A Markov approach to characterizing the PK-PD relationship of anti-migraine drugs

Maas, H.J.

7

Prediction of attack frequency in migraine: a Markov approach

HJ Maas, N Snelder, M Danhof, OE Della Pasqua
Submitted to Cephalalgia

Clinical studies of acute anti-migraine drugs have been used as a source of information for characterising migraine attacks in terms of duration and intensity. The design of some of these studies is such that they also encompass the period between attacks. Using this type of information it is shown that, for a clinical population of migraineurs, alternating attack and interictal periods can be described as a single stochastic process.

An analysis was performed on clinical data derived from patients who had 2 or 3 subsequent migraine attacks that were all treated with single doses of oral placebo ($n = 73$ patients), naratriptan 2.5 mg ($n = 143$) or sumatriptan 100 mg ($n = 154$). It was proved that the distribution of interictal durations in each of these three populations can be described by the Gamma distribution. Likewise, the exponential distribution accurately describes the duration of an attack. Based on these findings, a simple stochastic process was developed and evaluated consisting of two exponential distributions to describe interictal durations and another one to describe attack durations.

According to this process, the mean duration of the interictal period is 26 (18–40), 20 (14–25) and 19 (13–24) days (mean, 95% confidence interval) for the placebo, naratriptan 2.5 mg and sumatriptan 100 mg groups, respectively. The estimated mean durations of a migraine attack are 0.86 (0.67–1.1), 0.54 (0.46–0.62) and 0.54 (0.46–0.62) days for the

three groups. Compared with the distribution analysis, the stochastic process predicted the duration between attacks less accurately.

Modelling of paroxysmal diseases such as migraine using distributions and stochastic processes provides valuable insight into the dynamics of the disease. The results of such analyses are directly applicable to questions in healthcare and the design of new clinical studies.

7.1 Introduction

Although migraine manifests itself in the form of attacks, pathophysiological events are continuously unfolding even in the period between attacks [1, 2, 3]. Electrophysiology studies show abnormal brain activity in migraine patients during this period [4]. Raised sensitivity to pain stimuli has been demonstrated [5]. Finally, premonitory symptoms precede the actual attack by several hours [2, 3]. A few hypotheses have been advanced to explain this alternating pattern of attacks and periods of neuronal instability. Among them is the energy imbalance model suggesting hypoxia as the cause of attacks [6, 7]. Another compares the condition with epilepsy, where a disturbed equilibrium between excitatory and inhibitory cortical activity is at the basis of the migraine attack [8]. Unravelling the dynamics of the different events has proved a significant challenge in understanding migraine.

Recently, an interest in time-series analysis has arisen with the intention to express the dynamics of headaches in a quantitative manner [9, 10]. This type of analysis has since long been common practice in the field of epilepsy. Among the first questions addressed using this methodology was the matter of independence of subsequent seizures [11, 12]. When seizures tend to cluster in time or conversely occur with a high degree of periodicity, this may indicate regulation by physiological control systems, such as feedback loops. A random distribution of times between seizures however would point at the involvement of chaotic systems or the influence of multiple external factors. Though the results of these analyses varies between studies, the majority of seizures patterns was found to be random. Knowledge thus obtained has contributed to the understanding of epilepsy and paroxysmal diseases in general.

A necessity for time series analysis to be performed is the availability of suitable data. In particular, it requires observations on large series of events (attacks) within the same patient. In epilepsy there is a relative abundance of detailed data sets accounting the status of individual epilepsy patients on a day-by-day basis. This is partly due to the disease dynamics. Short durations of events and short intervals between events facilitate the collection of detailed data. Furthermore, the urgent need for prophylactic anti-epilepsy medication warranted longitudinal studies to be performed.

Perhaps the largest data source on migraine is constituted by the randomised clinical trials of acute anti-migraine medication. It has been exploited scientifically to study the course of migraine attacks in the presence and absence of drug treatment [13, 14, 15]. Long-term studies investigating patients' efficacy and tolerability to these drugs are also available [16, 17], yet these have not been utilised for the study of migraine disease

dynamics. This study explores the possibilities of applying clinical trial data on triptans to advance the understanding of the dynamics of migraine. An analysis of the alternating pattern of migraine attacks and interictal periods can also be of use in providing answers to practical problems arising in healthcare [18]. Examples include the calculation of the expected number of migraine attacks in a population within a certain period or the calculation of the total number of headache days.

The current analysis starts by showing that, within a limited time window, the time between migraine attacks can be described by the two parameters of the Gamma distribution, whereas the duration of an attack can be described using the exponential distribution. Using this result, it is demonstrated that the alternation of attacks and attack-free periods can be thought of as a chain of transitions, forming a simple stochastic process. Finally, the effect of acute anti-migraine therapy on migraine dynamics is investigated and the usefulness of inference with stochastic processes is demonstrated.

7.2 Methods

7.2.1 Data

Data analysis was performed on a sample selected from a total of 1288 patients participating in a clinical study investigating the efficacy of naratriptan and sumatriptan over multiple attacks. 394 patients included in the analysis received single doses of placebo, naratriptan 2.5 mg or sumatriptan 100 mg in each attack. Furthermore, only patients were selected for which two or three subsequent migraine attacks had been recorded. Patient characteristics are summarised in Table 7.1.

Headache was measured on a 4-point scale with scores 0, 1, 2 and 3 representing no pain, mild pain, moderate pain and severe pain. Measurement started once a patient's first migraine attack reached maximum pain intensity. The observation times were 0, 0.5, 1.0, 1.5, 2.0, 2.5, 3.0, 3.5, 4.0, 6.0, 8.0, 12.0, 24.0, 42.0 and 48.0 h. Second and third attacks were recorded similarly. To accommodate longitudinal analysis on these headache recordings, observations were reformatted in 3 steps. First, observation times of subsequent migraine attacks were expressed relative to the first measurement of the first attack. In the next step, the time unit was converted to days. Lastly, time intervals between observations were discretised by 1) preserving only observations at integer time values and 2) imputing '0' scores for days on which headache was not measured (the periods between two recorded attacks).

7.2.2 Distribution analysis

The durations of inter-attack periods were extracted from the reformatted data set by determining the lengths of all '0'-score sequences (*i.e.* headache free periods) in the data. This set of durations was then modelled using the Gamma distribution. This distribution is typically used to characterise processes involving waiting times. It is bounded at the lower end by zero and is unbounded at the upper end. Its appearance is determined by

two parameters, a scale parameter reflecting the degree of shallowness of the distribution curve and a shape parameter reflecting the position of the top.

The durations of attacks were found by determining the lengths of all non-zero score sequences in the reformatted data. These durations were then fit to the exponential distribution. This distribution is characterised by only a scale parameter. The position of the maximum is at time zero, reflecting the observation that most patients experience resolution of headache within one day after headache has reached maximum intensity.

The parameters of the Gamma and exponential distributions were estimated separately for the naratriptan, sumatriptan and placebo data sets using maximum likelihood estimation in the statistical software package S-Plus.

7.2.3 Stochastic process

Using the results of the distribution analysis, a simple stochastic process was constructed in the form of a Markov chain (Figure 7.1). It describes the process of migraine as a cycle of transitions leading up to and away from a migraine attack. By definition, the duration of an individual transition in a Markov chain is given by an exponential distribution. As in the distribution analysis, a single exponential distribution accounts for the duration of attacks. In Figure 7.1, this is represented by the transition from state A (Attack) to $NA1$ (No Attack 1). The duration of the period between attacks is modelled as the sum of two identical transitions, to wit from state $NA1$ to $NA2$ (No Attack 2) and from state $NA2$ to A. This part of the model conveniently uses the statistical property that summing over a number of Markov transitions amounts to using a Gamma distribution with the shape parameter equal to the number of the transitions.

Therefore, as in the distribution analysis, a Gamma distribution accounts for the duration of the period between attacks. The only difference is that, in the stochastic process, the shape parameter is restricted to positive integer values whereas the actual Gamma distribution allows the shape to be any positive real number. In total, the model contains only two scale parameters: one to characterise the duration of attacks and one to characterise the duration between attacks. These parameters were estimated separately for the sumatriptan, naratriptan and placebo data sets.

In contrast to static distributions, stochastic processes evolve over time. Thus, instead of extracting the attack durations (non-zero headache scores) and inter-attack durations (zero headache scores) prior to analysis, the reformatted time series could be analysed directly. However, it had to be verified whether the model could automatically and correctly allocate non-zero headache scores to state A and zero scores to states $NA1$ and $NA2$. This was achieved by applying a Hidden Markov model (HMM) [15], which, for every state in the model (A, $NA1$, $NA2$), determines the headache scores (0, 1, 2, 3) associated with it. The model was implemented in open-source software [19] which was operated from within S-Plus 6.2.1 on a Linux workstation (Fedora Core 3).

7.2.4 Goodness-of-fit statistics

The results of the distribution analysis were assessed graphically. The accuracy of the Gamma distribution fits was determined by inspection of quantile plots, showing the correspondence between the sample distribution and the distribution hypothesised by the model. The accuracy of the exponential distribution fits was analysed by comparing the predicted distributions with the histograms of the data. For each bar in the histogram, the area under the distribution curve in that interval should match that of the bar. This is because the total area under the histogram and under each curve is scaled to 1.

The estimates from the stochastic process were expressed in terms of their distributions: exponential distributions for the attack durations and Gamma distributions with integer-value shape parameter for the durations between attacks. These distributions were also compared by means of quantile plots and histograms.

Parameters and their standard errors were estimated. Estimates are reported including 95% confidence intervals. These were derived from the standard errors using log-normal approximations for scale parameters and normal approximations for shape parameters [20]. Significant differences between parameters are indicated by non-overlapping

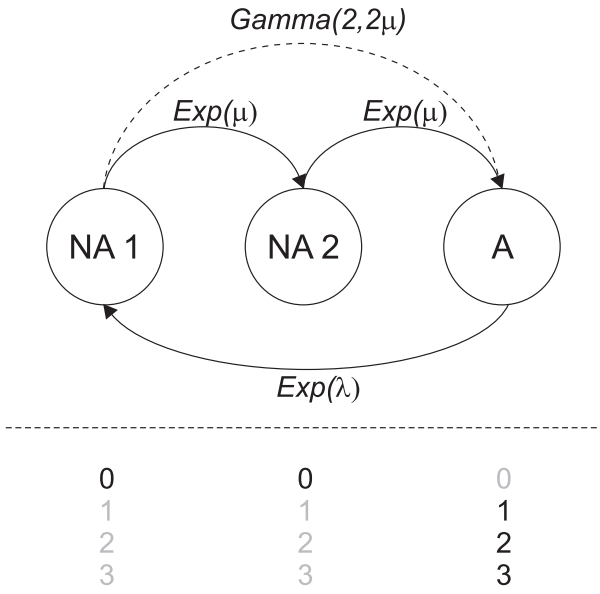

Figure 7.1: *Structure of the proposed stochastic process describing migraine in a clinical setting. Patients start in state A (attack) which is characterised by headache scores 1, 2 and 3. Their pain resolves (state $NA1$, headache score 0) according to rate $Exp(\lambda)$ which is exponentially distributed with scale parameter λ. A process leading up to a new attack via intermediary state $NA2$ is then initiated. This process is described by two identical exponentially distributed rates $Exp(\mu)$ with scale parameter μ . These rates correspond to a single Gamma distribution with shape 2 and scale 2 μ (as it covers 2 transitions).*

Table 7.1: *Clinical characteristics of the patient groups used in the data analysis.*

		placebo	naratriptan 2.5 mg	sumatriptan 100 mg
Number of patients	male	9	32	24
	female	64	111	154
Age (y)	mean	39	40	40
	range	21 – 60	18 – 64	18 – 63
Percentage of attacks	duration = 1 day	73	89	89
	duration $>$ 1 day	27	11	11
Duration between	mean	24	22	20
attacks (days)	range	2 – 78	1 – 133	1 – 88

95% confidence intervals.

To demonstrate the model's practical value, an example is given of inference on the stochastic process. Using a numerical procedure [21] in S-Plus, the stationary or long-term proportion of patients experiencing an attack was calculated based on the parameter estimates of the placebo, naratriptan 2.5 mg and sumatriptan 100 mg models. Mean and confidence intervals for this prediction were derived from the standard errors.

7.3 Results

It was first verified whether the HMM program running the stochastic process had correctly allocated no pain scores to states $NA1$ and $NA2$ and pain scores 1, 2 and 3 to state A. This was case for all analyses. This observation provides further evidence that exponential transitions adequately fit the time-series data.

Table 7.2 shows the distributions of the between-attack durations as estimated by the Gamma distribution analysis and the stochastic process. The estimates of the scale parameter in the distribution analysis are similar across treatments. The shape parameters for naratriptan and sumatriptan are somewhat smaller than that for placebo (1.7 and 1.8 *vs* 2.0, respectively). The overlap in the confidence intervals indicates that this difference does not reach significance. As a result, the means of the distributions, which is defined as the product of the scale and the shape, does not differ between the treatments either. Though not significant, the estimates in the stochastic process are clearly different from those in the distribution analysis. In particular, the naratriptan and sumatriptan scale parameters are smaller than their corresponding values in the distribution analysis (9.9 and 9.7 *vs* 12 and 13) Due to the model's structure (Figure 7.1), all shape parameters are confined to 2. The differences between the two models are amplified in the means. The actual mean values calculated from the data (Table 7.1) lie approximately between those estimated for the two methods.

To determine the accuracy of both models with respect to all observed data, quantile plots were constructed (Figure 7.2). Although the agreement is reasonable in both analyses, it immediately becomes evident from these plots that the distribution analysis predicts the data better. Particularly for the longer durations ($>$ 30 days), the stochas-

Table 7.2: *Estimates of distribution parameters describing the durations between attacks following the three treatments. The parameters of the stochastic process are expressed as their corresponding Gamma-based parameters to facilitate comparison with the Gamma distribution fit. The means of the estimated distributions are given by the equation:* $mean = scale * shape$*. The values on the lower lines represent the 95% confidence intervals of the estimates. The shape of the stochastic process was constrained at value 2 as the model contains 2 transitions leading up to an attack (see Figure 7.1).*

Model:	Gamma distribution			Stochastic process		
Parameter:	scale (days)	shape	mean (days)	scale (days)	shape	mean (days)
Placebo	12	2.0	24	13	2	26
	9.2 – 16	1.5 – 2.5	17 – 34	8.9 – 20		18 – 40
Naratriptan	12	1.8	23	9.9	2	20
2.5 mg	10 – 15	1.5 – 2.1	18 – 29	6.8 – 13		14 – 25
Sumatriptan	13	1.7	22	9.7	2	19
100 mg	11 – 16	1.4 – 2.0	17 – 28	6.6 – 12		13 – 24

tic process seems to underestimate the drug treatment observations and overestimate the placebo data.

Figure 7.3 illustrates the results of modelling the between-attack durations in terms of distribution plots. From this figure it also appears that the distribution analysis results agree best with the data. In contrast with the quantile plots, the difference is clearest in the relatively short durations (< 20 days).

Table 7.3 gives the modelling results for the durations of the attacks. The attack durations are significantly shorter in the presence of drug, irrespective of the type of analysis and drug. Consistent with the duration between attacks, the stochastic model predicts a higher value for the placebo scale parameter and lower values for the drug scale parameters, compared with the distribution analysis (0.86, 0.54 and 0.54 days *vs* 0.77, 0.61 and 0.61 days, respectively).

As the observed attack durations are either one or two days, quantile plots cannot be constructed from these data. Instead, the accuracy of the predictions with respect to the data is assessed using the areas under the distribution plots (Figure 7.4). From day 0 to 1 in the placebo plot, the areas are 0.73, 0.73 and 0.69, for the observed data, distribution analysis and stochastic process, respectively. In the naratriptan plot, these areas respectively are 0.89, 0.81 and 0.84. The areas in the sumatriptan plot are identical to those in the naratriptan plot. The areas between day 1 and 2 need not be calculated since they are the complements of the areas between day 0 and 1. Thus, the stochastic process performs slightly better than the distribution analysis in predicting the attack durations after sumatriptan and naratriptan. It performs poorer in predicting the duration after placebo.

Lastly, an example is given of how the dynamics of the stochastic process can be used for statistical inference. This is done using the property that on the long term the proportions of patients residing in each of the states of this model become stationary.

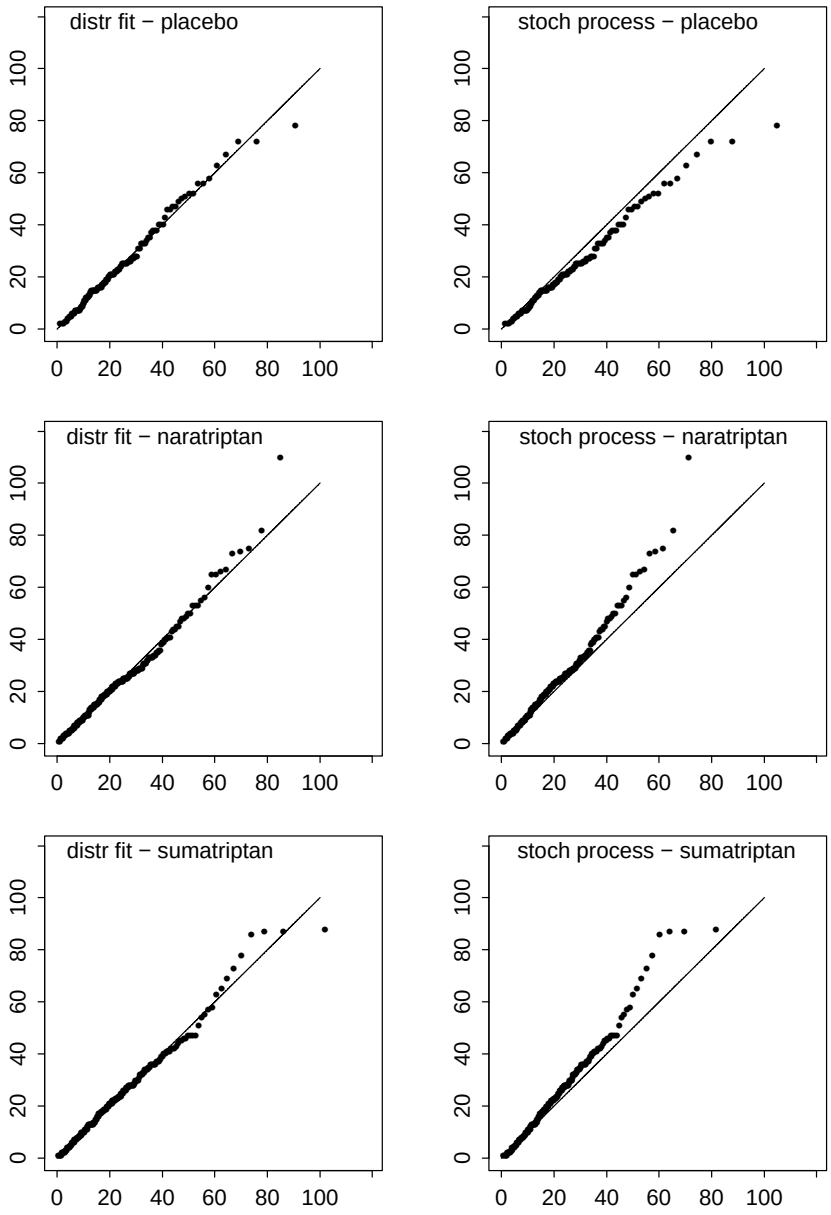

Figure 7.2: *Quantile plots showing the accuracy of the predicted durations between attacks. The distribution hypothesised by the model is plotted on the horizontal axis, the distribution of the data sample is plotted on the vertical axis. The diagonal represents equality between the hypothesised and the sample distribution. Left: results of the distribution fits for the three treatments. Right: results of the stochastic (Markov) process for the three treatments.*

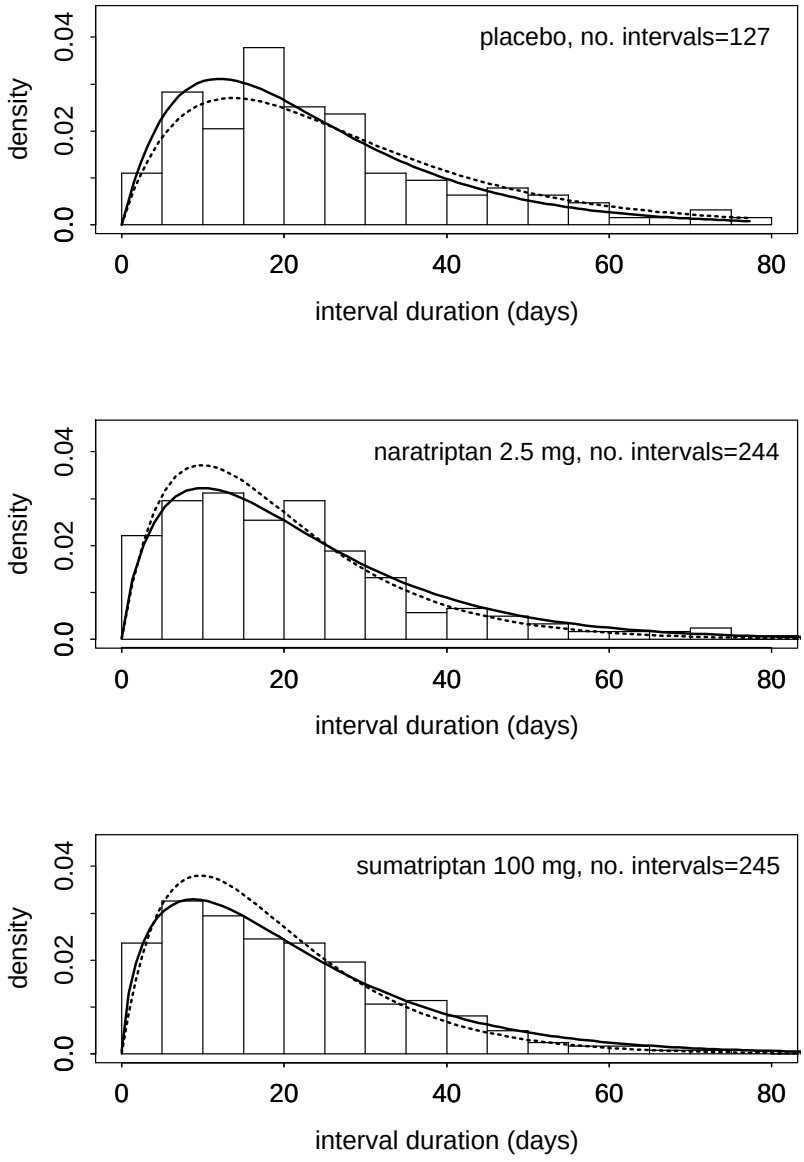

Figure 7.3: *Predicted and observed distributions of duration between attacks. The results of the distribution analysis (solid line) and the stochastic process (dotted line) are superimposed on the histograms of the observed between-attack durations. The upper, middle and lower panels show the analyses for placebo, naratriptan 2.5 mg and sumatriptan 100 mg, respectively. For comparison, the distributions are scaled so that the area under every curve and under the histogram adds up to 1.*

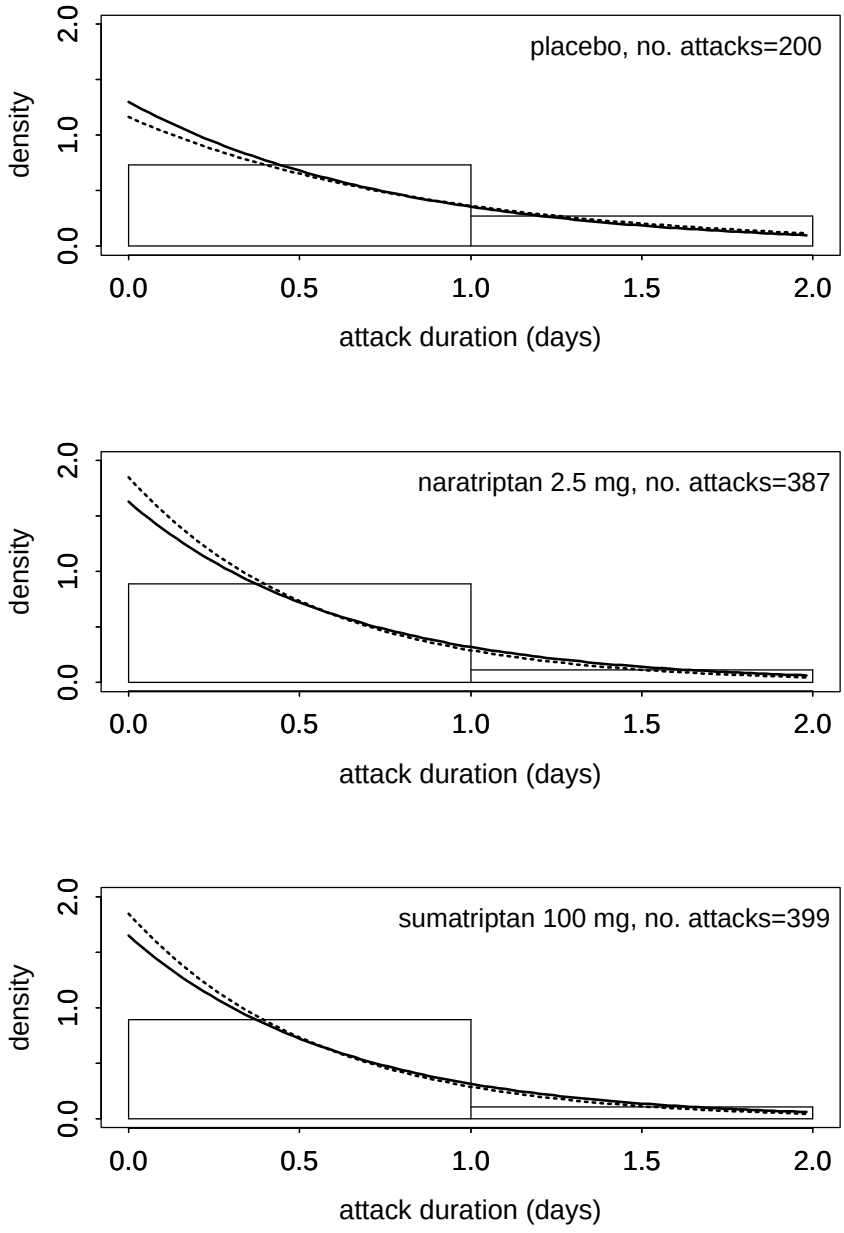

Figure 7.4: *Prediction of the distributions of attack durations. The results of the distribution analysis (solid line) and the stochastic process modelling (dotted line) are superimposed on the histograms of the observed attack durations. The upper, middle and lower panels show the analyses for placebo, naratriptan 2.5 mg and sumatriptan 100 mg, respectively. The area under every curve and under the histogram adds up to 1.*

Table 7.3: *Estimates of distribution parameters describing the durations of attacks following the three treatments. By definition, the means of the estimated distributions are equal to the scale parameters. The values on the lower lines represent the 95% confidence intervals of the estimates.*

Model:	Exponential distribution	Stochastic process
Parameter:	scale (days)	scale (days)
Placebo	0.77	0.86
	0.69 – 0.92	0.67 – 1.1
Naratriptan 2.5 mg	0.61	0.54
	0.55 – 0.68	0.46 – 0.62
Sumatriptan 100 mg	0.61	0.54
	0.55 – 0.67	0.46 – 0.62

Therefore, given the dynamics of alternating attack and no-attack periods in the stochastic process, the proportion of patients suffering migraine pain at any time in the future can be derived. The percentages of the clinical population having migraine pain at any given moment were found to be 5.5 (3.8–7.8)% for placebo, 4.9 (3.8–6.3)% for naratriptan 2.5 mg and 4.9 (3.9–6.3)% for sumatriptan 100 mg.

7.4 Discussion

A cyclical stochastic process was developed to characterise the dynamics of migraine. Although simple, the model captures the alternating pattern of periods of attack and no-attack using Gamma and exponential distributions. The choice of these distributions was confirmed by a distribution analysis performed on the attack durations and between-attack intervals. Inspection of different graphical goodness-of-fit criteria learned that the distribution analysis predicted the durations between attacks more accurately than the stochastic process. Most likely, the variable shape parameter enabled the Gamma distribution to better characterise the shape of the distributions of durations. As can be expected, the estimated attack durations were significantly shorter in the drug treatment groups, relative to the placebo group. No significant differences were found for the between-attack durations, indicated by the scale parameters, but there was a trend towards a smaller duration with drug treatment. Although this trend was not found in the distribution analysis, here a downward trend in the shape parameter, defining the position of the distribution, was observed in the presence of drug treatment. This difference in trends is likely due to the fact that the shape in the stochastic process was confined to the integer 2. Any changes in the distribution curve could therefore only be effectuated by changing the scale parameter.

The model parameters were estimated based on data from a randomised clinical trial

of the acute anti-migraine drugs sumatriptan and naratriptan. For each patient in this trial, up to three subsequent migraine attacks were recorded. This number of events is too small to determine migraine dynamics on an individual basis. For an analysis of a population of migraineurs however, the 394 patients included in the analysis provide sufficient information to allow parameter estimation. Durations and frequencies of migraine attacks are commonly reported in demographic studies using mean values and ranges. Yet, no attention has been given to the distribution of these measures. A statistical distribution captures the complete information of a measure using only a few parameters. Recently, the individual distributions of migraine and tension-type headaches have been described by Houle *et al*. [10]. To clearly distinguish between these types of headache, attacks and headache-free episodes were considered parts of a single distribution, which was found to be closer to normal when the headache condition was more chronic and more bimodal (two-peaked) otherwise. In this paper the emphasis is on the timing of migraine attacks. To this purpose attacks were treated as separate events characterised by distributions which, in contrast to bimodal distributions, are quantitatively well-defined.

In formulating the model, the terms headache-free and attack-free have been used interchangeably. This assumption implies that the acute anti-migraine drugs act by terminating the attack, rather than only suppressing the pain caused by it. This is visualised in Figure 7.1, where "no headache" scores are only associated with attack-free states $NA1$ and $NA2$.

There is no obvious interpretation for intermediary state $NA2$. It was required to characterise the period between attacks which spans over two identically distributed time intervals (A–$NA1$ and $NA1$–$NA2$). It is however not inconceivable that the time between two attacks can be subdivided into periods according to their susceptibility to precipitating factors or occurrence of premonitory symptoms. These periods are not likely to have equal durations.

Statistical distributions are important tools for making inference on data samples. They summarise many properties of data sets using a few parameters. However, they cannot describe by themselves more complex dynamics such as alternating sequences of events. In order to describe these, a number of distributions needs to be linked by stochastic processes. Once a stochastic process has been defined, a system's dynamics can be investigated. With regard to this type of analysis, the property of interest is usually the stationary distribution of the process, reflecting its long-term behaviour. In epilepsy, simple Markov chains have been applied to test for the existence of stationary behaviour in a patient's seizures [12]. A Poisson process has been used to model the number of cardiac arrests that can be expected in a hospital population on a given time interval. Although migraine and headaches in general are characterised by more complicated time distributions, developing statistical descriptors for these conditions will proof useful in a practical sense and will help to understand them.

References

[1] J. Afra, A. Mascia, P. Gerard, A. Maertens de Noordhout, and J. Schoenen. Interictal cortical excitability in migraine: a study using transcranial magnetic stimulation of motor and visual cortices. 1998, *Ann.Neurol.*, 44, 209–215.

[2] J.N. Blau. Classical migraine: symptoms between visual aura and headache onset. 1992, *Lancet*, 340, 355–356.

[3] E.L. Spierings, M. Sorbi M, G.H. Maassen, and P.C. Honkoop. Psychophysical precedents of migraine in relation to the time of onset of the headache: the migraine time line. 1997, *Headache*, 37, 217–220.

[4] A. Ambrosini and J. Schoenen. The electrophysiology of migraine. 2003, *Curr.Opin.Neurol.*, 16, 327–331.

[5] R. Burstein, D. Yarnitsky, I. Goor-Aryeh, B.J. Ransil, and Z.H. Bajwa. An association between migraine and cutaneous allodynia. 2000, *Ann.Neurol.*, 47, 614–624.

[6] W.K. Amery. Brain hypoxia: the turning-point in the genesis of the migraine attack? 1982, *Cephalalgia*, 2, 83–109.

[7] W.K. Amery. Migraine and cerebral hypoxia: a hypothesis with pharmacotherapeutic implications. 1985, *Cephalalgia*, 5, S131–S133.

[8] A.E. Eggers. New neural theory of migraine. 2001, *Med.Hypotheses*, 56, 360–363.

[9] A.W. Fox. Time-series data and the “migraine generator”. 2005, *Headache*, 45, 920–925.

[10] T.T. Houle, D.B. Penzien, and J.C. Rains. Time-series features of headache: individual distributions, patterns, and predictability of pain. 2005, *Headache*, 45, 445–458.

[11] J.G. Milton, J. Gotman, G.M. Remillard, and F. Andermann. Timing of seizure recurrence in adult epileptic patients: a statistical analysis. 2006, *Stat.Med*, 25, 2398–2426.

[12] E. Tauboll, A. Lundervold, and L. Gjerstad. Temporal distribution of seizures in epilepsy. 1991, *Epilepsy Res.*, 8, 153–165.

[13] C. Allen, K. Jiang, W. Malbecq, and P.J. Goadsby. Time–to–event analysis, or who gets better sooner? An emerging concept in headache study methodology. 1999, *Cephalalgia*, 19, 552–556.

[14] H. Hassani and A. Ebbutt. Use of a stochastic model for repeated binary assessment. 1996, *Stat.Med.*, 15, 2617–2627.

[15] H.J. Maas, M. Danhof, and O.E. Della Pasqua. Prediction of headache response after migraine treatment using a Markov model. 2006, *Cephalalgia*, 26, 416–422.

[16] M.A. Bomhof, J. Heywood, A. Pradalier, H. Enahoro, P. Winter, and H. Hassani. Tolerability and efficacy of naratriptan tablets with long-term treatment (6 months). Naratriptan Long-term Study Group. 1998, *Cephalalgia*, 18, 33–37.

[17] J. Heywood, M.A. Bomhof, A. Pradalier, L. Thaventhiran, P. Winter, and H. Hassani. Tolerability and efficacy of naratriptan tablets in the acute treatment of migraine attacks for 1 year. Naratriptan Long-Term Study Group. 2000, *Cephalalgia*, 20, 470–474.

[18] E. Skogvoll and B.H. Lindqvist. Modeling the occurrence of cardiac arrest as a poisson process. 1999, *Ann.Emerg.Med.*, 33, 409–417.

[19] A. Bureau, J.P. Hughes, and S.C. Shiboski. An S-Plus implementation of Hidden Markov Models in Continuous Time. 2000, *J.Comp.Graph.Stat.*, 9, 621–632.

[20] W.Q. Meeker and L.A. Escobar. *Statistical Methods for Reliability*. John Wiley, 1998.

[21] V. Anisimov, H.J. Maas, M. Danhof, and O.E. Della Pasqua. Analysis of responses in migraine modelling using hidden Markov models. 2007, *Stat.Med.*, published online.