Stochastic Study on the Sales Pattern of Influenza Anti-Viral Agents at Pharmacies

Mizuho Takahashi,[a] Kazushige Ijuin,[b] Kazuo Iwaki,[c] Rieko Matsuda,[d] Yuzuru Hayashi,*,[d] and Takehiko Yajima[e]

[a]Nikko Citizen's Hospital, 1752–10, Arasawa, Kiyotaki, Nikko, Tochigi 321–1441, Japan, [b]Tanashi Yakuhin Co,. Ltd., 4–25–5, Tanashi, Nishi-Tokyo, Tokyo 188–0011, Japan, [c]School of Pharmaceutical Sciences, Ohu University, 31–1 Misumido, Tomitamachi Koriyama , Fukushima 963–8611, Japan, [d]National Institute of Health Sciences, 1–18–1 Kami-Yoga, Setagaya, Tokyo 158–8501, Japan, and [e]Faculty of Pharmaceutical Sciences, Toho University, Miyama 2–2–1, Funabashi, Chiba 274–8510, Japan

(Received February 17, 2006; Accepted April 19, 2006; Published online April 20, 2006)

This paper focuses on the stochastic analysis of the daily variations in formulations dispensed at pharmacies. A recent study showed that the cross-correlation functions between the time variations can be used for estimating the route and speed of influenza propagation process in society. In reality, however, the peak-shaped signals of the time variations and cross-correlation function are so complicated and noisy that it is almost impossible to know the original shape of the signals. The aim of this paper is to give an answer to this problem by considering the mathematical relationship between the time variations and cross-correlation function. Influenza anti-viral agents for adults and children are taken as an example.

Key words —— correlation, pharmacy, influenza, spectral analysis

INTRODUCTION

The spatial propagation of disease in human society has been studied in the fields of medical and pharmaceutical sciences.[1–13] The focus in the dimension of the propagation is on time and space with much regard for humans, culture, society, environment, geography, *etc*. The stochastic properties such as auto-correlation, power spectral density, 1/f fluctuation and chaos have been extracted from the time series of disease processes.[1–8] For the spatial understanding of the processes, choropleth maps and geographic information systems have been developed in epidemiology.[9–13]

The information necessary for the above study is basically the number of patients diagnosed by doctors and can be obtained through the nation-wide surveillance, notification in communities, *etc*.[1–5,9–12] Recently, a new information source was proposed for the same purposes[9,10]: the daily variations in formulations dispensed at pharmacies and drug stores. These data were analyzed with a method of spectral analysis (cross-correlation function).[9,10]

The present paper pays attention to the practical use of the cross-correlation functions. Their typical application is concerned with meteorological problems.[14] For example, the precipitation in a mountain area can be related with the water level at the lower reaches of a stream. The cross-correlation function between the time series of the precipitation and water level has its maximum when the correlation of the phenomena is the strongest. Because of the definite causality between the precipitation and water level, the position (day) of the maximum correlation implies the time of the water flow from the mountain to the lower reach.

Similarly, the route and speed of influenza infection in Tokyo and its vicinity were estimated, though tentatively, from the time lags of the cross-correlation functions of the time series of an influenza anti-viral agent dispensed at distant pharmacies.[7] Two formulations of the influenza agents are on the market: Tamiflu® capsule for adults and Tamiflu® dry syrup for children. The above analysis was performed for the same formulation but among different pharmacies. If the time series of the different formulations at a pharmacy are analyzed, the following question can be handled: which are infected earlier with influenza, adults or children.[8]

The upper figure of Fig. 1 shows the time series of the sales of the influenza agents for adults (...) and children (___) at a community pharmacy and the lower figure their cross-correlation function. The time series start on November 1st in 2003 and the pharmacy is located in Kanagawa prefecture. The X-axis (τ) of the cross-correlation function denotes the time by which one of the time series is sifted to calculate the correlation coefficient (Y-axis). The

*To whom correspondence should be addressed: National Institute of Health Sciences, 1–18–1 Kami-Yoga, Setagaya, Tokyo 158–8501, Japan. Tel.: +81-3-3700-1141; Fax: +81-3-3707-6950; E-mail: fumi@nihs.go.jp

shape of the peaks in the time series is so complicated and noisy that it is difficult or almost impossible to estimate the original shape of the peaks by the visual inspection. If possible, we can know something (here, a feature of the life style of people) which hides behind the sales pattern of drugs. The aim of this paper is to attempt to get a solution of this problem.

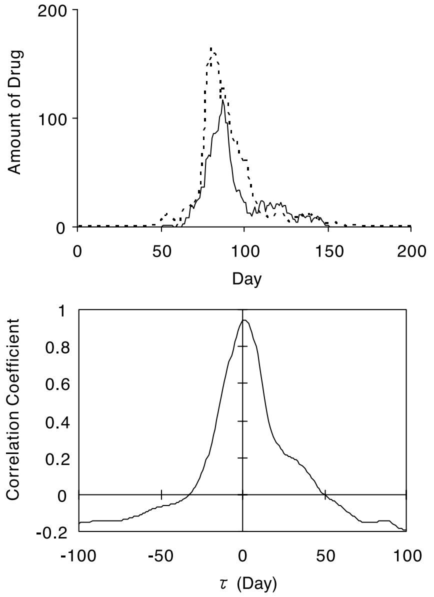

Fig. 1. Time series of Tamiflu® Capsule (...) and Tamiflu® Dry Syrup (___) and their Cross-Correlation Function at a Pharmacy

No. 6 of Table 1 lists the location of the pharmacy, time period of the drug sales and results of analysis. The Y axis denotes the number of capsules or weight (g) of dry syrup (the scale is arbitrary).

MATERIALS AND METHODS

The information about the prescriptions of influenza drugs was collected from the pharmacies listed in Table 1. The pharmacy (No. 2, 9) is located near an emergency hospital and the other pharmacies are near general clinics (but not emergency ones). The pharmacy data were available throughout the year, including those of weekends as well.

The cross-correlation function, $R(\tau)$, between time series, $A(t)$ and $B(t)$, describes the correlation coefficients between $A(t)$ and $B(t+\tau)$ as a function of τ:[14)]

$$R(\tau) = \frac{E[A(t)B(t+\tau)]}{\sqrt{E[A(t)^2]E[B(t+\tau)^2]}} \tag{1}$$

where E[.] denotes the mean over time, t, of the random variables inside the square brackets.

The moving average method[15)] used here has a window of seven days and the value at the central point of the window is replaced by the calculated average over the window. The time series shown in Figs. 1–3 are smoothed by the method.

By the definition of the cross-correlation functions, we can see that if two time series have the

Table 1. Difference between Days of Time Series Maxima (Δ), Days of Cross-Correlation Function Maxima (CF) and Difference between Gravity Centers of Time Series (DG) for Tamiflu® Capsules and Dry Syrup

No. Pharmacy	Location	Δ	CF	DG
Period: 11/01/2003–10/31/2004				
1	Shinmachi, Hanno, Saitama	3	1	2
2	Tanashi, Nishi-Tokyo, Tokyo	12	6	8
3	Tanashi, Nishi-Tokyo, Tokyo	3	2	3
4	Kouhoku, Yokohama, Kanagawa	6	7	1
5	Nakagawa, Adachi, Tokyo	5	0	−2
6	Tana,Sagamihara, Kanagawa	6	2	1
7	Asaka-machi,Koriyama, Fukushima	0	−1	−1
Period: 11/01/2004–10/31/2005				
8	Shinmachi, Hanno, Saitama	−14	−11	−11
9	Tanashi, Nishi-Tokyo, Tokyo	−1	−2	−5
10	Tanashi, Nishi-Tokyo, Tokyo	4	4	1
11	Kouhoku, Yokohama, Kanagawa	−19	−2	−1
12	Nakagawa, Adachi, Tokyo	−4	−6	−7
13	Tana,Sagamihara, Kanagawa	0	0	−2

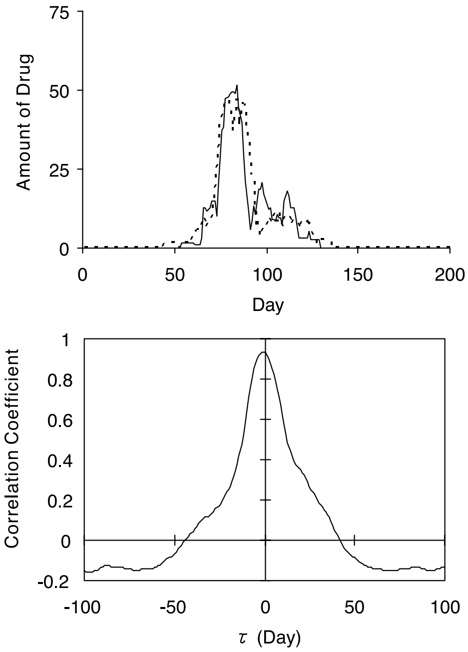

Fig. 2. Time Series of Tamiflu® Capsule (...) and Tamiflu® Dry Syrup (___) and their Cross-Correlation Function at a Pharmacy

No. 7 of Table 1 lists the location of the pharmacy, time period of the drug sales and results of analysis.

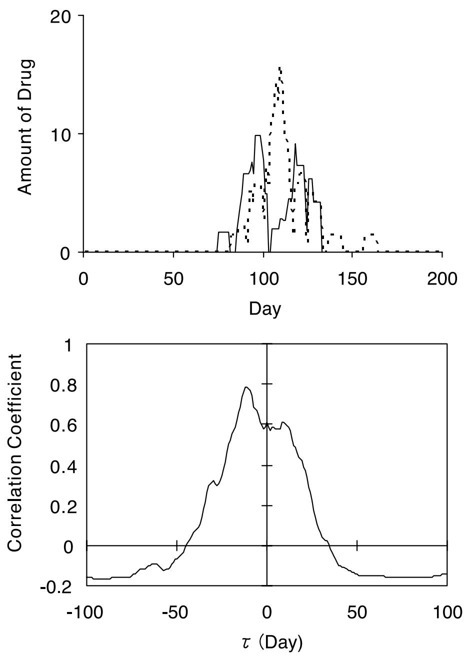

Fig. 3. Time Series of Tamiflu® Capsule (...) and Tamiflu® Dry Syrup (___) and their Cross-Correlation Function at a Pharmacy

No. 8 of Table 1 lists the location of the pharmacy, time period of the drug sales and results of analysis.

same peak shape, the difference between the peak maximum positions is equal to the maximum position (time lag) of the cross-correlation between the time series. If not, they are not necessarily the same. First, we examine how the asymmetry of time series model peaks affects the time lag and second, take some practical examples with close reference to the effects of the model experiment.

RESULTS AND DISCUSSION

Figure 4 shows the four types of model peaks which are expressed as time series, $A(t)$ and $B(t)$ in Eq. (1): peak 1: symmetrical narrow peak; peak 2: symmetrical wide peak; peak 3: tailing peak; peak 4: leading peak. The peak tops are all fixed at the same position. The cross-correlation functions, $R(\tau)$, (having gray backgrounds) are calculated between every combination of the peaks. In this paper, the time series are assigned to $A(t)$ and $B(t)$ such that if $B(t)$ lags behind $A(t)$, $R(\tau)$ has the maximum where $\tau > 0$ (positive time lag) and that if not, $R(\tau)$ has the maximum where $\tau \leq 0$ (zero or negative time lag).

From the diagonal elements of Fig. 4, we can see that the cross-correlation function has the maximum when $\tau = 0$, if the time series peaks are the same in shape, irrespective of asymmetry [see (1, 1), (2, 2), (3, 3) and (4, 4)]. The cross-correlation functions between the symmetrical peaks, even if the width is different, have no time lag [see (1, 2) and (2, 1)]. If the shapes of the peaks are different and if at least one of them is deformed, the cross-correlation functions have the maxima when $\tau \neq 0$ [see (1, 3), (1, 4), (2, 3), (2, 4), (3, 1), (3, 2), (3, 4), (4, 1), (4, 2) and (4, 3)].

The center of gravity of the symmetric peaks, 1 and 2, coincides with the positions of the peak tops, but that of the asymmetric peaks, 3 and 4, is 7 data points away from the peak top positions. The width of the fatty half of the asymmetrical peak is 50 data points and so is the half width of the symmetrical wide peak. Then, we can see the level of asymmetry as follows: the center of gravity is 14% of the half width away from the peak top.

Table 2 lists the time lags of the cross-correlation functions of Fig. 4. The time lags observed are 0, ± 5, ± 7 and ± 13 data points. The signs, plus and minus, depend on the assignment of the peaks to $A(t)$ and $B(t)$ as mentioned above. Since the peak top for every model peak is located in the same position in Fig. 4, we can assert that the cross-correla-

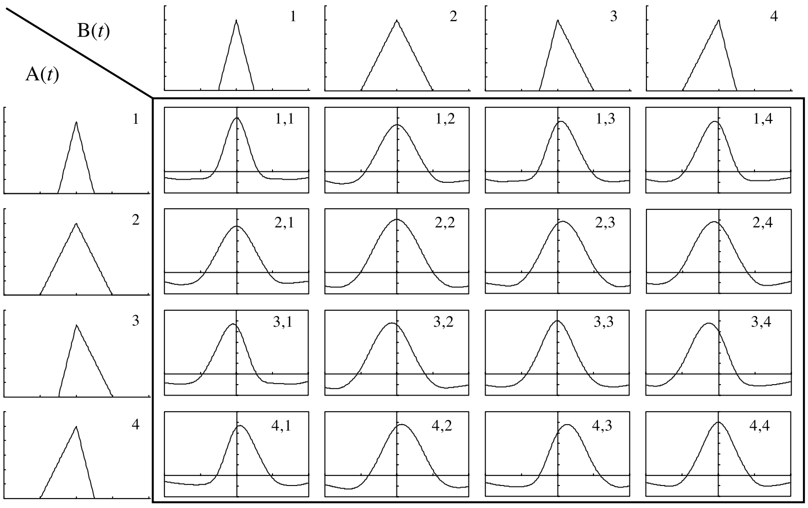

Fig. 4. Influences of Peak Shape of Time Series on Shape of Cross-Correlation Function

The triangle peaks denote the time series. Peak 1 is a symmetrical narrow peak; 2 symmetrical wide peak; 3 tailing peak; 4 leading peak. The round peaks with the gray background denote the cross-correlation functions of all the combinations of the time series. Figure (i, j) denotes the correlation function between time series i [= A(*t*)] and j [= B(*t*)]. The edge-to-edge width of the symmetrical narrow peaks is 50 data points, that of the symmetrical wide peaks is 100 data points and that of the asymmetrical peaks is 75 data points.

Table 2. Maximum Positions (Time Lags) of Cross-Correlation Functions between Model Peaks of Fig. 4

Peak A(*t*)	Peak B(*t*)			
	1	2	3	4
1	0	0	5	−5
2	0	0	7	−7
3	−5	−7	0	−13
4	5	7	13	0

tion functions refer to the centers of gravity of the peaks, rather than the difference of the peak top positions (see below).

As an example, the cross-correlation functions of the symmetrical narrow peak and tailing peak give a time lag of ± 5 points [see (1, 3) and (3, 1)] and the time lag is nearly the same as the difference between the centers of gravity (= ± 7). The time lags of the symmetrical wide peak and tailing peak are equal to the difference between the centers of gravity [= ± 7; see (2, 3) and (3, 2)]. The cross-correlation function between the tailing and leading peaks has its maximum when the time lag is ± 13 points [see (3, 4) and (4, 3)]. The time lag (= ± 13) corresponds to the distance between the centers of gravity of the asymmetrical peaks (= ± 7 × 2).

Referring to Fig. 4 and Table 2, we reconsider the time series of Fig. 1. Here, the time series of the influenza anti-viral agent for adult is assigned to $A(t)$ and that for children is $B(t)$. From the real data of Fig. 1, we can safely assume the asymmetry of the time series peaks. Then, there remain four possible situations in Fig. 4: (3, 3), (3, 4), (4, 3) and (4, 4). It is observed that the maximum positions of the time series differ by + 6 days, whereas the time lag is + 2 days. If the time series had the same peak shape, the time lag should be + 6. However, the observation implies that the actual peak asymmetry makes the time lag shift from + 6 to + 2 so that the time lag should take a minus value in Table 2. Then, the case (3, 4) is the only possibility. The time series of the adults agent will be a tailing peak (peak 3) and that for children will be a leading peak (peak 4). This peak shape is relative interpretation.

There can be various etiologies of the estimated peak shape (3, 4). We should note that the Tamiflu® agents cannot be administered repeatedly to the same patient. As for Fig. 1, the relatively faster decrease after the peak top of the children prescriptions might suggest that in the area examined (Tana, Sagamihara, Kangawa), the influenza epidemic subsided more rapidly for the children than for the adults. The temporary close of classes at elementary schools and kindergartens would play an important role.

It is not easy to find a general explanation of the pattern of the time series shown in Fig. 1. However, the thirteen case of Table 1 are indicative. Eight cases, including the case of Fig. 1, have the time lags smaller than the differences between the peak top positions of the time series. Three have the equal values and two have the larger values. Figure 2 (see No. 7 of Table 1) demonstrates the same sales pattern as Fig. 1. However, the shift due to the peak asymmetry is the minimum and it is difficult to discern the difference of the peak shape of the time series (... and ___). In Fig. 3 (also see No. 8 of Table 1), the time series of the children drug has two peaks (___) and the methodology of this paper is not applicable.

Table 1 shows that the differences between the gravity centers of the time series are close to the time lags, rather than the differences between the peak top positions of the time series, also in the real situations. This can corroborate the above-mentioned property of the cross-correlation functions.

In conclusion, it is difficult or almost impossible to find some essential elements hiding behind the time series of drug sales at pharmacies, *e.g.*, Fig. 1, by visual inspection. The comparison of the peak top positions of the time series with the time lags of their cross-correlation functions is one of the methods for this purpose.

Acknowledgements The authors would like to acknowledge Dr. Takeuchi, Hisako at Kamome Yakkyoku, Kitasato Kenkoukan (Triad Japan Co., Ltd.) and Mr. Yamaguchi, Jin (Yamaguchi Yakuhin Co., Ltd.) for their offering the pharmacy data.

REFERENCES

1) Sumi, A., Ohtomo, N., Tanaka, Y., Sawamura, S., Olsen, L. F. and Kobayashi, N. (2003) Prediction analysis for measles epidemics. *Jpn. J. Appl. Phys.*, **42**, 7611–7620.
2) Sugihara, G. and May, R. M. (1990) Nonlinear forecasting as a way of distinguishing chaos from measurement error in time series. *Nature* (London), **344**, 734–741.
3) Jose, M. V. and Bishop, R. F. (2003) Scaling properties and symmetrical patternsin the epidemiology of rotavirus infection. *Philos. Trans. R. Soc. Lond., B Biol. Sci.*, **358**, 1625–1641.
4) Segal, A. O., Crighton, E. J., Moineddin, R., Mamdani, M. and Upshur, R. E. (2005) Croup hospitalizations in Ontario: A 14-year time-series analysis. *Pediatrics*, **116**, 51–55.
5) Sumi, A., Olsen, L. F., Ohtomo, N., Tanaka, Y. and Sawamura, S. (2003) Spectral study of measles epidemics: The dependence of spectral gradient on the population size of the community. *Jpn. J. Appl. Phys.*, **42**, 721–733.
6) Ijuin, K., Hatanaka, N., Segawa, K., Nakano, T., Nakata, K., Tohara, A., Sato, M. and Hayashi, Y. (2006) Stochastic properties of prescriptions for infectious and non-infectious diseases dispensed by a city pharmacy. *Jpn. J. Pharm. Health Care Sci.*, **32**, 51–54.
7) Ijuin, K., Matsuda, R. and Hayashi, Y. (2006) A method for estimating infection route and speed of influenza. *Yakugaku Zasshi*, **126**, 161–165.
8) Ijuin, K., Matsuda, R. and Hayashi, Y. (2006) A Method for estimating the order of influenza infection between adults and children. *Yakugaku Zasshi*, **126**, 311–314.
9) Moore, D. A. and Carpenter, T. E. (1999) Spatial analytical methods and geographic information systems: use n health research and epidemiology. *Epidemiol. Rev.*, **21**, 143–161.
10) Glass, G. E. (2000) Update: Spatial aspects of epidemiology: The interface with medical geography. *Epidemiol. Rev.*, **22**, 136–139.
11) Suzuki, H., Saito, R., Shoubugawa, Y. and Sakai, T. (2005) Epidemiology of influenza. *Virus Report*, **2**, 81–87.
12) Suzuki, H., Sakai, T., Saito, R., Komata, O. and Sato, I. (2002) Use of geographic information systems for determining the epidemiology of influenza virus infections. *Kagaku Ryoho no Ryouiki*, **18**, 1801–1807.
13) Carrat, F. and Valleron, A. J. (1992) Epidemiologic mapping using the “Kriging” method: Application to an influenza-like illness epidemic in France. *Am. J. Epidemiol.*, **135**, 1293–1300.
14) Hino, M. (1982) *Spectral Analysis (Supekutoru Kaiseki)*, Asakura Shoten, Tokyo.
15) Massart, D. L., Vandeginste, B. G. M., Deming, S. N., Michotte, Y. and Kaufman, L. (1988) *Chemometrics: a textbook*, Elsevier Science Publishing Company Inc., New York.