required to establish the existence of a cycle of period \( T \) with significance \( p \) using the definition of the Lomb periodogram, it can easily be used to do this numerically. We generated a sine wave of amplitude \( A \) and period \( T \) with added Gaussian distributed white noise with variance \( \sigma^2 \), and determined the relation between the number \( N \) and the significance level \( p \) with which \( T \) could be determined for a given ratio of the amplitude of the sine wave to the standard deviation \( \sigma \). This ratio is called the signal-to-noise ratio (SNR) and is usually expressed in dB and in terms of root-mean-square values. In the present case, the SNR can be written as

\[
\text{SNR} = 20 \log_{10} \left( \frac{A}{\sqrt{2} \sigma} \right). \tag{7}
\]

Fig 5 shows the significance level \( p \) with which the period \( T \) of these noisy test signals can be determined as a function of the number of data points \( N \) and three SNRs. (The lower line is that of a pure sine wave without added noise and fixes the lower bound on the significance level.) The SNRs used were 1, 5 and 10 dB (corresponding to ratios \( A/\sigma \) of about 1·6, 2·5 and 4·5). For a pure sine wave (no noise), a minimum of 13 data points is required to detect a periodic oscillation with \( p \leq 0·05 \). Similarly, a minimum of 16 data points is required for a typical 5 dB signal. If one wishes to establish a cycle with a period of 2 months and with a significance level of 0·5, then the following criteria should be satisfied. Blood samples should be obtained at least once a month (Nyquist frequency) for a minimum of 16 months, twice a month for a minimum of 8 months, or once a week for 4 months.

The data of Fig 5 indicate that the detection PCML would require a sampling frequency that is higher and more regular than is currently standard practice. The detection of these oscillations in CML patients may help to take a more informed decision concerning the choice of treatment.
ACKNOWLEDGMENTS

We thank Professor B. J. Kennedy for kindly discussing some of his published data and sharing unpublished data with us. Our colleagues Mlle Caroline Haurie and Professor David C. Dale made valuable suggestions at various phases of this research for which we are grateful. This work was supported by the Natural Sciences and Engineering Research Council (NSERC grant OGP-0036920, Canada), and Le Fonds pour la Formation de Chercheurs et l’Aide à la Recherche (FCAR grant 98ER1057, Québec).

REFERENCES


