Ann. Inst. Statist. Math.
32 (1980), Part B, 291-301

DOSE-RESPONSE PROBLEMS IN DRUG-INDUCED DISEASES AFTER CONSECUTIVE INTAKE OF LONG DURATION

SHIRO YAMAZOE, KAZUO FUKUTOMI AND TAKEMI YANAGIMOTO

(Received Aug. 24, 1978; revised June 2, 1980)

1. Introduction

Many drugs nowadays are administered for a long term in succession. In the treatment of a chronic disease like hypertension and diabetes, patients continue to take the drugs for the rest of their lives since the onset of the diseases. When a drug is suspected to induce a disease after its coming into the market, epidemiological studies, as well as animal experiments, are usually conducted to prove the causality by using patient charts in hospitals [1]. The existence of a dose-response relationship is one of the most important criteria. In epidemiological surveys, however, doses in a conventional sense can not be controlled, of course. Some devices are needed to estimate a dose-response curve and to test the existence of a dose-response relationship.

In the case of the relation between SMON (Subacute Myelo-Optico Neuropathy) and CQL (Clioquinol), epidemiologists got confused at dose-response curves obtained by a conventional method of epidemiological surveys, as reviewed in Section 2. By some researchers (for example, Meade [2]) this fact was considered to be one of serious inconsistencies which characterized the SMON controversy. Nakae and Fukutomi [3] pointed out that the conventional method is not appropriate for data of SMON and CQL. They proposed the use of some estimators in the saturated model, which were originally developed by Kaplan and Meier [4].

In this paper, we present a method for testing the existence of a dose-response relationship, which is an extension of the method of [3]. The method is a kind of test of goodness of fit and is based on the assumption that patients were given a constant dose per day.

The test statistic and model are applied to some data of SMON and CQL. In Section 7, an example on retinopathy and chloroquine is further added to show that the applicability of our method is not limited to the data of SMON and CQL.

2. SMON and CQL

SMON is a disease newly recognized in Japan in the middle of the 1950's and the number of its cases amounts to approximately 10,000. It is characterized by paresthesia in the limbs and occasionally visual disorder preceded by abdominal symptoms. It had been regarded as an infectious disease, because of clustering of cases in some areas and a regular summer peak in incidence, until the green substance extracted from the urine of patients with SMON was identified as the iron chelate of CQL in 1970. The epidemiological studies also supported the etiology that SMON was caused by CQL intake, and the Japanese Government banned the sale of CQL on 8 September, 1970. Since then, the incidence of SMON rapidly diminished to zero.

Many surveys on the dose-response problem concerning SMON and CQL have been conducted. Here we will discuss the data of females obtained by Kasai, Kanemitsu and Ito [5] and given in Table 1. The table consists of total doses taken by patients from the start of CQL intake either to the onset of neurological symptoms or to the withdrawal out of observation without any appearance of SMON. In the table, m_i denotes the number of patients suffering from SMON at a total dose from t_{i-1} to t_i, and n_i the number of those withdrawn at a total dose from t_{i-1} to t_i.

The column (4) in Table 1 shows the ratio $m_i/(m_i+n_i)$, which has been used as an estimate of the response rate at a total dose of $t_{i-1}\sim t_i$. Fig. 1 illustrates the estimated dose-response rates obtained by this conventional method. The curve is not monotone increasing, but looks rather decreasing in high doses. The curves derived from any

Table 1 Estimates of response rates from data of females of SMON and Clioquinol [5]

		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Total Dose (g)	No. of SMON m_i	No. of withdrawal h_i	Total (1)+(2)	(1)/(3)	N_i*	(1)/(5) $\hat{\pi}_i$	Estimates of P_i (Saturated)	Estimates of P_i (Probit)
1	1- 20	8	1678	1686	.0047	2215	.0036	.0036	.0047
2	21- 40	18	280	298	.0604	529	.0340	.0375	.0277
3	41- 60	8	106	114	.0702	231	.0346	.0708	.0646
4	61- 80	3	54	57	.0506	117	.0256	.0947	.1087
5	81-100	3	32	35	.0857	60	.0500	.1399	.1553
6	101-120	1	24	25	.0400	25	.0400	.1743	.2021
	Total	41	2174	2215		3177			

* $N_i=\sum_{j=i}^{k}(m_j+n_j)$

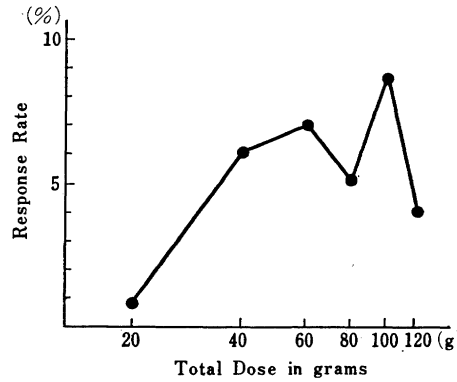

Fig. 1. Estimated response rates by a conventional method.

other data also show similar patterns. This often misled to the erroneous conclusion that a dose-response relationship between SMON and CQL was not recognizable.

3. Formulation and estimation

According to the analysis of patient charts of CQL intake, the daily dose for patients was kept constant during the treatment in each hospital.

Based on this observation, we assume that the daily dose was constant. This assumption implies that a total dose administered is proportional to the length of time and the problem can be considered as a survival problem with censored data.

Let T be a random variable which represents the total dose required by the time of the onset of SMON. Let k be the number of classes of dose levels. Let π_i and P_i be defined by

$$\pi_i=\Pr\{t_{i-1}<T\leqq t_i \mid t_{i-1}<T\}$$

(1) and

$$P_i=\Pr(T\leqq t_i) \qquad (i=1, 2, \cdots, k).$$

The probability P_i denotes a response rate corresponding to the dose t_i, which can be written in another form

$$P_i=1-\prod_{j=1}^{i}(1-\pi_j). \tag{2}$$

We assume here that the withdrawal occurs independently of the onset of SMON and only at t_i during the interval $(t_{i-1}, t_i]$. The latter assumption will be discussed briefly in Section 6. Let ω_i be the rate of the withdrawal at t_i. Then the likelihood function of the present

model by the outcome is given by

$$(3)\qquad L \propto \pi_1^{m_1} \prod_{i=2}^{k} \left[\left\{ \prod_{j=1}^{i-1} (1-\pi_j)(1-\omega_j) \right\} \pi_i \right]^{m_i} \cdot \prod_{i=1}^{k} \left[\left\{ \prod_{j=1}^{i} (1-\pi_j) \prod_{j=1}^{i-1} (1-\omega_j) \right\} \omega_i \right]^{n_i} \propto \prod_{i=1}^{k} \pi_i^{m_i} (1-\pi_i)^{N_i - m_i} .$$

This means

$$(4)\qquad L \propto \prod_{i=1}^{k} (P_i - P_{i-1})^{m_i} (1-P_i)^{n_i} ,$$

where $N_i = \sum_{j=i}^{k} (m_j + n_j)$, that is, N_i denotes the number of patients exposed beyond the total dose t_{i-1}.

The maximum likelihood estimators $\hat{\pi}_i$ and $\hat{P}_i$ of π_i and P_i, respectively, are given by

$$(5)\qquad \hat{\pi}_i = m_i/N_i \quad \text{and} \quad \hat{P}_i = 1 - \prod_{j=1}^{i} (1-\hat{\pi}_j) ,$$

provided $N_i > 0$.

$\hat{P}_i$ is the estimator of the response rate, proposed by Nakae and Fukutomi [3] in place of $m_i/(m_i+n_i)$. It is easily obtained that $\hat{\pi}_i$ is unbiased and also that $\hat{P}_i$ is unbiased, provided that there does not exist $j<i$ such that $n_j = N_j - m_j > 0$ ([4], [6]). The column (7) in Table 1 presents the estimated response rates defined by (5).

4. Test for existence of dose-response relationship between SMON and CQL

In the previous section, no restrictions on the parameters P_i or π_i are assumed. That is, $\hat{P}_i$ in (5) is the maximum likelihood estimator of P_i in the saturated model.

The estimated probability $\hat{P}_i$ is monotone increasing in i even though there exists no dose-response relationship between SMON and CQL. This means that a model which represents non-existence of a dose-response relationship is needed to test the existence of a dose-response relationship.

Suppose that SMON is caused not by CQL but by unknown factors affecting every patient, independently of total doses of CQL. Then it is reasonable to assume that the incidence of SMON depends only on the length of time. This assumption implies that the random variable T follows the exponential distribution, of which distribution function

$F(x)$ is $1-e^{-\lambda x}$ with a positive parameter λ. This model is called the exponential model.

The likelihood function of the model is then given by

$$L=\prod_{i=1}^{k}(e^{-\lambda t_{i-1}}-e^{-\lambda t_i})^{m_i}(e^{-\lambda t_i})^{n_i}\,. \tag{6}$$

Without loss of generality, we assume for simplicity that the length of the interval $t_{i-1}\sim t_i$ is a constant h. This implies that π_i is constant for every i. Let $\pi=1-e^{-\lambda h}$. Then the likelihood function can be written as

$$L=\prod_{i=1}^{k}\pi^{m_i}(1-\pi)^{N_i-m_i}\,, \tag{7}$$

which implies that the MLE $\hat{\pi}$ of π is given by

$$\hat{\pi}=\sum_{i=1}^{k}m_i\Big/\sum_{i=1}^{k}N_i\,. \tag{8}$$

The following χ^2-statistic can be used for the test of goodness of fit of the exponential model

$$\chi^2=\sum_{i=1}^{k}(m_i-e_i)^2/e_i(1-e_i/N_i)\,, \tag{9}$$

where e_i is the estimate of the expected number $N_i\hat{\pi}$ of patients with SMON under the assumption of the present model. The χ^2-statistic is asymptotically distributed according to the chi-square distribution with $k-1$ degrees of freedom under the null hypothesis.

Table 2 shows that for the data of Table 1 the null hypothesis is rejected at the significance level .01. Here the last four classes are combined so that each e_i is larger than 4, and the degrees of freedom of the χ^2-statistic is 2.

Table 2 Goodness-of-fit test applied to data in Table 1. The null hypothesis H_0 is the exponential model

	Total Dose (g)	Observed No. of SMON m_i		Expected No. of SMON e_i		$\frac{(m_i-e_i)^2}{e_i(1-e_i/N_i)}$
1	1- 20	8		28.59		15.02
2	21- 40	18		6.83		18.53
3	41- 60	8		2.98		
4	61- 80	3	15	1.51	5.59	16.05
5	81-100	3		0.77		
6	101-120	1		0.33		
	Total	41		41.00		49.60*

* $P<0.01$

5. Probit model

In toxicological studies, the threshold dose required to cause a response is usually assumed to be lognormally distributed. This assumption corresponds to the well-known probit model. While π_i in the exponential model is represented in a simple form, π_i in the probit model is represented in a rather complex form. This π_i is not necessarily monotone increasing in i.

The probability π_i is given by $(F(t_i)-F(t_{i-1}))/(1-F(t_{i-1}))$, and is related with the hazard rate function $f(x)/(1-F(x))$, where $f(x)$ is the density function of $F(x)$. We recall here that the hazard rate function of the lognormal distribution increases at first and eventually decreases for large x [7].

Under the assumption of the probit model, the likelihood function of the model is given by

$$L=\prod_{i=1}^{k}\{\Phi(a+b\log t_i)-\Phi(a+b\log t_{i-1})\}^{m_i}\cdot\{1-\Phi(a+b\log t_i)\}^{n_i}, \tag{10}$$

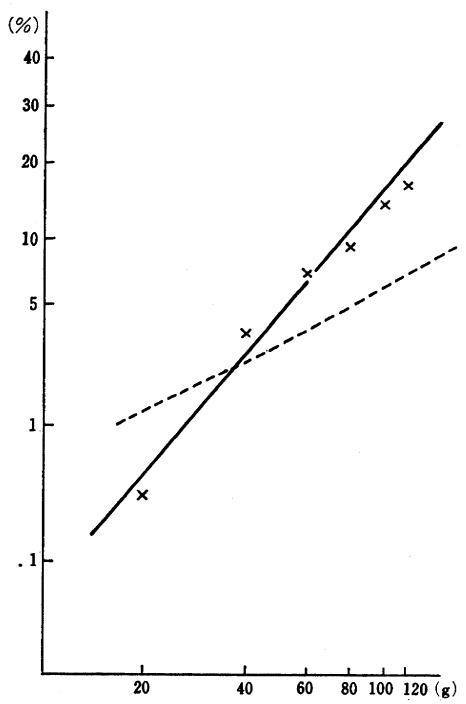

Fig. 2. Estimated response rates in the saturated model and the regression lines in the probit model (——) and in the exponential model (------).

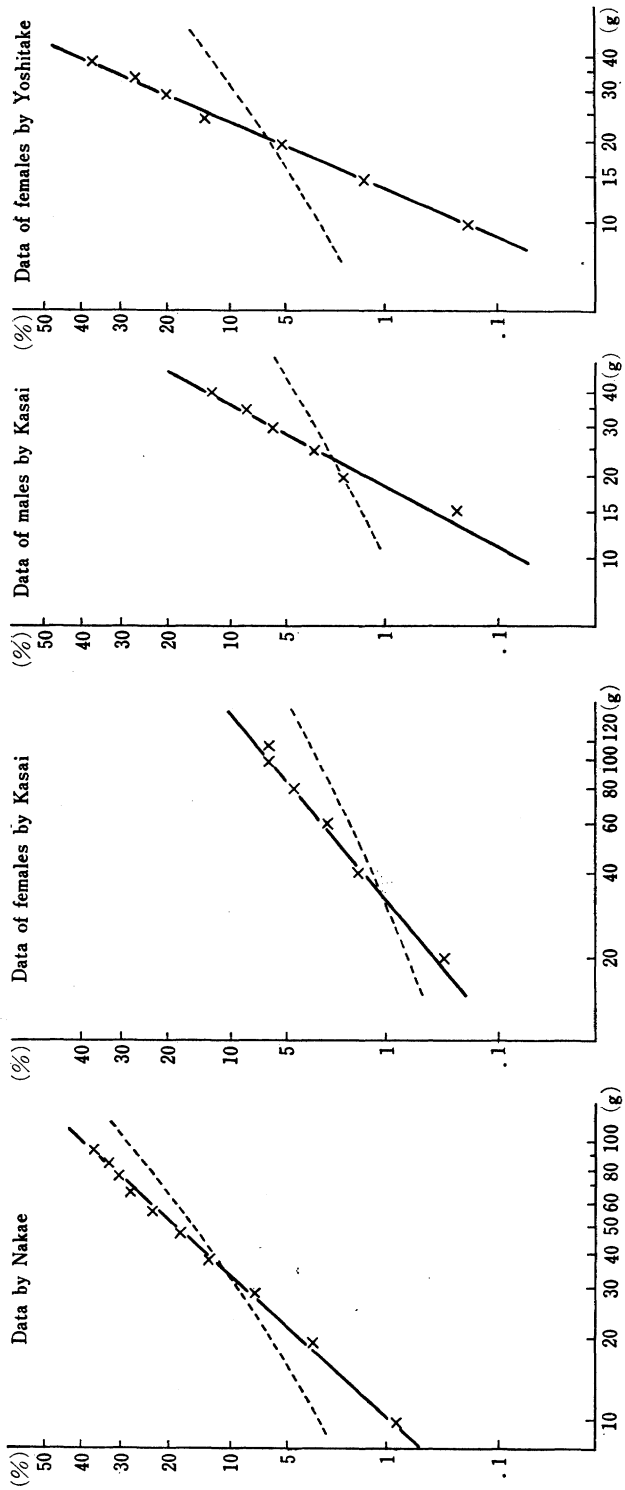

Fig. 3. Estimated response rates in the saturated model and the regression lines in the probit model (——) and in the exponential model (------).

where $\Phi(\cdot)$ is the distribution function of the standard normal distribution, and a and b are parameters to be estimated.

The maximum likelihood estimators of the parameters a and b can be calculated by the Newton-Raphson iterative method, with the initial values $a_0=b_0=0$.

The column (8) in Table 1 gives the estimated response rates obtained by the probit model. Fig. 2 shows that the estimated regression line by the probit model fits well to the response rates estimated by the saturated model, while the estimated line by the exponential model does not. In this case, the χ^2-value is 6.11 with 3 degrees of freedom.

6. Results on other data and discussions

1. As shown in Section 4, the null hypothesis that T is exponentially distributed is rejected. This suggests that the occurrence of SMON is related to the accumulated dose of CQL. The same type of analysis has been applied to other data of SMON and CQL: data of males obtained by Kasai [5] and by Nakae [8], and data of males and females by Yoshitake [9]. The results are presented in Table 3 and Fig. 3. The tests of the hypothesis of the exponential model all produced highly significant results. This consistency confirms the existence of some dose-response relationship.

2. In obtaining the estimator of the hazard rate π_i, it is recommended that the sample size N_i be modified as follows (see, for example, Littell [10]):

$$m_i/(N_i-n_i/2)\,, \qquad \text{in place of } m_i/N_i\,. \tag{11}$$

The analysis made after this modification did not produce any significant changes to all the data mentioned above. The tests for the exponen-

Table 3 Results of χ^2-tests of the exponential and the probit models

	Sex	Daily dose (g)	χ^2-value of the exponential model	d.f.	χ^2-value of the probit model	d.f.
Nakae et al. [8]	Male & Female	1.2	38.17*	6	1.48	8
Kasai et al. [5]	Male	0.90-0.94	9.64*	1	.79	1
	Female	0.90-0.94	49.60*	2	6.11	3
Yoshitake et al. [9]	Male	1.35	29.29*	2	.87	1
	Female	1.35	62.64*	2	.95	3

* $P<0.01$

tial model in all the data were found to be significant.

3. The slope of the regression line in the probit model varies fairly significantly from data to data as seen in Figs. 2 and 3. This variation can partly be explained by the fluctuation of daily doses.

7. Retinopathy and chloroquine

The present procedure is applicable to some other cases of diseases induced by consecutive constant intake of a drug. Here we review another example of a dose-response relationship between retinopathy and chloroquine.

Nylander [11] studied the incidence of retinopathy in relation to the duration of chloroquine therapy. Patients suffering from rheumatoid arthritis or lupus erythematosus disseminatus were given continuous, long term chloroquine medication. Table 4 presents the data of females by Nylander [11] and some estimates calculated for the case of retinopathy data excluding the suspected ones.

The fitted exponential and probit models are given by $P_t=1-e^{-0.3757t}$ and $P_t=\Phi(-3.684+1.697\log t)$, respectively, which are illustrated in Fig. 4. As shown in Fig. 4, the probit model shows a good fit, while the exponential model does not. In fact, chi-square values after combining classes suitably are 63.25 with 5 degrees of freedom for the exponential model and 1.02 with 3 degrees of freedom for the probit model.

In the case of retinopathy data including suspected ones, similar results were obtained. These results support a dose-response relationship between retinopathy and chloroquine, and are concordant with

Table 4 Estimates of response rates from data of females of retinopathy exclusive of the suspected ones and chloroquine [11]

Duration yrs.	No. of retinopathy m_i	(No. of suspected)	Total no. m_i+n_i	N_i	Estimates of P_i (Saturated)	Estimates of P_i (Exponential)	Estimates of P_i (Probit)
<1	0	1	25	294	.0000	.0369	.0001
1 - 2	3	4	42	269	.0112	.0724	.0061
2 - 3	5	5	52	227	.0329	.1066	.0344
3 - 4	8	2	54	175	.0771	.1395	.0916
4 - 5	11	3	42	121	.1610	.1713	.1704
5 - 6	10	3	46	79	.2672	.2018	.2601
6 - 7	5	5	21	33	.3783	.2312	.3514
7 - 9	3	0	7	12	.5337	.2869	.5180
>9	0	1	5				
Total	45	24	294				

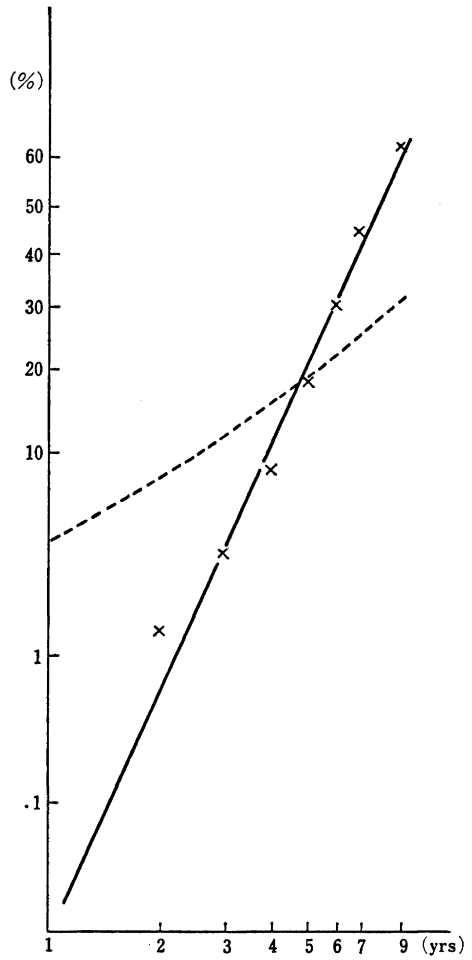

Fig. 4. Estimated response rates of the saturated model and the regression lines of the probit model (——) and of the exponential model (------).

other evidences which demonstrate the causality between them [11].

8. Conclusion

A procedure is proposed from the standpoint of the life-table analysis to evaluate the dose-response relationship between a disease and a drug by using patient charts. The models employed here are the saturated, the exponential and the probit.

The procedure is applied to the data of SMON and clioquinol and that of retinopathy and chloroquine. The results support the positive dose-response relationships between them in both the examples.

The authors would like to thank Editors for their helpful comments

in revising the paper.

SIGA UNIVERSITY
INSTITUTE OF PUBLIC HEALTH
INSTITUTE OF STATISTICAL MATHEMATICS

REFERENCES

[1] Mac Mahon, B. and Pugy, T. F. (1970). *Epidemiology, Principles and Methods*, Little, Brown and Company, Boston.

[2] Meade, T. W. (1975). Subacute myelo-optic neuropathy and clioquinol, an epidemiological case-history for diagnosis, *Brit. J. Prev. Soc. Med.*, **29**, 157-169.

[3] Nakae, K. and Fukutomi, K. (1974). Epidemiological study on the relation between subacute myelo-optic neuropathy (S.M.O.N.) and clioquinol (Report 1)—A study on dose-response relationship between S.M.O.N. and clioquinol, *Japanese J. Pub. Health*, **21**(8), 441-447. (in Japanese)

[4] Kaplan, E. L. and Meier, P. (1958). Nonparametric estimation from incomplete observations, *J. Amer. Statist. Ass.*, **53**, 457-481.

[5] Kasai, M., Kanemitsu, M. and Ito, Y. (1972). A study of dose-response relationship between doses of clioquinol and SMON incidences, Clioquinol group, SMON research commission, February. (in Japanese)

[6] Chiang, C. L. (1968). *Introduction to Stochastic Processes in Biostatistics*, John Wiley & Sons, New York.

[7] Barlow, R. E. and Proshan, F. (1965). *Mathematical Theory of Reliability*, John Wiley & Sons, New York.

[8] Nakae, K. and Shibata, T. (1975). Epidemiological study on the relation between subacute myelo-optic neuropathy (S.M.O.N.) and clioquinol (Report 2)—Relation between S.M.O.N. and clioquinol in Y hospital, Okayama Prefecture, *Japanese J. Pub. Health.*, **21**(11), 607-611. (in Japanese)

[9] Yoshitake, Y. and Ishiyama, I. (1971). On the relationship between SMON and clioquinol administration, *Igaku no Ayumi*, **77**, 148-150. (in Japanese)

[10] Littell, A. S. (1952). Estimation of the T-year survival rate from follow-up studies over a limited period of time, *Human Biology*, **24**, 87-116.

[11] Nylander, U. (1967). Ocular damage in chloroquine theraphy, *Acta Ophthalmologica*, Supplementum 92.