INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

Bell & Howell Information and Learning
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA
800-521-0600

Change-point models in the risk of anti-asthmatic drugs

Hong Yang
Department of Epidemiology and Biostatistics
McGill University

June 1998

A thesis submitted to
the Faculty of Graduate Studies and Research
in partial fulfillment of the requirement of the degree of
Master of Science

TO MY LOVING PARENTS AND JIA-I

AND TO THE MEMORY OF MY GRANDPARENTS

ACKNOWLEDGMENT

I would like to express my deepest appreciation to both of my supervisors, Dr. Samy Suissa and Dr. Robert Platt, who made such a great effort to help me complete this thesis.

I would like to thank Dr. Suissa who find this interesting research topic for me. I am particularly grateful for his continued encouragement and enthusiasm. He helped me overcome many difficulties I have faced since the very beginning of my studies. I feel very proud and fortunate to have him as my supervisor; his ability to integrate statistics into medical research and to work as a team with epidemiologists and physicians will undoubtedly have a great impact throughout my future career.

In addition, I would like to give many thanks to Dr. Platt for his constant support and valuable advice, without which this thesis may never have been completed. I wish to express my appreciation for his great patience in correcting the work in his spare time over and over again. I would particularly like to thank Dr. Platt for giving me the opportunity to present this work at the annual meeting of Statistical Society of Canada, which gave me great motivation and confidence in my current job working with him.

I would also like to thank Dr. Lawrence Joseph and Dr. Ming-Gao Gu for their critical advice and comments on the statistical methodology used in this work.

Furthermore, I would like to thank Dr. Lucie Blais for her valuable time in explaining the database and providing important references for me. I wish her the best for her career. I would also like to thank Odile Sheehy for her indispensable computer support and Diane Gaudreau for her effort in arranging meetings with Dr. Suissa as soon as possible whenever he was out of town.

Finally, I would like to thank Rick Le Grand for all the proof reading he did in order to make this thesis not only readable but also less dull.

Abstract

Epidemiologic studies indicate that excessive use of short acting beta-2 agonists is associated with increased risk of asthma mortality. We used data from a cohort of 12,301 asthmatics to fit a change-point Poisson regression model to estimate the maximum safe dose of these beta-2 agonists and its confidence limit. Using the profile likelihood method, the maximum likelihood estimate of the change-point is 1.8 canisters/month, and excessive rate of fatal or near fatal asthma attack is 3.7 (2.7-4.7) and 7.0 (3.5-10.4) per 1,000 asthmatics per year before and after the change-point. Its bootstrap 95-percentile intervals are (2, 64) and (2, 71) canisters/year respectively for non-parametric and parametric approaches. Simulation studies found the profile likelihood and bootstrap methods useful for inference of the change-point in providing safe dose information for these drugs. Future studies are needed to obtain more precise bootstrap intervals and to assess the confounding effects of covariates.

Résumé

Des études épidémiologiques révèlent que l'utilisation excessive d'agonistes du récepteur bêta-2 est associé à un risque accru de mortalité par asthme. Nous avons exploité les données tirées d'une cohorte de 12 301 asthmatiques dans le cadre d'un modèle de régression de Poisson avec point de rupture (change-point) pour évaluer la dose maximale d'agoniste du récepteur bêta-2 qu'il est possible d'administrer en toute innocuité, de même que son intervalle de confiance. Au moyen de la méthode des profils de vraisemblance, l'évaluation de la vraisemblance maximale du point de rupture est de 1,8 cartouches/mois et le taux excessif de crise d'asthme fatale ou quasi-fatale est de 3,7 (2,7-4,7) et de 7,0 (3,5-10,4) pour 1 000 asthmatiques par an avant et après le point de rupture. Les intervalles "bootstrap" (95^{e} percentile) sont de 2 à 64 et 2 à 71 cartouches/an respectivement pour les analyses non paramétrique et paramétrique. Les études de simulation révèlent que le profil de vraisemblance et les méthodes "bootstrap" sont utiles pour l'estimation du point de rupture et pour fournir des données sur les doses sûres de médicaments. D'autres études s'imposent pour obtenir des intervalles "bootstrap" plus précis et pour évaluer l'effets confoundants des autres covariables.

TABLE OF CONTENTS

1. INTRODUCTION 4

1.1 QUESTION 4
1.2 SOLUTION 5
1.3 OUTLINE OF THE THESIS 7

2. LITERATURE REVIEW 9

2.1 LIMITATIONS IN PRE-MARKETING DOSING STUDIES 9
2.2 POST-MARKETING DRUG SAFETY EVALUATION 11
2.3 ASTHMA AND ITS TREATMENT WITH BETA AGONISTS 14
2.3.1 About asthma 14
2.3.2 Beta-receptor agonists for the treatment of asthma 16
2.3.3 Asthma mortality and its association with beta-agonist bronchodilators 21
2.3.4 Summary 26
2.4 GENERAL REVIEW OF THE CHANGE-POINT PROBLEM 27
2.4.1 Introduction 27
2.4.2 Inference for the change-point problem without covariates 28
2.4.3 Multiphase regression 30
2.4.4 Other approaches 36

3. METHODS 38

3.1 ASTHMATIC COHORT FROM SASKATCHEWAN ASTHMA EPIDEMIOLOGY PROJECT (SAEP) 38
3.1.1 General description 38
3.1.2 The main outcome 39
3.1.3 The main drug exposure 40
3.2 STATISTICAL INFERENCE 40
3.2.1 Model 40
3.2.2 Maximum likelihood estimation for τ and β's 43
3.2.3 Confidence interval estimation for the change-point in model 3.4 44
3.2.4 Simulation 46

4. RESULTS 49

4.1 GENERAL DESCRIPTION: 49
4.2 POINT ESTIMATION OF THE CHANGE-POINT: 49
4.3 CONFIDENCE INTERVAL ESTIMATION FOR THE CHANGE-POINT IN MODEL 3.4: 54

5. CONCLUSION 56

5.1 DISCUSSION 56
5.2 CONCLUSION 63

6. BIBLIOGRAPHY 66

LISTS OF TABLES AND FIGURES

Table 1.	Rate of fatal or near fatal asthma attack by ordinal classification of exposure dose.
Table 2	Parameter estimates: analysis of the asthmatic cohort data
Table 3	Bootstrap confidence intervals
Table 4	Parameter estimates: simulated data
Figure 1	Asthma deaths of persons aged 5-34 compared with sales and prescriptions of asthma preparations in England and Wales for 1959-1968
Figure 2	Asthma mortality (5-34year olds) and fenoterol market share (%) in New Zealand 1974-1990
Figure 3	Plot of pollen concentration vs. depth of a lake sediment core
Figure 4	Plot of $(\text{serum-creatinine})^{-1}$ vs. time for one patient
Figure 5	Observed dose-response plot
Figure 6	Profile log likelihood plot of change-point model with identity link
Figure 7	Profile log likelihood plot of change-point model with identity link and constant rate before the change-point
Figure 8	Observed and predicted dose-response curves
Figure 9	Profile log likelihood plot of change-point model with log link
Figure 10	Observed dose-response plot on log scale
Figure11	Dose-response curve from one simulated data
Figure 12	Profile log likelihood plots for the change-point—one simulated data
Figure 13	Dose-response curve on log scale from one simulated data
Figure 14	Histogram of MLE's for the change-point estimated from simulated data

Figure 15 Histogram of non-parametric bootstrap estimates for the change-point

Figure 16 Histogram of parametric bootstrap estimates for the change-point

Figure 17 Histogram of non-parametric bootstrap estimates of the change-point in simulated data

1. Introduction

1.1 Question

One of the major challenges in drug development is to determine the recommended dose for a medication. This dose must be sufficiently high to be effective, but low enough not to be harmful. This determination is usually the result of extensive research in pre-marketing drug development. However, such a safe dose limit can not be fully established in pre-marketing clinical trials due to limitations in sample size, time and ethics.

The application of pharmacoepidemiology through post-marketing studies has several advantages. Not only can it prove the effectiveness of the drug in real practice but it can also identify certain adverse outcomes, rare and/or severe, which were undetectable in the pre-marketing phase. Post-marketing studies can adequately evaluate the adverse drug effects using post-marketing drug surveillance databases. They are collected for medical research on an on-going basis, including useful information on drug utilisation in a broad spectrum of patients, and large enough to identify rare outcomes and high dosing events. This allows us to evaluate the dose-response curve in a wider dose range so that the maximum effective dose, free of severe toxicity, can be established.

One situation where post-marketing studies could be beneficial is in the treatment of asthma with the short-acting beta-2 agonists (fenoterol and salbutamol) delivered by metered dose inhaler (MDI). Asthma is a common

disease in developed countries and its prevalence has been increasing since the turn of the century. Its treatment with these two short-acting MDI beta-2 agonists provides fast symptom relief. They are relatively safe and effective when used in lower doses. However, epidemiologic evidence suggested that excessive use of such medications may cause sudden death or near death from asthma and the dose-response curve indicated that the excessive rate of severe asthma attack dramatically increases beyond certain dose level. Thus, while the treatment of asthma with the short-acting MDI beta-2 agonists is highly effective, there is a high risk of asthma mortality associated with an excessive dose. Even though it is possible that excessive use of these medications can be a marker for asthma severity, the confounding effect of which further is discussed in Chapter 5, it is still clinically important to identify the maximum safe dose for these medications. Since clinical trials are not feasible to identify such dose limit, a large-scale post-marketing cohort study is necessary.

1.2 Solution

There are several ways of modelling an increasing dose-response curve. First, in the case of a binary outcome, the probit link or the logit link is often used to model the relationship between the probability of outcome and dose, which assumes the dose-response curve has the shape of cumulative normal or logistic distribution function, respectively. Secondly in the case of count data, Poisson regression with log link is usually used to model the rate, assuming an exponential relationship between the rate and dose. However, all of these approaches assume a

smooth increasing dose-response relationship, whereas a change-point dose-response model assumes a change-point connecting two lines with the slope after the change-point higher than the one before. With respect to the drug safety question mentioned above, such a change-point can be regarded as the maximum safe dose limit beyond which asthma mortality rate increases much faster. This can be very useful clinical information in order to treat asthmatic patients safely and effectively. Therefore, to identify such maximum safe dose, a change-point dose-response curve can be constructed from a cohort analysis to directly model the rate of fatal or near fatal asthma attacks as a function of exposure level of the short-acting MDI beta-2 agonists. The change-point problem in statistics has been studied since the 1950's, and has been found useful in various fields from operational control in industry to birth defects in paediatric epidemiology. One of the applications of the change-point problem is in multiphase linear regression, where the coefficient of X, β_x, changes at unknown point, X_τ. Most of the theoretical work for the inference of change-point focuses on linear regression with normal error and their application to real life situations has been hindered by its mathematical complexity. However, Stasinopoulos[97] wrote a macro in GLIM to calculate the maximum likelihood estimate of the change-point using the profile likelihood method, and its confidence interval was estimated by Ulm[101] using the bootstrap method. Their methods have been successfully applied to epidemiological data with binomial and Poisson error and have performed well in simulations. This thesis applies their numerical methods to give point and

confidence interval estimate of the change-point in the dose-response curve of short-acting MDI beta-agonists and fatal or near fatal asthma attacks.

1.3 Outline of the thesis

In Chapter 2, a general review of dosing studies in drug development is first given, showing the advantages of post marketing pharmacoepidemiologic studies in the evaluation of safe drug doses. This is followed by a review of beta-2 agonists in the treatment of asthma and their association with asthma mortality, which presents a typical situation in which a large scale post-marketing cohort study is necessary to investigate the safe dose limit. Finally, a critical review of the change-point problem in statistics with focus on multiphase regression is given for the inference of the change-point based on the maximum likelihood theory and alternative numerical methods.

Chapter 3 describes the data and the methods used in this thesis. First a general description of the asthmatic cohort from the Saskatchewan Asthma Epidemiology Project (SAEP) is given, and then the outcome of interest and main drug exposure are defined. Secondly, the statistical models used to fit the dose-response curve are described, which include change-point models with either identity link or log link and the general log linear model. The profile likelihood method to obtain the point estimate of the change-point is presented followed by the bootstrap method for its interval estimation. Finally, a simulation study to evaluate the performance of these numerical methods is presented.

Chapter 4 presents the results of the analysis. Parameter estimates from

various models are given and their goodness of fit are compared. 95% percentile intervals using both parametric and non-parametric bootstrap techniques are presented. As a comparison, confidence intervals based on normal theory and likelihood ratio test are also given. Finally, results from the simulation study to evaluate the goodness of point and interval estimation of the change-point are given.

In the final chapter, the advantages of modelling the dose-response curve with a change-point in post-marketing drug dosing studies are discussed. The estimated safe dose limit for the short-acting MDI beta-2 agonists is compared to the corresponding maximum recommended dose level and its implications are discussed. Finally, limitations in numerical estimation of the change-point and its confidence intervals are also evaluated.

2. Literature Review

2.1 Limitations in pre-marketing dosing studies

Pre-marketing drug development can be broadly divided into two periods, namely preclinical animal studies and clinical trials.

In order to ensure that a drug is both safe and effective when taken by humans, it goes through many testing steps in animals before being given to humans. Many drugs are abandoned at this stage for lack of effectiveness or serious toxic effect. Information concerning starting doses in humans can be obtained from these studies. However, the proper and complete integration of pre-clinical data from various test models and animal species into a single comprehensive toxicity statement of the drug, which is then extrapolated to humans, is a challenging process.[1] In general, extrapolating from animals to humans is risky. Animal studies are able to predict acute and short term adverse events in human with some degree of success,[2,3] but the predictive value of chronic toxicity such as cancer is poor.[4]

Early stage clinical trials (phase I and phase II trials) are usually clinical pharmacology studies designed to determine the pharmacokinetic and pharmacodynamic profile of drugs in human. One of the objectives of these trials is to establish safe effective drug doses to be used in the later stages of clinical testing. The studies are usually conducted in healthy individuals (phase I) to avoid severe toxicity or patients who have the target disease with fair conditions (not the most or the least sick) to be able to demonstrate initial drug efficacy (phase II)

with minimum risk of having adverse effects. The design used in these early stage trials is usually a sequential escalating dose study, which attempts to establish the dose-response relationship. The maximum tolerable dose (MTD) of the drug, with which the numbers of patients experiencing a given degree of toxicity meets some set criteria, is also obtained. The sample size of these trials are usually small (10-20 subjects). Due to the small sample size, stochastic nature of design and heterogeneity of the study subjects, statistical properties of the dose-response curve, and MTD estimation are unpredictable.[5,6,7]

Phase III trials provide conclusive information on the efficacy of the drug for specified indications and continue to evaluate other safety profiles such as long term adverse effects. These trials usually employ multicenter, parallel, double-blinded, controlled, randomised, and complete block designs with centres as blocks. The number of patients involved may range from several hundreds to more than a thousand, and the duration of such trials are relatively longer. For example, phase III of a drug trial for a chronic disease may last from 3 to 12 months or more. Phase III trials are generally less restrictive in the choice of patient selection than phase II trials because they may need to ensure representation of several distinct populations in order to establish claims in each group. However, because of ethical reasons, the inclusion-exclusion criteria in a Phase III protocol usually does not or can not include representation of all portions of the targeted population (ex. children, pregnant women etc.). Also, to increase statistical efficiency in detecting to differences between study groups,

patients with concomitant diseases or patients receiving other drugs are often excluded.

In summary, pre-marketing drug trials are necessarily limited in size, time and by ethics. As a result, some drug effects that are severe, unpredictable, rare and have longer latent periods can not be identified in these trials. Furthermore, patients in clinical trials are closely monitored, and very few have high doses of drugs, which makes it difficult to study drug overdose effects. All these make it difficult to determine the maximum effective dose, free of toxicity, in pre-marketing dosing studies. This can only be assessed after marketing when pharmacoepidemiology studies are performed using large drug surveillance data collected as part of ongoing medical care.

2.2 Post-marketing drug safety evaluation

After a drug is marketed, its safety profiles are continuously evaluated in pharmacoepidemiologic studies. The following section is a brief summary of the first part of a monograph edited by Strom.[8]

Pharmacoepidemiology is the study of use and effects of drugs in a large number of people. It is a relatively new field resulting from the union of clinical pharmacology and epidemiology. Instead of aiming at the individual level as in clinical pharmacology, pharmacoepidemiology applies epidemiologic methods to provide risk/benefit assessment of the drug at the population level. The primary focus of pharmacoepidemiology is to study adverse drug effects (ADEs), trying to identify risk factors for ADEs across the target population. The need for

pharmacoepidemiologic research has arisen from drug regulation aimed at protecting public interests.

A brief history of drug regulation in the U.S., similar to most developed countries, can serve as an example of the evolution of pharmacoepidemiology in the past several decades. Federal regulation on drug supplies began as early as 1848 with the Import Drug Act. However, it was not until 1938 that the first regulation, The Food, Drug, and Cosmetic (FD&C) Act, passed. Following the result of 100 deaths due to a sulfanilamide preparation containing diethylene glycol,[9] the Act required safety approval prior to marketing, with adequate labels and warnings. The issue of post-marketing drug safety was first addressed in the 1939 Annual Report to the FD&C Act. Following the discovery that chloramiphenicol aplastic causes anaemia,[10] adverse drug reactions received more and more attention in the 1950s. During the following decade (1960s), the new field of pharmacoepidemiology began to develop. Several in-hospital drug monitor programs were established in U.S. which explored the short term effects of drugs used in hospitals.[11,12] After the "thalidomide disaster"[13] in 1968, the Kefauver-Harris Amendments were passed in the US requiring more vigorous evaluation of drug safety and efficacy. It led to many previously approved drugs being removed from the market and also dramatically prolonged the drug approval process. Several serious and uncommon ADEs in the late 60s[14] had stimulated rapid development of this new field. In fact, since the early 1970s the

FDA has required post-marketing studies at the time of approval for about one-third of drugs.[15]

As mentioned above, pharmacoepidemiology applies epidemiologic methods to study the content area of clinical pharmacology. Epidemiologic study designs can be broadly divided into two categories, namely descriptive studies and analytic studies. Both of them can be used in the field of pharmacoepidemiology. Descriptive studies include case reports, case series and analysis of secular trend. These studies are usually inexpensive and fast, and are used for generating hypotheses. For example, the question of whether oral contraceptives (OCs) cause venous thromboembolism was first suggested by case reports and case series. It was explored in greater detail in secular trend analysis.[16] However, the lack of controls and difficulties in controlling confounding effects mean that these studies can not be used in hypothesis testing, which can only be assessed using analytic designs such as the case-control study, cohort study and clinical trials. For the same example, because of the safety concerns about OCs, a series of case-control studies[17] were carried out to investigate its causal relation with the adverse outcome. Furthermore, because of the importance of this drug and the number of women using it (most of them are healthy), a large-scale long-term cohort study[18] was conducted to further confirm the previous findings. One of the advantages of cohort analysis is the ability to investigate dose-response relationship whether an increased drug exposure (either in time or dose) causes an excessive risk of ADEs. Not only does this allow us to confirm their causal

relationship, but it also provides useful information with respect to the safe dose limit in real life. For oral contraceptives, it is neither feasible nor ethical to answer these questions in randomised clinical trials.

2.3 Asthma and its treatment with beta agonists

2.3.1 About asthma

The National Asthma Education Program Expert Panel Report defines asthma as: "Asthma is a lung disease with the following characteristics: (1) airway obstruction that is reversible (but not completely so in some patients) either spontaneously or with treatment; (2) airway inflammation; and (3) increased airway responsiveness to a variety of stimuli".[19] One of the main features of asthma is bronchial hyperactivity. Usually asthma patients tend to have a 'twitchy' airway. The major symptom of asthma abnormality is airway obstruction caused by contraction of bronchial muscles contained within the walls of the airway. An allergic inflammation can lead to swelling of the airway epithelium, accumulating inflammation products and increasing the number of mucus glands, all of which can narrow the bronchial airway causing an asthma attack.

Objective asthma diagnosis can be made using lung function tests such as peak expiratory flow rate (PEFR), or forced expiratory volume in one second (FEV1). They are simple, safe and reproducible.[20] Usually a reduction of more than 15% in PEFR or FEV1, after being challenged by a standard exercise challenge, histamine or methacholine, is normally regarded as a positive

diagnosis.[21] Skin tests and blood tests are also used for diagnosis of asthma caused by allergy. Subjectively, any two of the five symptoms, cough, tightness of chest, nocturnal misery, chest pain and wheezing, are strong indicators of asthma.[22] However since there is controversy in defining asthma and its wide rapid variation in airway obstruction, misclassification can be a potential source of bias in asthma survey study. Chronic bronchitis, lung infections as well as certain heart diseases are sometimes mislabelled as asthmatics.

There are several important risk factors for asthma at the *host level* as well as at the *environmental level*. **Allergy**[23] is increasingly being recognised as playing an important role in asthma. Allergens land within the lungs on the airway making contact with mast cells. The mast cells then lock onto the allergens and produce inflammation products such as histamine causing asthma attack. **Genetic factors:** asthma is well known to cluster in families[24] and a number of genetic mechanisms have been proposed.[25] **Birth factors** such as lower birth weight, pre-maturity, neonatal lung disease are also associated with an increased level of airway obstruction and bronchial reactivity.[26,27,28] *At the environmental level*, **smoking** is a well recognised respiratory irritant. Several studies have reported that adolescent asthmatics smoke at a level equal to or higher than the general population.[29] Tobacco smoke also increases the severity of asthma.[30] Another environmental factor is **air pollution**; incidence of asthma has been reported to increase in polluted area as compared to non-polluted areas[31] and in urban areas

vs. rural areas.[32] Studies of migrants and ethnic groups show some mild **race** differences, but **living environment** is much more important; the children of migrants have been reported to acquire the prevalence of asthma in the area where their parents moved.[33] **Socioeconomic status** is a possible factor to explain that in the US, blacks of any age group have more asthma and asthma hospitalisation compared to whites.[34] **Seasonal variations** in asthma mortality and morbidity have been reported; in general, early summer and fall demonstrate the highest rate of asthma attacks[35]. Finally, **infections** are well recognised to be a factor precipitating asthma attacks[36]; Viral infections could enhance the release of inflammatory mediators from mast cells and basophiles in the lungs.[37] Other respiratory tract infections, such as bronchitis and pneumonia, have been shown to be associated with later onset of asthma.[38]

2.3.2 Beta-receptor agonists for the treatment of asthma

Available treatment for asthma can be broadly divided into two categories: 1) bronchodilators to relieve the symptoms (i.e. bronchoconstriction) during acute asthma attacks, and 2) prophylactic therapy to prevent the onset of asthma attacks and to achieve long-term control. The most widely used bronchodilators are beta-receptor agonists, also called sympathomimetic agents. These drugs exert their actions through stimulating beta-receptors, which in turn activate a cascade of events to control the tone of bronchial smooth muscle. For prophylactic therapy, a very important group of drugs are corticosteroids. Their actions are achieved by

blocking the synthesis and release of various mediators, and reducing airway inflammation. They can be administered either orally (oral corticosteroid) or by inhalation (inhaled corticosteroid). These drugs have proven to be very effective in long-term suppressive therapy for asthma.[39]

Beta-agonists belong to a group of asthma medications with a basic amine structure. Because the pharmacological effects of these medications are similar to those observed when sympathetic autonomic nerves are stimulated, they are called sympathomimetic bronchodilators. Sympathomimetic agents have been important drugs in the treatment of asthma for the most part of the 20th century. In 1900, Solis-Cohen discovered a bronchodilating 'adrenal substance',[40] which was then purified as adrenaline and given subcutaneously to asthmatics.[41] Later, adrenaline was applied directly to the bronchial mucosa as an inhaled solution. The concentration in the inhalation solution was subsequently increased to improve its effectiveness, and it had been widely used for the treatment of asthma till the 1950s with side effects such as palpitations and tremor.[42]

From the structure-activity studies carried out during the 1940s, Boehringer-Ingelheim in Germany synthesised a derivative of adrenaline.[43] This drug effectively relieved bronchial-spasm in asthma and was more or less free of the troublesome side effects of adrenaline. The drug was later introduced under the name isoprenaline. For over 20 years isoprenaline was considered to be the drug of choice for the relief of acute asthmatic attacks with mainly cardiac side effects and tremor.

The introduction of isoprenaline played a key role in the discovery of alpha-beta receptors, first reported by Ahlquist in 1948.[44] From a summary of observations with a number of sympathomimetic agents (including the new drug isoprenaline), he proposed that these compounds could act on two kinds of receptors, which he named alpha and beta receptors. Alpha-receptors were associated with responses such as vasoconstriction and uterine contraction whereas beta-receptors were associated with certain responses such as cardiac stimulation and vasodilation. The inhibition of bronchial smooth muscle (i.e. bronchodilation) was classified by Ahlquist as a beta receptor response. It was considered at that time that future sympathomimetic bronchodilators should be specific beta-receptor agonists. Many new bronchodilators described as beta-receptor agonists were developed during the 50s and early 60s. As a result, at the beginning of the 1970s, the main beta-agonist aerosols available in the United States were adrenaline, isoprenaline, isoetharine and orciprenaline.[45]

Subsequent structure-activity studies on these four drugs were carried out to seek new beta-receptor agonists with reduced cardiac stimulating effect. In Britain, researchers at Allen and Hanburys (now part of the Glaxo group) obtained a compound called salbutamol. A structural modification of isoprenaline, this compound had increased bronchodilating potency with fewer cardiac side effects.[46,47,48] Independently, chemists at Draco (now part of the Astra group) in Sweden found a derivative of orciprenaline named terbutaline. Not only did it have oral activity and longer duration, but also, as with salbutamol, fewer cardiac

stimulating effects.[49] In Germany, Boehringer-Ingelheim also produced a series of relatives of orciprenaline. Among these was the compound named fenoterol, which again was a potent beta-agonist with less cardiac potency and longer duration shown in many clinical studies.[50,51]

The discovery of these 3 new anti-asthmatic drugs helped confirm the beta1-beta2 receptor subtype hypothesis first proposed by Lands et al.[52] in 1967. It is now clear that cardiac stimulating side-effects belong to the beta-1 receptor responses whereas bronchodilation, as well as vasodilation and muscle tremor side effects, falls into the responses of beta-2 receptor. Therefore, all these new drugs were classified as beta-2 selective agonist bronchodilators and were widely used for treatment of asthma.

The beta receptor subtype hypothesis promoted the search for compounds with better separation between bronchodilating (beta-2 response) and cardiac stimulating (beta-1 response) potency. Many were studied and were shown to have good beta-2 selectivity in animal studies. However, none were able to replace the original three beta-2 agonists in practice because of their toxicity in human. In the mean time, other research was carried on to look for beta-2 selective drugs with fewer side effects caused by vasodilation. It is a troublesome beta-2 response producing tachycardia through a reflex response to the drop in blood pressure (an indirect effect on the heart), and fine skeleton muscle. However, improvements were difficult to achieve. For unlike before, no evidence

from animal tissues suggested a difference between beta-2 receptor in respiratory tissue and that in vascular smooth muscle or skeleton muscle.

Only in recent years have several alternative beta-2 selective bronchodilators been introduced with improved pharmacokinetic profiles. Among them are salmeterol and formoterol for inhaled use, and bambuteral for oral use. These drugs have longer duration of action, increased chemical stability and increased bioavailability.

Sympathomimetic bronchodilators have been used for the treatment of asthma for most of this century. The discoveries of first beta-specific agonists in the 1950s and then beta-2 agonists in the 1970s were milestones in the history of anti-asthma drug therapy. Throughout the 70's and 80's, considerable effort in both basic and clinical research has continued to search for better specific beta-2 agonists with ideal pharmacokinetic profiles and no unwanted side effects. Yet 20 years after the development of salbutamol, terbutaline and fenoterol, these drugs were still the main bronchodilators used for acute asthma attacks. Inhaled beta-2 agonists allow a small but effective dose of drug to be delivered directly to the airways and produce a fairly quick bronchodilating response. When taken properly, these agents are relatively safe with few cardiac and muscular side effects. However, concerns[53]arose from epidemiologic evidence that increased asthma morbidity and mortality was associated with excessive use of these short acting inhaled beta-2 agonists.

2.3.3 Asthma mortality and its association with beta-agonist bronchodilators

Most recent epidemiological reviews of asthma mortality have restricted their periods of observation to approximately the last 30 years. A number of methodological issues may arise when longer-term trends are examined. Changes in the past three decades in asthma mortality appear to show two patterns: 1) sudden epidemic increases associated both in time and place with the widespread introduction of high-dose preparations of two beta-agonists, isoprenaline and fenoterol. 2) gradual increases in asthma mortality in several countries including the US, Canada, Australia, and the UK.[54] Although studies of asthma mortality over long periods of time are likely to be influenced by many factors, there is a strong belief that these trends can't be wholly explained by changes in classification or health care factors and that a real and unexplained increase has indeed occurred.[54] It is interesting to note that there was a large increase in prescription of beta-agonists aerosols between 1977 and 1985,[55] during which asthma mortality increased significantly.[56] This parallel has been recently discussed as the 'asthma paradox'[57], that is, steady increases in asthma mortality at a time of introduction and increasing sales of low-dose beta selective asthma medications.

And there was also a steady increase in asthma mortality at a time of introduction and increasing sales of low-dose beta selective asthma medications.[57]

There have been two well-documented epidemics in asthma mortality, one during the mid 1960s in six countries,[58] and one during the mid 1970s in New Zealand.[59] During these epidemics, the introduction and widespread use of two beta-agonists, isoprenaline and fenoterol respectively, paralleled in time with increase in asthma mortality. Both drugs, which are effective bronchodilators, were marketed in high-dose formulations.

The most striking aspect of the 1960s' epidemic was its abrupt onset, but only in certain countries. A striking increase in asthma mortality was first noticed in England and Wales in the mid 1960s when asthma mortality showed a sevenfold increase in younger people.[60] Several other countries such as Ireland, Scotland, Australia, New Zealand and Norway also experienced a sudden asthma mortality increase during the period of 1965-1967. In contrast, other countries (Germany, US, Canada and Belgium) showed no changes in mortality during the same period. Speizer and Doll[60]examined the asthma mortality patterns in the UK and pointed out that the abrupt increase in mortality coincided with the introduction of a new therapy, the hand-held pressurised nebulizer containing a potent non-selective beta-agonist, isoprenaline. Plotting sales of these nebulizers against deaths over time by Inman and Adelsteit[61] (Figure 1) showed excellent correspondence in England and Wales. This nebulizer hypothesis was further studied by Stolley[62] who examined the difference in constituents of these nebulizers marketed in different countries. Results indicated that all the epidemic countries had licensed and used an extremely potent form of the isoprenaline

nebulizer. The dosage form (0.4mg/puff), called "forte" by the manufacturers, was 5 times the strength of the usual dosage form (0.08/puff) licensed in non-epidemic countries. The high-dose nebulizer hypothesis, also called the "forte hypothesis", was quickly supported by the British Medical Journal with a lead editorial.[63]

A second epidemic increase in asthma mortality began in 1976. This time solely in New Zealand, the epidemic reached its highest peak in 1979 and subsided rapidly. A 2-year national survey of asthma mortality was undertaken between 1981 and 1983 by an asthma task force to investigate all deaths from asthma in New Zealand since 1976. The study confirmed the high mortality rate and suggested that rather than changes in death coding, inadequate assessment of asthma severity and/or inadequate use of oral corticosteroids may have played an important role. From the study of a cluster of asthma deaths in Auckland, Wilson et al.[64] suggested that fatal asthma attacks were a result of the increasing use of a combination of inhaled beta-agonists and slow-released theophyllines. Such treatment was more common in New Zealand than in many other countries. The association with bronchodilators was also raised by Grant et al.,[65] who suggested that the increase in mortality may have been related to widespread use of beta-agonists delivered by home nebulizers in New Zealand. Secular trend studies[66] demonstrated a dramatic increase in the sale of total beta-agonists as well as all asthma medications in New Zealand during the same period. In particularly, fenoterol, which was first marketed in 1976, had a rapid increase in market share strikingly parallel to the increase in asthma mortality (Figure 2).

In addition to such strong parallel secular trends, excessive repeated use of fenoterol was found in patients who died of asthma. Furthermore, this drug was marketed in high dose formula. They all suggested a specific hypothesis that the use of fenoterol might be associated with an increase risk of fatal asthma attack. The hypothesis was examined by three case-control studies carried out subsequently between 1989 and 1991

In the first study,[67] cases were obtained from the New Zealand Asthma Mortality Survey, including all asthma deaths in 1981-83 with age ranging from 5-45. Asthmatic controls were selected and matched on the time when the case patient died. Information on prescribed drug therapy for self-administration at the time of the last attack was documented for cases and controls. The only anti-asthma drug that was associated with a significant increased risk of asthma death was fenoterol. In the subgroup analysis defined by asthma severity markers, the odds ratio associated with fenoterol was markedly increased ranging from 2.2 to 13.3 in the most severe group indicating effect modification rather than confounding by severity. Similar findings were obtained by a second case control study which examined fatal asthma attacks during the period of 1977-1981.[68] To overcome the potential problem of information bias encountered in the first study, more accurate information on prescribed drug therapy was collected from the hospital records for both cases and controls.

A third national case-control study was done using all cases between 1981and 1987.[69] It used two control groups to explore the methodological issues

of control selection. Group A controls were chosen in the same manner as previous studies. Group B controls were selected randomly from all patients with an admission and a discharge diagnosis of asthma in the hospitals surveyed during the study period and matched for age and hospital. The purpose of control group B was to be representative of all subjects who had been admitted to the chosen hospitals irrespective of their later outcome. It was an ideal control group as a sample of the original incidence cohort. For both control groups, an increased risk was associated with the prescription of fenoterol. The subgroup analysis repeatedly showed an increased relative risk for the most severe subgroup, suggesting that the findings were not confounded by asthma severity.

The epidemiologic findings in New Zealand were confirmed in Canada by a series of studies from the Saskatchewan Asthma Epidemiology Project (SAEP).

First a nested case-control study[70] was carried out where each control was required to be at risk at the case index date. In addition, cases and controls were matched on several social demographic factors as well as asthma severity. From this study, the use of beta-2 agonists in general was found to be associated with increased risk of combined death or near-death from asthma. For asthma death only, the use of fenoterol gave an O.R. of 5.4 per canister compared with 2.4 per canister for salbutamol. After taking into account the dose difference between the two drugs (i.e. one canister of fenoterol equals to two canisters of salbutamol), their increased risks were similar on an equivalent weight basis.

Following this study, a cohort analysis[71] was carried out using the entire cohort. Using an additive model, a linear dose-response curve between MDI beta-2 agonists use and the rate of fatal or near fatal asthma attack was fitted. The background event rate in this cohort was estimated, and there were significant increased risks associated with the excessive use of fenoterol and salbutamol. The severity of asthma marked by the use of oral corticosteroid and previous asthma hospitalisation in the last 2 years was well controlled and evaluated in the analysis. It is also very interesting to note that the observed dose-response curve showed a sudden increase in the excessive risk around the recommended limit of these medications.

2.3.4 Summary

Asthma is a common airway disorder which affects up to 10% of the population in the United States,[72] and beta agonists are widely used for its treatment. However, ever since their introduction into clinical use, beta-receptor agonist drugs have been the subject of controversy with respect to their safety in asthmatics. Especially when used excessively, these drugs may cause a fatal asthma attack. The causal relation with such severe yet rare adverse outcome can not be assessed feasibly and ethically in pre-marketing clinical trials. However, until the primary defects underlying asthma are fully understood, there will be a need for bronchodilator drug therapy. The need for such therapy at the present time is not in doubt; but the way in which they are prescribed and used in asthmatics needs to be continually reassessed as our knowledge increases.

Therefore, a pharmacoepidemiologic study is needed to evaluate the safe dose limit from a large cohort.

2.4 General review of the change-point problem

2.4.1 Introduction

The change-point problem in statistics has a long history. It was first proposed by Page in 1954[73, 75, 76] in a situation where a series of observations were taken in order and the whole set of observations could be divided into subsets. Each of the subsets could be regarded as a random sample from a common distribution, which was different between the subsets. For example, here is the simplest type of change-point problem: consider an ordered (by time or by exposure level) sequence of independent observations $x_1, x_2, \ldots, x_n$ with distribution functions $F_1, F_2, \ldots, F_n$. If $F_1 = F_2 = \ldots = F_n$, then the data is homogenous, that is $x_1, x_2, \ldots, x_n$ constitute a random sample from a fixed distribution. However, suppose that for some integer τ ($1<\tau<n$), $x_1, x_2, \ldots, x_\tau$ have a common cumulative distribution function (c.d.f), F, while $x_{\tau+1}, x_{\tau+2}, \ldots, x_n$ have a c.d.f, G, then a change in distribution has taken place and τ is called the change point. Statistical inference for the change-point problem usually includes 1) testing the null hypothesis H_0: F=G against H_a: F≠G or 2) testing H_0: $\tau=\tau_0$ vs. H_a: $\tau\neq\tau_0$ and 3) point and interval estimators for τ.

The change-point problem arises in various practical fields such as epidemiology and toxicology. For example, Worsley[74] has considered the incidence of the birth defect Talipes in a region of northern New Zealand for the

years of 1960 through 1976. He assumed that the number of Talipes births each year was binomial distributed with parameter n_i (total number of births for i-th year) and p_i (the unknown probability of a Talipes birth). In 1965 the herbicide 245-T was first used in the region and he was interested to see whether this coincided with an apparent increase in the incidence of such birth defect.

2.4.2 Inference for the change-point problem without covariates

Suppose we have a sequence of independent continuous observations, x_1, x_2, ..., x_n, obtained in order, and the hypothesis is:

H_0: there is no change-point of the mean response among them

H_a: there is a change-point

A test statistic as a function of cumulative sum, defined as

$$c_k = \sum_{i=1}^{k} (x_i - \bar{x}_n)$$

where k = the number of observations before the possible change-point, was proposed. Basically, if there is no change-point, this cumulative sum should be small, and the null hypothesis is rejected for large value of the test statistic Hawkins[77] developed a recursion formula to calculate percentage points for the exact null distribution of the likelihood ratio test. His method was extended by Worsley[78] to yield percentage points in case of unknown variance using the Bonferroni inequalities. A highly accurate and easily computable approximation

to its tail probabilities was also given by Worsley.[79] To eliminate the nuisance parameter, Worsley[74,80] developed a conditional distribution for the cumsum test and the likelihood ratio test and implemented his techniques in both binomial and Poisson cases. Siegmund[81] also discussed tests to detect a change-point in the drift of Brownian motion.

When the null hypothesis (H_0: no change-point) is rejected, one wishes to estimate the change-point with its confidence interval. An asymptotic distribution of the maximum likelihood estimator (MLE) for the change-point (τ) was given by Hinkley[82,83] for the normal and binomial cases using the random walk properties of the likelihood function. He pointed out that the MLE was not a sufficient statistic for τ, even asymptotically. Therefore, he suggested estimating its confidence interval based on the likelihood ratio (LR) test. Hinkley also remarked that the MLE was not consistent because increasing the sample size n gave 'negligible' information except in the immediate neighbourhood of the true τ. More recently, Siegmund[84] extended the method of C.I. estimation based on LR to the general exponential family.

One of the extensions of the simple change-point problem is regression with a change-point or more commonly called multiphase linear regression. The British coal mining data[85] can be used to illustrate the importance of such an extension. Assuming a single change-point in Poisson process without covariates, the data has been analysed by Worsley[80] using sampling theory and by

Raftery&Akman[86] and Carlin[87] using Bayesian theory. The time of sudden change in the rate of mining accident was estimated along with a tentative explanation in terms of changes in safety practices implemented around that time. Such approaches, however, may suggest an abrupt change with exaggerated magnitude. On the other hand, regression of the accident rate on time with a change-point might lead one to attribute long run changes of accident rates to gradual changes in the coal mining industry.

2.4.3 Multiphase regression

To illustrate the multiphase regression model, consider the simple regression of y against x in which the regression function, $E[y|x] = f(x;\theta)$, takes different forms over intervals, i.e.

$$f(x;\theta) = \begin{cases} g_1(x_1;\theta_1) & x \le \tau_1 \\ g_2(x_2;\theta_2) & \tau_1 < x \le \tau_2 \\ \cdots & \\ g_D(x_D;\theta_D) & \tau_{D-1} < x \end{cases} \qquad (2.1)$$

Here the D-1 change-points and the D phase-models or regimes are unknown and needed to be estimated. Here are some examples:

1. Esterby and EI-shaarawi[88] studied the relationship between pollen concentration and the depth of a lake sediment core (Figure 3). A two-phase polynomial with one unknown change-point model was fitted to the data. Since there is an apparent discontinuity between the first 12 points and the remaining

points, they did not impose the continuity assumption that the two curves meet at the change-point. This discontinuity represented a fairly abrupt change in pollen concentration.

2. Smith and Cook[89] examined the rejection time of kidneys following renal transplants. They plotted the reciprocal of serum-creatinine versus time following surgery for one patient (Figure 4), and the change of the slope from positive to negative indicates rejection. Unlike the previous example, the graph are rather continuous, reflecting a gradual rejection process, which is more common in a biological system. To represent this, the continuity constraint at the change-point τ, i.e. $f(\tau;\theta_1) = f(\tau;\theta_2)$ for some τ satisfying $x_d \leq \tau \leq x_{d-1}$, is introduced into the model. Furthermore, the assumption of high-order derivative continuity at the change-point τ, more commonly employed in spline regression,[90] can be imposed representing a more smooth change. The same situation arose from the field of toxicology to fit a dose-response curve where the rate of response varies with dose. Usually it is reasonable to assume that there is a change-point (or threshold) beyond which the rate of toxic response increases dramatically. Again continuity constraints are made to account for the gradual biological responses. In fact, this continuous multiphase regression method is used in this thesis to fit a dose-response curve, which is doses of the short-acting MDI beta-2 agonists (fenoterol and salbutamol) vs. the rate of fatal or near fatal asthma attacks.

Consider again model 2.1, where θ_d is a vector with dimension p_d for the d-th regime (d=1,2, ...D). A continuity constraint is imposed at the change-point, τ_d, so that

$$f(\tau_d; \theta_d) = f(\tau_d; \theta_{d+1}) \quad \text{for } d = 1,2, \ldots, D-1$$

Moreover, to allow for a greater smoothness of the transition m_d-th order derivative continuity at the change-point (τ_d) can also be imposed, which means $\partial^s f / \partial^s x$ is continuous at τ_d for $s = 1, \ldots m_d$ but discontinuous for $s = m_d+1$. Now we have p regression parameters, D-1 change-point parameters, D-1 continuity constrains and m_T smoothing constrains. Thus the total number of 'free parameters to be estimated' is $p-m_T$, where

$$p = \sum^{D} p_d \text{ and } m_T = \sum^{D-1} m_d$$

An important special case of this model is the so called 'two-line segment problem', where

$$E[y_i|x_i] = \begin{cases} \beta_{10}+\beta_{11}x_i & x_i \leq \tau \\ \beta_{20}+\beta_{21}x_i & x_i > \tau \end{cases} \qquad (2.2)$$

with continuity constraint $\beta_{10}+\beta_{11}\tau=\beta_{20}+\beta_{21}\tau$ but no smoothing (i.e. first derivative discontinuity $m_{d=1}=0$). Here we have two phases (D=2) and one change-point with a total of 4 'free parameters' to be estimated (P=4 and $m_T=0$).

Model 2.2 was first studied by Sprent[91] under the assumption of identical independent normal error with $var(\varepsilon_i)=0$. He derived the likelihood ratio statistics for testing H_0: $\tau=\tau_0$ vs. H_a: $\tau\neq\tau_0$ and discussed its application in biometry. To

estimate the unknown change-point τ, Hudson[92] gave a concise method of calculating the least square estimate τ_{lse}, which is also the maximum likelihood estimate in the normal case. The distribution of parameter estimates was first studied by Hinkley.[93,94] He found that asymptotic normal distributions for β's are good approximations for moderate sample size, whereas the usual normal approximation for τ tended to be poor. In fact he derived an alternative asymptotic distribution of the MLE which gave better small sample properties. Feder[95,96] gave a rigorous treatment for the more general case (model 2.1). In his paper, in addition to the normal error assumption, further conditions were also required:

- $E[|\varepsilon_i|^{2-\delta}]<\infty$ for some $\delta>0$.
- The number of phases D must be known or in his words the two adjacent regimes must be 'identifiable' at the true θ, τ.
- As $n\rightarrow\infty$, the number of observations falling into in each interval must also tend to be infinity.

Under these conditions, Feder proved the consistency of the MLE for θ and τ, and also derived their asymptotic distributions if all m_d's are odd or 0. In fact, the MLE for θ_d, the regression parameters for d-th segment, has a multivariate normal distribution with

$$\theta_{d,\,mle} \sim N_{pd}(\theta_d, \sigma^2(\mathbf{X}_d'\mathbf{X}_d)^{-1}) \qquad \text{for } d = 1,2\ldots D \text{ if all } m_d\text{'s} = \text{odd or } 0.$$

where the design matrix $\mathbf{X}_D$ can be obtained from all the data points within the d-th interval. σ^2 can be estimated by $S^2(\theta_{mle}, \tau_{mle})/(n-p)$, where S^2 is the residual sum

of squares and p is the number of "free parameters". For the special case (model 2.2) where D =2 and $m_d = 0$, using Feder's theorems[95], we can get, for d = 1,2

$$\beta_{d1,\,mle} \sim N(\beta_{d1},\ \sigma^2/[\Sigma_j(x_{dj}-x_{d.})^2])$$

$$\beta_{d0,\,mle} \sim N(\beta_{d0},\ \sigma^2\Sigma_j x_{dj}^2/[n_j\Sigma_j(x_{dj}-x_{d.})^2])$$

$$\mathrm{cov}(\beta_{d0,\,mle}, \beta_{d1,\,mle}) = -\sigma^2 x_{d.}^2/[\Sigma_j(x_{dj}-x_{d.})^2]$$

$$\text{for } j = 1\ 2\ \ldots\ n_d$$

and $$\tau = (\beta_{10,\,mle}-\beta_{20,\,mle})/(\beta_{21,\,mle}-\beta_{11,\,mle})$$

$$\tau_{mle} \sim N(\tau, \sigma^2_\tau)$$

where $$\mathrm{var}(\tau_{mle}) = \sigma^2_\tau = (\beta_{21,\,mle}-\beta_{11,\,mle})^{-2}\{(\mathrm{var}[\beta_{10,\,mle}]+\mathrm{var}[\beta_{20,\,mle}])$$
$$+2\tau_{mle}(\mathrm{cov}[\beta_{10,\,mle}, \beta_{11,\,mle}]+\mathrm{cov}[\beta_{20,\,mle}, \beta_{21,\,mle}])$$
$$+\tau_{mle}^2(\mathrm{var}[\beta_{11,\,mle}]+\mathrm{var}[\beta_{21,\,mle}])\}$$

For confidence intervals for τ, both Hinkley[94] and Feder[96] recommended the use of LR statistics. Suppose we have a null hypothesis, H_0: $\tau = \tau_0$ and $\theta = \theta_0$. It has q independent restrictions reducing parameter space to P-q, where $P = p-m_T$ is the parameter space for model 2.2, then the likelihood ratio statistic is defined as

$$LR = -2\log\{L(\theta_0, \tau_0)/L(\theta_{mle}, \tau_{mle})\},$$

where $(\theta_{mle}, \tau_{mle})$ are the MLE of (θ, τ) and (θ_0, τ_0) are the MLE of (θ, τ) subject to restrictions imposed under the H_0. Feder proved that, asymptotically under the null hypothesis, that

$$LR \sim \chi^2_q$$

Several conditions are required 1) all m_d's must be odd or 0, 2) $E[|\varepsilon_i|^{2-\delta}]<\infty$ for some $\delta>0$ but need not to be normal, and 3) the parameters of the model must remain 'identifiable' under H_0. These restrictions rule out one very important hypothesis: no change-point under the null. Under such a hypothesis, the intersection between two phases becomes unidentifiable. So far theoretical inferences for θ, τ in model 2.1 or even model 2.2 for non-normal error term, say binomial or Poisson have not been worked out, although Bayesian solutions are available[87].

Statistical inference for the change-point can be made numerically. Stasinopoulos[97,98] wrote a macro in GLIM[99] to calculate the MLE's for θ and τ for piecewise polynomials with one change-point by maximizing the profile log likelihood function. Suppose for a given τ_0, the corresponding log likelihood function is $l(\tau_0)$, which could subsequently be maximized over θ. The MLE for the change-point τ is the value which gives the maximum of these l's (or the so-called 'profile log likelihood function') and the corresponding θ_{mle} is the MLE for θ. For a continuous change-point parameter, the macro first locates the interval where the maximum of the profile likelihood function lies through a rough search and then use a Golden Section Search[100] to get the maximum point. The Golden Section Search is efficient in one dimension, which is the case for one change-point situation. The method has been applied by Ulm[101] and Stasinopoulos[102] to Poisson and binomial data from several epidemiologic studies. To estimate the

confidence interval for τ, Ulm[101] applied the Bootstrap resampling technique[103] to the change-point data and showed its percentile interval performed well in the simulation.

2.4.4 Other approaches

The method of deriving Bayesian type change-point test statistics to detect the location change at an unknown point was first introduced by Chernoff & Zacks[104] and subsequently studied by Garder,[105] MacNeill,[106] Sen & Srivastava,[107] and Jandhyala & McaNeill.[108] Bayesian analysis in multiphase regression was studied by Bacon & Watts,[109] Ferreira,[110] Chin Choy & Broemeling,[111] Smith & Cook[89] Moen[112] and Jandhyala & MacNeill.[113] In general, using both non-informative and informative priors for parameters τ and θ, the marginal posterior probabilities of the change point at various possible points, $1 \leq \tau \leq n$, were derived. However, such an approach usually involves considerable analytical effort. Carlin et al[87] investigated the hierarchical Bayesian analysis using a Gibbs sampler[114] avoiding sophisticated high dimensional integration of the posterior joint distribution. The procedure was then applied to regressions, Poisson processes and Markov chains with change-points at unknown points. Bayesian analysis is also proved to be useful in the 'unidentifiable situation', where the null hypothesis is 'no change-point'.[87]

Non-parametric inference for the change-point problem was generally based on the rank statistics. It was first studied by Pettitt[116] who suggested a

version of Mann-Whitney statistics with its application to binary as well as continuous observations. More general cases were studied by Lambard[117] and Sen.[118] Wolfe & Schechtman[119] made a small-sample power comparisons between various statistics based on the rank test, which were reviewed by Lambard.[120] In the context of regression, non-parametric response curve or surface can be fitted using smoothing spline techniques[121] such as kernel smoothers with the point connecting the two segments referred as a knot.

3. Methods

3.1 Asthmatic cohort from Saskatchewan Asthma Epidemiology Project (SAEP)

3.1.1 General description

The computerized files of the Saskatchewan Prescription Drug hold over 20 million prescriptions for drugs listed in the Saskatchewan formulary that had been dispensed to its eligible residents between ages of 5-54. This large database includes 68,813 subjects who had at least one prescription of commonly used asthma medication from 1980 to 1987. These drugs are fenoterol, salbutamol, metaprotetrenol, terbutaline, any compound of theophylline, ipratropium bromide, cromolyn, and inhaled beclomethasone. From this geographically defined cohort, a subset of 12,301 patients who had at least 10 prescriptions over the 10 years from 1978 to 1987 were selected. The entry date for this cohort was defined as the date of the subject's 10^{th} dispensed prescription, the subjects 5^{th} birthday, or Jan.1, 1980, which ever was the latest. The exit date from the cohort was the subject's 55^{th} birthday, the date of the outcome (asthma death or near death) or April 30 1987, whichever was the earliest. As a result, the total follow-up time for the entire cohort is 574,103 person-months (or 47,849 person-years). For each month of every patient's follow-up time, information on the use of the two beta-agonists, fenoterol and salbutamol, in the previous one year (our main exposure of interest defined in section 3.1.3) is recorded. In addition, in the same month, his/her main outcome status (defined in the next section) is also available in this cohort

3.1.2 The main outcome

The primary outcome we are interested in is the combined outcome of asthma death and near-fatal attacks during the 10-year study period. If a subject who died of asthma had a previous near-fatal asthma attack, then death was chosen as the outcome. If a patient experienced several near-fatal asthma attacks, then the very first episode was counted as the outcome. This definition of the end-point in case of repeated episodes is inconsistent, and is one of the problems in this cohort data set. Ideally, the very first near-fatal asthma attack should be defined as the end-point for patients with multiple episodes. From this cohort, there are 180 deaths identified through death certificates, coroner's reports, autopsy results and hospital discharge summaries. Three physicians with special expertise in asthma reviewed all the information for 165 deaths and no documents could be found for the remaining 15 deaths. 44 deaths were categorized as probably due to asthma, among which 40 deaths received complete agreement independently in their classification. The remaining 4 were classified by consensus. The criteria for near-fatal asthma attack are either hypercarbia, or non-elective intubation during an acute asthma attack. To identify such near-fatal asthma, cohort information on procedures and billings corresponding to cardiopulmonary resuscitation, airway intubation, or assisted ventilation in the hospitalized patients, whose discharged diagnosis suggested airway disease, were searched from the database. Based on this information, the three consultants reached complete agreements independently on 80 cases of near-fatal asthma and

5 cases by consensus. Therefore, the entire asthmatic cohort contains a total of 129 cases of asthma death or near death from 1980 to 1987.

3.1.3 The main drug exposure

The main drug exposure we are interested in is chronic use of inhaled fenoterol and salbutamol delivered by a metered-dose inhaler (MDI). To evaluate such drug exposure for a given patient the total number of canisters of fenoterol or salbutamol dispensed in the previous one year was computed for each of his/her follow-up months. The one-year exposure window also controlled for the seasonal variation in the use of these anti-asthmatic drugs. Such a definition of drug exposure provides an average dose profile dispensed monthly in one year. This definition is limited, however, in that it assumes the same outcome risk for the patients with the same average drug exposure level. This may not, however, be true, since they may have different drug use patterns in the one-year exposure window and their risk of fatal or near fatal asthma attack may not be the same[115]. Under this definition of drug exposure, we redistributed the total number of person-months and outcome cases in the entire cohort into each observed dose level and use it as the working data set for our cohort analysis.

3.2 Statistical inference

3.2.1 Model

First, the simple two-line regression model (model 2.2) is chosen to fit the dose response curve. The response is the rate of fatal or near fatal asthma attack, and the dose level is measured by the total number of MDI canisters of both

fenoterol and salbutamol dispensed in the previous year. One change-point among the observed dose levels is assumed to reflect the hypothesis that excessive use of these medications is associated with a marked increase in fatal or near fatal asthma attack rate. Moreover, the continuity constraint (i.e. the two regression lines meet at the change point) is imposed to reflect a gradual continuous rate increase at the change-point. Since we model the rate here, Poisson regression is used. From model 2.2, we have

$$g(E[rate_i|dose_i]) = \begin{cases} \beta_{10}+\beta_{11}dose_i & dose_i \leq \tau \\ \beta_{20}+\beta_{21}dose_i & dose_i > \tau \end{cases} \qquad (3.1)$$

with continuity constraint $\beta_{10}+\beta_{11}\tau=\beta_{20}+\beta_{21}\tau$ and no smoothing (i.e. first derivative discontinuity at the change-point), where $rate_i$ is the rate of fatal or near fatal asthma attack at i-th dose level, τ is the change-point, and g is the link function equal to identity or log link. Combining the two equations of model 3.1 into one, we get

$$g(E[rate_i|x_i]) = \beta_{10}+\beta_{11}(dose_i-(dose_i-\tau)I)+\beta_{21}(dose_i-\tau)I \qquad (3.2)$$

where

$$I = \begin{cases} 1 \text{ if } x_i > \tau \\ 0 \text{ otherwise} \end{cases}$$

from this model, we have

$$g(E[rate_i|dose_i]) = \beta_{10}+\beta_{11}dose_i$$

and

$$g(E[rate_i|dose_i]) = (\beta_{10}+(\beta_{11}-\beta_{21})\tau)+\beta_{21}dose_i$$

before and after the change-point respectively. The two lines meet at the change-point with total of 4 'free unknown parameters' to be estimated, β_{10}, β_{11}, β_{21} and τ. For a given change-point τ', model 3.2 is fitted with identity link in the following way: from model 3.2, we have

$$E[rate_i] = E[events_i/P\text{-}M_i]$$

$$= \beta_{10}+\beta_{11}(dose_i-(dose_i-\tau')I)+\beta_{21}(dose_i-\tau')I \qquad (3.3)$$

therefore

$$E[events_i] = \beta_{10}(P\text{-}M_i)+\beta_{11}(dose_i-(dose_i-\tau')I)(P\text{-}M_i)$$

$$+\beta_{21}(dose_i-\tau')I(P\text{-}M_i) \qquad (3.4)$$

where $events_i$ is the total number of fatal or near fatal asthma attack at i-th dose level, and $P\text{-}M_i$ is the total person-months of follow-up at that level. Assuming $events_i \sim$ Poisson distribution with

$$\lambda = \beta_{10}(P\text{-}M_i)+\beta_{11}(dose_i-(dose_i-\tau')I)(P\text{-}M_i)+\beta_{21}(dose_i-\tau')I)(P\text{-}M_i),$$

model 3.4 is fitted by using SAS 6.12 PROC GENMOD procedure with identity link, Poisson error and no intercept.

Secondly, to model a dose-response curve assuming an exponential structure, ordinary Poisson regression with its canonical log link is also fitted to the data. To do so, we have

$$\log(E[events_i/P\text{-}M_i]) = \beta_{10}+\beta_{11}(dose_i) \qquad (3.5)$$

and therefore

$$\log(E[events_i) = \beta_{10}+\beta_{11}(dose_i)+\log(P\text{-}M_i) \qquad (3.6)$$

similarly assuming $events_i \sim$ Poisson distribution with

$$\lambda = \beta_{10}+\beta_{11}(dose_i)+\log(P\text{-}M_i)$$

model 3.6 is fitted by Poisson regression with log link and $\log(P\text{-}M_i)$ being an offset variable.

Both the change-point model with identity link (model 3.4) and the log-linear model (model 3.6) are used to model the dose-response curve where the absolute outcome risk increases rapidly at high dose levels. However, the two can not be used together to model such a dose-response structure. In other words, the change-point model with log link can be used in a situation where there is a rapid increase in relative risk (as opposed to absolute risk) beyond a certain dose level. This is not the case in our study and to demonstrate this, such a model is also fitted. Again from model 3.2, we have

$$\log(E[events_i/P\text{-}M_i]) = \beta_{10}+\beta_{11}(dose_i-(dose_i-\tau')I) + \beta_{21}(dose_i-\tau')I \qquad (3.7)$$

which equals to

$$\log(E[events_i]) = \beta_{10}+\beta_{11}(dose_i-(dose_i-\tau')I) + \beta_{21}(dose_i-\tau')I+\log(P\text{-}M_i) \qquad (3.8)$$

3.2.2 Maximum likelihood estimation for τ and β's

To fit a change-point dose-response model, the maximum likelihood estimate for the change-point is obtained by maximizing the profile likelihood in τ. For each fixed value of the change-point τ, the β parameters in model 3.2 and model 3.4 are estimated by the maximum likelihood method. They are then substituted into the likelihood. The resulting function of τ is called the profile

likelihood for τ, which can be plotted. Because of the chronic one-year exposure window and one canister increment between almost all the observed doses, especially at the range where the change-point is more likely to occur, it is reasonable to assume the change-point could only occur at one observed dose level. Therefore, the Golden Section Linear Search, which is useful to search for a change-point lying somewhere between the two adjacent observed dose levels, is not used here. Furthermore, the range of possible change-points is restricted between 2 to 72 canisters per year in which the change-point is believed to occur. The MLE for the change-point (τ) is the dose level among this range which gives the highest value of log likelihood. And at this dose level, the corresponding fitted change-point model (model 3.4 or model 3.8) gives the MLEs for βs.

3.2.3 Confidence interval estimation for the change-point in model 3.4

Consider the null hypothesis: H_0 $\tau=\tau_0$, for the change-point parameter. The corresponding likelihood ratio test is:

$$\lambda_{obs} = -2\log\{(L(\tau_0)/L(\tau_{MLE})\}$$

where L(.) is the likelihood function. Using the theorem given by Feder[96]

$$\lambda_{obs} \sim \chi^2_1$$

and the rejection region is $\lambda_{obs}>\chi^2_{1,\alpha}$, which is equivalent to

$$l(\tau_0)<l(\tau_{MLE})-0.5*\chi^2_{1,\alpha}$$

where l(.) is the log likelihood function. Thus the null hypothesis is not rejected if the log likelihood value under the null is not more than $0.5\chi^2_{1,\alpha}$ units less than the maximum log likelihood value at τ_{MLE}. Therefore, the values of τ_0 that satisfy this

requirement are a $(1-\alpha)\%$ likelihood ratio test-based confidence interval for τ.[122] It can be easily obtained from the corresponding profile log likelihood plot.

The bootstrap resampling technique was also used to construct confidence intervals for the change-point. Usually for the non-parametric bootstrap, one bootstrap sample is obtained by sampling with replacement from the original observed data. From this bootstrap sample, we would need to redistribute the total number of person-months and outcome cases into each dose level, and both of them are assumed random at each dose level. However, such resampling is very slow, since we have to resample 574,103 person-months from the original entire cohort to get one bootstrap replicate.

To overcome this problem, an alternative sampling approach is taken assuming that the person-months at each observed dose level are fixed rather than random. It is a reasonable assumption in the sense that the higher the dose level is, the fewer person-months are at that level. Under such an assumption, a bootstrap sample is obtained by sampling with replacement from the person-months at each observed dose level rather from the entire cohort. Since the only random quantity we need to know is the number of outcome cases at each observed dose level, this turns out to be a very simple process, which is equivalent to generating a binomial random variable x_i, satisfying

$$x_i \sim \text{Bin}(n_{i\,obs}, p_{i,\,obs}) \quad i=1,2,\ldots,D$$

where $n_{i,\,obs}$ and $p_{i,\,obs}$ are the observed number of person-months and outcome risk respectively at the i-th dose level, and D is the maximum observed dose level.

Therefore, to get one non-parametric bootstrap sample, D observations were generated. Each observation has $n_{i,\,obs}$ and x_i, where $n_{i\,obs}$ is the number of observed person-months at i-th dose level and x_i is the number of outcome cases generated at that dose level by the binomial distribution described above. This bootstrap replication is then fitted by model 3.4 to get one bootstrap estimate of the change-point.

In addition, a parametric bootstrap is also done. By fitting model 3.4 to the observed data with the change-point equal to its MLE, the predicted outcome risk at each observed dose level, $p_{i,\,pred}$, is obtained. Instead of using $p_{i,\,obs}$, $p_{i,\,pred}$ is used as p_i to generate the binomial variable x_i to give a parametric bootstrap sample. For each type of bootstrap sampling, 1500 samples are generated from which the 95-percentile confidence interval is obtained.

Finally, as a comparison, confidence intervals based on normal approximation of the bootstrap estimates are also given.

3.2.4 Simulation

To assess the performance of point and interval estimation for the change-point in model 3.4 using profile likelihood and bootstrap methods respectively, a simulation study is done. To get one simulated change-point data set, D binomial observations are generated each representing the number of cases, x_i, at i-th dose level, i.e.

$$x_i \sim \text{Bin}\,(n_{i,\,simu},\, p_{i,\,simu}),\ i=1,2,\ldots,\ D$$

where D is the maximum observed dose level in the cohort.

First, the binomial parameter $p_{i,\,simu}$ at i-th dose level is determined by the following the change-point model:

$$p_{i,\,simu}=\begin{cases}\beta_{10}+\beta_{11}dose_i & \text{if} \quad dose_i<=\tau \\ \beta_{20}+\beta_{21}dose_i & \text{if} \quad dose_i>\tau\end{cases} \tag{3.9}$$

we assume:

1) before the change point, a constant rate of 4 fatal or near fatal asthma attacks per 10,000 asthmatics per year with a rate difference equal to 0 or an odds ratio equal to 1 (i.e. $\beta_{10}=0.0004$ and $\beta_{11}=0$) for any excessive use of the drugs,
2) after the change-point, a sudden increase of rate difference to 8 per 10,000 asthmatics per year for every additional canister expensed per month, which gives an odds ratio of 3 for the very first monthly additional canister immediately after the change-point (i.e. $\beta_{21}=0.0008$),
3) and the change-point is assumed to be at the MLE of the change-point estimated from model 3.4 using the profile likelihood method (the best guess for the change-point),

Thus we have $\beta_{11}=0$, $\beta_{21}=8X10^{-4}$, $\beta_{10}=4X10^{-4}$ and β_{20} can be calculated using the continuity constraint, which is

$$\beta_{10}+\beta_{11}\tau_{mle} = \beta_{20}+\beta_{21}\tau_{mle}$$

Secondly, the binomial parameter $n_{i,\,simu}$ is large enough in order to get at least 5 cases at each dose level so that the simulated sample size is optimal for our

methods used for inference of the change-point. The way to do it is to let $n_{i,\,simu}$ equals to round $(5/p_{i,\,simu})$.

Therefore, one simulated change-point data set has D observations (or D dose levels) each with $n_{i,\,simu}$, $x_{i,\,simu}$ representing the total number of person-months and cases at i-th dose level.

100 such simulated change-point data sets are generated and subsequently fitted by model 3.4. The bias and variance of the MLE for the change-point (τ_{mle}) obtained by the profile likelihood method is assessed from the distribution of the 100 τ_{mle}s

Usually in practice, the non-parametric bootstrap method is used because it only relies on the empirical distribution of observed data. Therefore, to evaluate its 95-percentile interval estimation for the change-point in model 3.4 in ideal sample size situation, a simulated data set is randomly generated by the structure described above with seed=10 in SAS. 1500 non-parametric bootstrap samples were obtained from the simulated data set to construct the 95-percentile interval for the change-point.

4. Results

4.1 General description:

From this asthmatic cohort, there is a total of 47,842 person-years of follow-up for the 12,301 patients, in which 44 asthma deaths and 85 near asthma occurred. The overall crude rate of fatal or near fatal asthma attacks (the main outcome of interest) for this cohort is 2.70 per 1,000 asthmatics per year.

Table 1 shows the frequency distribution of main outcome and follow-up time by ordinal classification of the short-acting MDI beta-2 agonists. The corresponding observed outcome rates at each grouped dose level are also given. The observed dose-response plot (Figure 5) shows a clearly increasing trend, gradual at lower dose levels and more dramatic later on, particularly starting from 20 canisters per year. This dose-response pattern is supported by previous review of asthma medications. It was believed that the short-acting MDI beta-2 agonists, in this case fenoterol and salbutamol, are relatively safe with potential beneficial effects for asthma when used at recommended dose levels but the risk of asthma mortality increases when used excessively.

4.2 Point estimation of the change-point:

The asthmatic cohort was fitted by three models, the change-point model with identity link (model 3.4), the log-linear model (model 3.6), and the change-point model with log link (model 3.8). Table 2 gives the parameter estimates for each of these 3 models.

First, for model 3.4, the profile log likelihood plot for the change-point obtained is shown in Figure 6. Between 2 canisters/year and 72 canisters/year, the searching limits where the change-point is believed to exist, the model with the change-point at 21 canisters/year has the highest value of the log likelihood. Additional restriction was imposed by assuming a constant rate of fatal asthma attack (β_{11}=0) before the change-point. This model resulted in no change-point among the searching range shown in Figure 7, where the maximum profile log-likelihood is obtained at the lower limit.

Also from model 3.4, the maximum likelihood estimates of the excessive rates before (β_{11}) and after (β_{21}) the change-point were obtained with the change-point being at 21 canisters per year (the MLE for the change-point). From this model, the predicted dose-response curve, i.e. the predicted rate of fatal or near fatal asthma attack at each drug exposure level, is shown in Figure 8. Here the base-line rate is 2.4 (0.21-4.7) per 10,000 asthmatics per year. Before the change-point, there is a significant excessive rate (β_{11}) of 3.7 (2.7-4.7) per 1,000 asthmatics per year for each additional canister of the drug dispensed monthly. After the change-point, the excessive rate (β_{21}) increases to 7.0 (3.5-10.4) per 1,000 asthmatics per year, nearly double that before the change-point.

Secondly, instead of using a change-point structure to model a dose-response structure in which the response rate increases slowly at lower doses followed by a dramatic jump at higher dose levels, the ordinary log-linear Poisson regression model (model 3.6) without change-point was also used. However, the

two have different assumptions. The change-point model assumes outcome rates increase linearly with two different slopes, whereas the log-linear model assumes an exponentially increasing rate structure. As a rough method of comparing the fitness of the two models, the scaled deviance of the two models is compared. The scaled deviance is defined as

$$\text{deviance}/\phi$$

where ϕ is the scale parameter, which is fixed at 1 in both models because of the assumption of Poisson error, and the deviance is *the value of the likelihood ratio test statistics for the fitted model compared to the saturated model.* That is

$$\text{deviance} = -2\log\{L(\text{fitted model})/L(\text{saturated model})\}$$

The scaled deviance is helpful in assessing the goodness of fit of a given model. In our case, the log-linear model has larger a scaled deviance compared to that from the change-point model shown in Table 2. This suggests that the change-point model explains the data better, although with one extra parameter for the change-point. Also, the log-linear model can not provide information on the maximum safe dose limit because it assumes a smooth increasing dose-response structure.

Thirdly, from model 3.8, the change-point model with log link, the profile log likelihood plot for the change-point is given in Figure 9, where the maximum peak is at 13 canisters/year. However, at this dose level the model shows that the relative risk ($\exp(\beta_{11})$) is 4.9 (3.7-6.6) before the change-point and reduces to 1.3 (1.1-1.5) after. The reason for such a model to pick up quite a different change-

point, after which the RR is decreasing, is because it assumes a change-point in relative risks (by using the log transformation) rather than excessive risks. From the scatter plot of log transformed observed outcome rate vs. dose level (Figure 10), we can see that the relative risk (i.e. the slope of the smoothing line) decreases slightly rather than increases after the change-point. In fact, natural logarithm transforms a gradual-then-rapid increasing response rate pattern to a straight linear one. Thus, model 3.8 with the log link masks the possible change-point on the real scale, which results in picking up a different change-point, which is not the one we are interested in.

Since only model 3.4, the change-point model with identity link, can provide the maximum safe dose limit we are looking for, we focused our analysis mainly on this model.

The bias and variance of the MLE for the change-point in model 3.4 obtained by the profile likelihood method were evaluated in the simulation study. Simulated change-point data were generated according to the dose-response structure described in the simulation section 3.2.4 assuming the true change-point being at 21 canisters/year (i.e. the MLE for τ from the observed data). The scatter plot of one randomly simulated data set is shown in Figure11.

As an example, the change-point model with identity link (model 3.4) was fitted to one simulated data set using seed=10 in SAS. The resulting profile likelihood plot is shown in Figure 12, which gives the MLE for change-point at 23 canisters/year. Again as a comparison, the same data was also fitted by model

3.8, the change-point model with log link. The corresponding profile likelihood plot (Figure 12) has the peak at 71 canisters/year, taken as the MLE for the change-point, much larger than the true change-point, 21 canisters/year. To see the effect of the transformation, Figure 13 shows the scatter plot of log transformed rate vs. dose level. Again it shows that after such transformation the true underlying change-point rate structure on the real scale is masked, resulting in a change-point quite different from the true one.

100 such change-point data were simulated. Each of them was then fitted by model 3.4 to give the distribution of the 100 MLEs for the change-point. The profile log likelihood plot is based on 71 data points, each corresponding to one likelihood value at one possible change-point from the range of 2 to 72 canisters/year (our searching range). However, since the identity link was used, at some possible change-points, model 3.4 could not be fitted due to the expected number of events being negative. In fact, for some simulated data, model 3.4 could not be fitted at more than 14 possible change-points. The corresponding profile likelihood plot had an irregular shape with many discontinuities due to more than 20% missing data points (14/71). Since the resulting MLE for the change-point was unreliable, such simulated data was then discarded. Therefore a selection criterion was made so that a profile likelihood plot obtained from a simulated data has to be based on at least 57 (71-14) data-points, i.e. it should not have more than 20% of data points missing. Among the 100 simulated data, 93 met the selection criteria when fitted by model 3.4. The distribution of these 93 MLEs for the change-point is shown in Figure 14, which has a range of 16-33

canisters/year with unsymmetrical shape slightly skewed to the left. It has a small standard deviation (STD=2.2) and both its mode and mean are at 21 canisters/year, equal to the true value of the change-point. This suggests that both bias and variance of the estimated change-point by the profile likelihood method are low.

4.3 Confidence interval estimation for the change-point in model 3.4:

Table 3 gives the distribution of both parametric and non-parametric bootstrap estimates of the change-point in model 3.4 and the corresponding 95-percentile intervals.

Among the 1500 non-parametric bootstrap samples, 1492 of them were used according to the selection criterion when fitted by model 3.4. The distribution of the MLEs for the change-point from these bootstrap samples is shown in Figure 15 with mean and median 25.3 and 22 respectively. It has a wide standard deviation (s.d. = 17.5) and is skewed to the left with nearly 25% falling below 10 canisters/year. The 95% percentile confidence interval based on this distribution is (2, 64). As a comparison, the confidence interval based on the normal approximation for the bootstrap estimates is (-9, 60). However, due to the fact that the change-point must be positive, the normal-theory based confidence interval has to be truncated to (0, 60).

Using the parametric bootstrap approach, 1485 bootstrap samples, when fitted to model 3.4, met the selection criterion. The distribution of the MLEs for the change-point from these bootstrap samples is shown in Figure 16, which is

less skewed to the left compared to that from the non-parametric approach. It also has a large standard deviation (s.d. = 17.7), which gives a wider bootstrap 95-percentile interval, (2, 71). And the corresponding truncated confidence interval based on normal approximation is (0, 60).

From the profile log likelihood plot (Figure 7), the difference between $l(\tau_{MLE})$ and $l(\tau_i)$, i.e. $l(\tau_{MLE})-l(\tau_i) = \log(L(\tau_{MLE})/L(\tau_i))$, is smaller than $0.5\chi^2_{1,0.05} = 1.92$, for all i= 2, 3, ..., 72. This indicates that the likelihood ratio test-based confidence interval contains the restricted interval (2, 72), the searching range. Therefore, the truncated confidence interval based on likelihood ratio test is (2, 72).

Finally, the non-parametric bootstrap 95-percentile interval estimation of the change-point in model 3.4 was evaluated. 1500 non-parametric bootstrap samples were generated from one simulated data with 21 canisters/year being the true change-point (seed=10 in SAS). Only 1371 of them were selected according to the missing value criterion, and the distribution of these bootstrap estimates of the change-point is shown in Figures 17. The 95-percentile confidence interval is (17, 26). The parameter estimates of the change point and the βs for this simulated data are given Table 4.

5. Conclusion

5.1 Discussion

Beta-2 agonists are by far the most useful bronchodilators for treating asthma, and are most effective when inhaled.[123] During the 70's and 80's, MDI-generated salbutamol and fenoterol had been the two most commonly used short-acting beta-2 agonists for rapid symptomatic relief. When inhaled at recommended dosages, these agents have very few side effects.

However, since the New Zealand asthma epidemic in mid 1970s, concerns were raised about the use of excessive doses of these short acting inhaled beta-2 agonists, especially fenoterol. Three case control studies found that high doses of fenoterol were associated with increased risks of death from asthma. Similar results were also found in Canada[70,71] from SAEP when investigating the association of excessive use of inhaled beta-2 agonists and increased rate of fatal or near fatal asthma attack. The dose-response curve using the combined doses of MDI fenoterol and salbutamol clearly demonstrated that the excessive rate of fatal asthma attack increased dramatically at high levels, especially when more than one canister was used per month.

In this thesis, we used the Saskatchewan asthmatic cohort to fit a change-point dose-response model using identity link. The MLE for the change-point parameter is at 21 canister/year (on average approximately equals to 1.8 canisters per month), beyond which the excessive rate of fatal asthma attack is almost doubled. However the maximum recommended dose for fenoterol in MDI is two

puffs (400ug) for up to 4 times a day.[124] For a MDI canister containing 200 puffs, this recommended dose limit is equivalent to 1.2 canisters of fenoterol (200ug/puff) per month. The estimated change-point is 50% larger. There are several possible explanations for this:

1) In our analysis, fenoterol and salbutamol were combined without taking into account the dose difference between the two MDI-generated beta-2 agonists. It has been shown on an equal weight basis, fenoterol could be considered to be equivalent to salbutamol in both its bronchodilator effect and cardiac side effects. Most patients are advised to take two puffs of their MDI-generated beta-2 agonists at a time for asthma control or symptom relief. However, one puff of fenoterol contains 200ug active substance whereas one puff of salbutamol delivers only 100ug. This dose difference is the most plausible explanation for why, given in the same number of puffs, fenoterol in general has more prolonged bronchodilating effects as well as more profound cardiac side effects compared to salbutamol.[125] However, in the analysis, canisters of fenoterol and salbutamol combined were considered to represent 2 canisters of MDI-generated beta-2 agonists, rather than 1.5 canisters of fenoterol (in other words, one canister of combined beta-2 agonists should equal to 0.75 canister in fenoterol). Thus after such unit change, 1.8 canisters of combined beta-2 agonists is equivalent to 1.35 (1.8x0.75) canisters of fenoterol, slightly but still larger than the recommended dose level for fenoterol.

2) The maximum recommended dose for fenoterol was obtained under well-controlled experimental conditions where the adverse outcome of interest could not be life threatening and drug overdose was unlikely to occur. However, in this retrospective cohort analysis, the outcome of interest is a fatal or near fatal asthma attack. Such severe rare events can be investigated with the large observational data where drug-overdosing events are likely to occur. Since the adverse outcomes are different, one being more severe than the other, the estimated change-point where the rate of fatal or near fatal asthma attack increases dramatically should and must be larger than the maximum recommended dose level.

3) Due to the limitations of our cohort data, other risk factors, markers for asthma severity in particular, were not controlled in our change-point model. Excessive use of these medications is very likely to be associated with patient's asthma severity. Their confounding effects may change the estimated location of the change-point. However, in future studies, it would be interesting to examine whether their confounding effects only affect the magnitudes of rate differences before and after the change-point or the location of change-point itself.

The relationship between the rate of fatal or near fatal asthma attack and excessive use of MDI-generated beta-2 agonists can be modelled in two ways: either by a change-point structure (model 3.4), or by an exponential structure without a change-point (model 3.6). Both structures can represent a dose-response relationship with slow increase at lower doses followed by a dramatic increase at higher dose levels. The simple log-linear model (model 3.6) can only provide the

single relative risk estimation. On the other hand, the change-point parameter and the excessive rates before and after the change-point estimated from model 3.4 even without controlling possible confounders still provide clinical useful information with respect to the safe use of MDI-generated beta-2 agonists. A model with a quadratic or higher terms could also be used to model such a dose-response curve. However such a model can not provide information on the safe dose limit for these medications.

From the change-point model (model 3.4), the excessive risk of fatal asthma attack increases almost two-fold if the monthly use of beta-2 agonists is over 1.8 canisters. Even though the parameter estimates can be confounded by other risk factors, the assumption of a change-point existing for the excessive rate of fatal asthma attacks seems to be reasonable based on the following account. Due to direct stimulation of cardiac beta-2 receptors, tachycardia was observed with higher doses of fenoterol and salbutamol.[126] Furthermore, an increased incidence of cardiac dysrhythmias had occurred at higher doses of beta-2 agonists administered by nebulizer.[127] Furthermore, excessive use of these medications is an indication of chronic use. It could lead to receptor desensitization,[128,129] which in turn, could develop tolerance to the bronchodilating effect in asthmatic subjects.[130,131] Finally, chronic use of beta-2 agonists could cause deterioration in asthma by allowing more allergen to be deposited in the airway and by inhibiting the mast cells' anti-inflammatory effects.[57]

One interesting thing to notice from the previous section is that to model such a dose-response structure with a change-point, the identity link should be used instead of the canonical log link for Poisson distribution. Otherwise, the underlying dose-response relationship is completely transformed on the log scale and the resulting change-point is not the one we are interested. Thus, such two-phase regression model can only be applied in cohort study where the absolute risk or rate can be directly modelled by Poisson regression with identity link whereas in case control study logit is the only link available to model the odds.

The maximum likelihood estimate for the change-point in model 3.4 was obtained by maximizing the profile likelihood. The simulation study shows that for 93 simulated change-point data, the corresponding maximum profile likelihood estimates are distributed with relatively small variation around its mode, equal to the value of the true underlying change-point. This suggests that the profile likelihood method is reliable in providing the point estimate of the change-point. Its slightly unsymmetrical shape, skewed to the left, may be likely due to the small sample size (n=93). However, our ability to use larger samples was limited by the time-consuming nature of the simulation. Further simulation studies can be performed to investigate the confounding effects of other risk factors, especially the severity of asthma and the use of other asthma medications. To do so, data can be simulated according to a structure where there are constant excessive rates (i.e. the constant slopes) for these factors, and a change-point for the excessive rate associated with the drug exposure. Such simulated data can

then be fitted model 3.4 with and without these covariates to evaluate their confounding effects.

One major drawback of the change-point model using the identity link (model 3.4) is that at certain possible change-points, model 3.4 can not be fitted. For a given change-point, using the identity link may give negative predicted numbers of events from the linear predictor during one iteration of the re-weighted least square procedure. Since positive predicted mean values are required for Poisson error, the iteration process stopped at that cycle and the model could not be fitted at this change-point. As a result, it led to a non-smooth discontinuous profile likelihood plot. In fact, for some simulated data, the model could not be fitted at more than 20% of all possible change-points, resulting in a profile likelihood plot with many of discontinuities, which could not be used. Thus it may give a biased result, since not all the simulated data were used. To overcome such problem, a SAS macro can be written to prevent a negative predicted mean value within each iteration.

The 95-percentile intervals for the change-point in model 3.4 based on bootstrap method are consistent with those based on normal approximation and likelihood ratio (LR) test. All of the intervals have a wide range, which is very close to the searching limits (2-72 canisters/year) where the change-point is believed to lie. However, both the normal approximation and the LR test based methods require asymptotic normality, and they are not subjected to the searching limits, resulting in truncated intervals. Furthermore, more precise bootstrap intervals such as BCa and ABC methods could be tried in future studies.

In general, when the bootstrap samples are generated according to a specified parametric structure rather than the empirical distribution, the distribution of the bootstrap estimates is more symmetric and less skewed. This is the case shown in our results (see Figure 15 and 16). However, the percentile interval based on parametric bootstrap samples is wider. There are more parametric bootstrap samples that could not be fitted properly than the non-parametric ones. This may be either due to chance or due to a greater possibility of having negative predicted mean values when fitting a more structured parametric model. According to our selection criterion, these samples were not used in constructing the corresponding percentile intervals.

Even though the bootstrap method gives a better confidence interval, it still has a wide range very close to the searching limits. This is most likely due to the small number of cases especially at the higher dose levels. In fact based on one simulated data, it has shown that if the number of cases at each dose level is at least 5, for the change-point in model 3.4 the non-parametric bootstrap method gave a pretty narrower percentile interval (17-25 canisters/year) containing the true change-point (21 canisters/year). In future studies, the same simulation can be done for a large number of simulated data, say 100, to see the proportion of these percentile intervals, each obtained based on 1000 bootstrap replications, containing the true change-point. Such a computing intensive simulation study will evaluate the coverage property of the percentile interval for the change-point, and require a more efficient program written in a more flexible language such as S-plus or C. Finally, other methods based on the Bayesian approach on a

modification of the 'inverted likelihood ratio' confidence intervals suggested by Worsley[80] could be tried, and might give useful results.

5.2 Conclusion

The change-point problem in statistics has existed for more than 40 years, and has been found in various areas of statistics. One of the interesting applications of the change-point problem is in linear regression where a change in β parameter for a given covariate occurs at an unknown point. The problem has been studied extensively since the 1950s. Unfortunately, because of the discontinuity and stochastic nature of the likelihood function, inference on the change-point parameter is complicated and hinders its application in real life. However, numerical methods using the profile likelihood and the bootstrap make such a new modelling approach readily applicable to interesting problems encountered in epidemiologic studies.

Premarketing dosing studies have limitations in size, time, money and more importantly, ethics. Therefore, information on effective drug dose with minimum adverse risk can not be adequately addressed in clinical trials. Post-marketing pharmacoepidemiologic studies, relatively free of these limitations, can provide this additional information which is crucial to the public health. Asthma is a common disease in developed countries and the most widely used medications for fast symptom relief are short-acting MDI beta-2 agonists. However, epidemiologic evidence suggested that excessive use of such medications might cause a severe adverse effect, i.e. asthma death or near death.

This provides a typical situation where a post-marketing epidemiologic study using a large-scale cohort is required to establish the safe dose limit.

Using the Saskatchewan asthmatic cohort database, a dose-response curve for short-acting MDI beta-2 agonists (fenoterol and salbutamol) was fitted assuming a change-point beyond which the excessive rate of fatal or near fatal asthma attack increases much faster. Compared to the log-linear model, the two-phase regression model with identity link gives a better fit to the data with a smaller deviance. More importantly, it directly models the absolute rate of the outcome under a more plausible assumption with respect to such asthma medications, and provides information on the change-point parameter that is clinical relevant.

The maximum likelihood estimate of the change-point in such model was obtained by maximizing the profile likelihood, which gave a relatively unbiased estimate shown in the simulation study. The estimated change-point is at 1.8 canisters/month, after which the excessive risk is two times higher than before. The point estimate is slightly higher than the maximum recommended dose, which is expected because of the dose difference of the two drugs being combined and the more severe outcome (fatal or near fatal asthma attacks) being looked at. The fact that our estimated change-point, which can be regarded as the maximum safe dose limit, is larger than the maximum recommended dose for these medications indicates that the maximum recommended dose provides an extra safet margin and should be followed in real medical practice.

There are several risk factors, which are considered to be markers for asthma severity. Due to the limitations in our cohort data, however, their effects as confounders and/or effect modifiers on the parameter estimates of the change-point model were not addressed here, and should be investigated in the future. Despite not controlling these factors, the estimated maximum safe dose limit (the change-point) still provides useful information for these beta-2 agonists as a warning sign to require more medical attention or to seek alternative treatment.

The confidence intervals for the change-point given by the both parametric and non-parametric bootstraps are consistent with the ones based on normal theory or likelihood ratio test and require few assumptions. All the confidence intervals are wide, which is very likely due to the small number of cases in the entire cohort (129 cases). From one simulation study with ideal sample size, the 95-percentile interval obtained by the non-parametric bootstrap is quite narrow and contains the true change-point. However, its real coverage probability needs to be evaluated in future studies.

In conclusion, the profile likelihood and the bootstrap methods are useful in making inference on the change-point in two-phase regression model. For post-marketing drug safety study, using cohort analysis with identity link, the change-point in such model can be regarded as the maximum safe dose limit. This dose information can not only be used to confirm the recommended dose level established in clinical trials but also provide a drug safety guideline in real practice. When the outcome of interest is not so rare, inference for the change-point based on the two methods is reliable.

6. Bibliography

1. Vocci F, Farber T. Extrapolation of animal toxicity to man. *Reg. Toxicol. Pharmacol.* 1988; 8:389-398.
2. Lumley CE. Clinical toxicity: Could it have been predicted? Pre-marketing experience In: Lumley CE, Walker SR. eds. *Animal Toxicity Studies: Their relevance for Men.* Lancaster UK: Quay; 1990: 49-56.
3. .Heywood R. Clinical toxicity: Could it have been predicted? Post-marketing experience In: Lumley CE, Walker SR. eds. *Animal Toxicity Studies: Their relevance for Men.* Lancaster UK: Quay; 1990: 57-70.
4. Schach Von Wittenau M. Strength and weakness of long term bioassay. *Reg. Toxicol. Pharmacol.* 1987; 7:113-119.
5. Storer B. Design and analysis of phase I clinical trials. *Biometrics* 1989; 45:925-937.
6. Gehan E., Schneiderman M. Historical and methodological developments in clinical trials at the National Cancer Institute. *Statistics in medicine* 1990; 9:871-880.
7. Geller N. Design of phase I and II clinical trials in cancer: a statistician view. *Cancer investigation* 1984; 2:483-491.
8. Strom B. ed. *Pharmacoepidemiology* 2nd. edn. Chichester, England: John Wiley & Sons; 1994: 1-65.
9. Geiling EMK, Cannon PR. Pathogenic effects of elixir of sulfanilimide (diethylene glycol) poisoning. *JAMA* 1938; 111:919-926.

10. Wallerstein RO, Condit PK, Kasper CK, *et al.* Statewide study of chloramphenicol therapy and fatal aplastic anemia. *JAMA* 1969; 208:2045-50.

11. Cluff LE, Thornton GF, Seidl LG. Studies on the epidemiology of adverse drug reactions. I. Methods of surveillance. *JAMA* 1964; 188:976-83.

12. Miller RR, Greenblatt DJ. *Drug effects in hospitalized patients*. New York: Wiley; 1976.

13. Lenz W. Malfunctions caused by drugs in pregnancy. *Am J Dis Child* 1966; 112:99-106.

14. Strom B. What is pharmacoepidemiology? In: Strom B. ed. *Pharmacoepidemiology* 2nd. edn. Chichester, England: John Wiley & Sons; 1994: 6.

15. Mattison N, Richard BW. Postapproval research requested by FDA at the time of NCE approval, 1907-1984. *Drug Info J* 1987; 21: 309-329.

16. Royal College of General Practitioners. *Oral Contraceptives and Health.* London: Pitman Publishing; 1974.

17. Strom BL, Stolley PD. Vascular and cardiac risks of steroid contraception. In: Sciarra JJ, ed., *Gynecology and obstetrics*. Vol. 6 New York: Harper and Row; 1989:1-17.

18. Vessey MP, Doll R, Peto R. *et al.* A long term follow-up study of women using different methods of contraception — an interim report. *J Biosoc Sci* 1976; 8:373-427.

19. Sheffer AL. Guidelines for the diagnosis and management of asthma. National Heart, Lung, Blood Institute National Asthma Education Program Expert Panel Report. *J Allergy Clin Immunol* 1991; 88:425-534.

20. Macdonald JB, Cole TJ. The flow-volume loop: reproducibility of air and helium-based tests in normal subjects. *Thorax* 1980; 35:64-69.

21. Balfour-lynn L, Tooley M, Godfrey S. Relationship of exercise-induced asthma and clinical asthma in childhood. *Arch Dis child* 1981; 56:450-454.

22. Jariwalla G. *Asthma.* Lancaster UK: MTP Press Limited; 1988:1.

23. Grossman J. One airway, one disease. Chest(supplement) 1997; 111(2):11s-16s.

24. Sibbald B, Horn MEC, Brain EA, Gregg I. Genetic factors in childhood asthma. *Thorax* 1980; 35:671-674.

25. Edfors-Lubs M-L. Allergy in 7000 twin pairs. *Acta Allergol* 1971; 26:249-85.

26. Bertrand J-M, Riley SP, Popkin J, Coates AL. The long-term pulmonary sequelae of prematurity: the role of familial airway hyperreactivity and the respiratory distress syndrome. *N Engl J Med* 1985; 312:742-745.

27. Bader D, Ramos AD, Lew CD, Platzker ACG, Stabile MW, Keens TG. Childhood sequelae of infant lung disease: exercise and pulmonary function abnormalities after bronchopulmonary dysplasia. *J pediatr* 1987; 110:693-699.

28. Chan KN, Noble-Jamieson CM, Elliman A, Bryan EM, Silverman M. Lun function in children of low birth weights. *Arch Dis Child* 1989; 64:1284-1293.

29. Martin AJ, Landau LI, Phelan PD. Asthma from childhood to age 21: the patient and his disease. *Br Med J* 1982; 284:380-382.

30. Murray AB, Morrison BJ. Passive smoking and the seasonal difference of severity of asthma in children. *Chest* 1988; 94:701-708.

31. Goren AL, Hellmann S. Prevalence of respiratory symptoms and diseases in school children living in a polluted and a low polluted area in Israel. *Environ Res* 1980; 45:28-37.

32. Keeley DJ, Neil P, Gallivan S. Comparison of prevalence of reversible airways obstruction in rural and urban Zimbabwean children. *Thorax* 1991; 46:549-553.

33. Waite DA, Eyles EF, Tonkin SL, O'Donnell TV. Asthma prevalence in Tokelauan children in two environments. *Clin Allergy* 1980; 10:71-75.

34. Weitzman M, Gortmaker S, Sobol A. Racial, social and environmental risk factors for childhood asthma. *Am J Dis Child* 1990; 144:1189-1194.

35. Khot A, Burn R, Evans N, Lenney C, Lenney W. Seasonal variation and time trends in childhood asthma in England and Wales 1975-8. *Br Med J* 1984; 289:235-237.

36. Minor TE, Dick EC, DeMeo AN, Ouellette JJ, Cohen M, Reed CE. Viruses as precipitants of asthmatic attacks in children. *JAMA* 1974; 227:292-298.

37. Calhoun WJ, Swenson CA, Dick EC, Schwartz LB, Lemanske RFJ, Busse WW. Experimental rhinovirus 16 infection potentiates histamine release after antigen brochoprovocation in allergic subjects. *Am Rev Respir Dis* 1991; 144:267-273.

38. Pullan CR, Hey EN. Wheezing, asthma pulmonary dysfunction 10 years after infection with respiratory syncytial virus in infancy. *Br Med J* 1982; 284:1665-9.

39. Clark TJH. Ed. *Steroids in asthma.* Auckland: ADIS Press; 1983.

40. Solis-Cohen S. The use of adrenal substances in the treatment of asthma. *JAMA* 1900; 34:1164-6.

41. Bullowa JGM, Kaplan DM. On the hypodermic use of adrenaline chloride in the treatment of asthma attacks. *Med News* 1902; 83:787.

42. Goodman Ls, Gilman A. *The pharmacological basis of therapeutics*. 2nd ed. New York: Macmillan; 1958: 497.

43. Konzett H. Neues zur Asthmatherapie. *Klin Wochenschr* 1940; 19:1303.

44. Ahlquist RP. A study of adrenotropic receptors. *Am J physiol* 1948; 153:586.

45. Leifer KN, Wittig HJ. The beta-2 sympathomimetic aerosols in the treatment of asthma. *Ann. Allergy*. 1975; 35:69.

46. Brittain RT, Farmer JB, Jack DJ. α-(t-butylamino)-methyl-4-hydroxy-m-xylene-α^1,α^3-diol (AH 3365): a selective β-adrenergic stimulant. *Nature* 1968; 219:862.

47. Harley D, Jack D, Lunts LHT, Ritchie AC. New class of stimulants of β-adrenergic receptors. *Nature* 1968; 219:861.

48. Kennedy MCS, Simpson WT. Human pharmacological and clinical studies on salbutamol: a specific β-adrenergic bronchodilator. *Br J Dis Chest* 1969; 63:165.

49. Freedman BJ. Trial of new bronchodilator, terbutaline in asthma. *Br Med J* 1971; 1 633.

50. Powles ACP. The bronchodilator effect of fenoterol (Berotec). *N Z Med J* 1971; 81:249.

51. Blackhall MI, Dauth M, Mahoney M, O'Donnell SR. Inhalation of fenoterol (Berotec) by asthmatic children: a clinical comparison with salbutamol, orciprenaline and isoprenaline. *Med J Aust* 1976; 2:439.

52. Lands AM, Arnold A, McAuliff JP, Ludena FP, Brown TJ. Differentiation of receptor systems activated by sympathomimetic amines. *Nature* 1967; 214:597.

53. Barnes PJ, Chung KF. Questions about beta-2 agonists in asthma. *Trends Pharmacol Sci* 1992; 13:20-3.

54. Buist AS, Vollmer WM. Reflections on the rise of asthma morbidity and mortality. *JAMA* 1990; 264:1719.

55. Bosco LA, Knapp DE, Gerstman B, Graham CF. Asthma drug therapy trends in the United States 1972 to 1985. *J Allergy Clin Immunol* 1987; 80:398.

56. Sly RM. Increases in deaths from asthma. *Ann Allergy* 1984; 53:20.

57. Page CP. One explanation of the asthma paradox: inhibition of natural anti-inflammatory mechanisms by beta-2 agonists. *Lancet* 1991; 1:717.

58. Stolley PD, Schinnar R. Association between asthma mortality and isoproterenol: a review. *Prev Med* 1978; 7:519.

59. Sears MR, Rea HH, Beaglehole R, Gillies AJD, Holst, P, O'Donnell TV, Rothwell RPG, Sutherland DC. Asthma mortality in New Zealand: a two-year national study. *N Z Med J* 1985; 98:271.

60. Speizer FE, Doll R. A century of asthma death in young people. *Br Med J* 1968; 3:245.

61. Inman WHW, Adelstein AM. Rise and fall of asthma mortality in England and Wales in relation to use of pressurised aerosols. *Lancet* 1969; 2:279.

62. Stolley PD. Asthma mortality: why the United States was spared an epidemic of deaths due to asthma. *Am Rev Respir Dis* 1972; 105:883.

63. Editorial. Asthma deaths: a question answered. *Br Med J* 1972; 4:443.

64. Wilson JD, Sutherland DC, Thomas AC. Has the change to beta agonists combined with oral theophylline increased cases of fatal asthma? *Lancet* 1981; 1:1235.

65. Grant IWB. Asthma in New Zealand. *Br Med J* 1983; 286:374.

66. Keating G. Mitchell EA, Jackson RT, Rea HH. Trends in sale for asthma in New Zealand, Australia and United Kingdom. *Br Med J* 1984; 289:348.

67. Crane J, Pearce N, Flatt A, Burgess C, Jackson R, Kwong T, Ball M, Beasley R. Prescribed fenoterol and death from asthma in New Zealand, 1981-83: case control study. *Lancet* 1989; 1:917.

68. Pearce NE, Grainger J, Atkinson M, Crane J, Burgess C, Culling C, Windom H, Beasley R. Case-control study of prescribed fenoterol and death from asthma in New Zealand, 1977-1981. *Thorax* 1990; 45:170.

69. Grainger J, Woodman K, Pearce, N, Crane J, Burgess C, Keane A, Beasley R. Prescribed fenoterol and death from asthma in New Zealand, 1981-1987: a further case control study. *Thorax* 1991; 46:105.

70. Spitzer WO, Suissa S, Ernst P, et al. The use of beta-agonists and the rate of death and near death from asthma. *N Engl J Med* 1992; 326:503-506.

71. Suissa S, Ernst P, Boivin J, et al. A cohort analysis of excess mortality in asthma due to use of inhaled beta-agonists. *Am J respir crit Care Med* 1994; 149:604-10.

72. Evens R III. Epidemiology and natural history of asthma, allergic rhinitis and atopic dermatitis. In: Middleton E Jr, Reed CE, Ellis EF, et al, eds. *Allergy principles and practice.* 4th ed. St. Louis: Mosby; 1993:1109-1136.

73. Page ES. Continuous inspection schemes. *Biometrika* 1954; 41:100-115.

74. Worsley KJ. The power of likelihood ratio and cumulative sum tests for a change in a binomial probability. *Biometrika* 1983; 70:455-465.

75. Page ES. A test for a change in a parameter occurring at an unknown point. *Biometrika* 1955; 42:523-526.

76. Page ES. On problem in which a change in a parameter occurs at an unknown point. *Biometrika* 1957; 44:248-252.

77. Hawkins DM. Testing a sequence of observations for a shift in location. *J Amer statist Assoc* 1977; 72:180-186.

78. Worsley KJ. On the likelihood ratio test for a shift in location of normal populations. *J Amer statist Assoc* 1979; 74:365-367.

79. Worsley KJ. An improved Boferroni inequality and applications. *Biometrika* 1982; 69:297-302.

80. Worsley KJ. Confidence regions and tests for a change-point in a sequence of exponential family random variables. *Biometrika* 1986; 72:91-104.

81. Siegmund D. Boundary crossing probabilities and statistical applications. *Ann Statist* 1986; 14:361-404.

82. Hinkley DV. Inference about the change-point in a sequence of random variables. *Biometrika* 1970; 57:1-16.

83. Hinkley DV, Hinkley EA. Inference about the change-point in a sequence of binomial variables. *Biometrika* 1970; 57:477-488.

84. Siegmund D. Confidence sets in change-point problems. *Intern Statist Rev* 1988; 56:31-48.

85. Maguire BA, Pearson ES, Wynn AHA. The time intervals between industrial accidents. *Biometrika* 1952; 38; 168-180.

86. Raftery AE, Akman AE. Bayesian analysis of a Poisson process with a change. *Biometrika* 1986; 73:85-89.

87. Carlin BP, Gelfand AE, Smith AFM. Hierarchical Bayesian analysis of change-point problems. *Appl Statist* 1992; 41:389-405.

88. Esterby SR, El-Shaarawi A. Inference about the point of change in a regression model. *Appl statist* 1981; 30:277-285.

89. Smith AFM, and Cook DG. Straight lines with a change-point: a Bayesian analysis of some renal transplant data. *Appl statist* 1980; 29:180-189.

90. Seber GA, Wild CJ. *Non-linear regression*. New York: John Wiley & sons; 1989.

91. Sprent P. Some hypotheses concerning two-phase regression lines. *Biometrics* 1961; 17:634-645.

92. Hudson DJ. Fitting the segmented curves whose join points have to be estimated. *J Amer Statist Assoc* 1966; 61:1097-1129.

93. Hinkley DV. Inference about the intersection in two-phase regression. *Biometrika* 1969; 56:495-504.

94. Hinkley DV. Inference in two phase-regression. *J Amer statist Assoc* 1971; 66:736-743

95. Feder PI. On asymptotic distribution theory in the segmented regression problems: identified case. *Ann Statist* 1975; 3:49-83.

96. Feder PI. The log likelihood ratio in segmented regression. *Ann Statist* 1975; 3:84-97.

97. Stasinopoulos DM. Using GLIM to fit split-line curves. *GLIM Newsletter* 1988; 17:24-31.

98. Stasinopoulos DM, Rigby RA. GLIM macros to fit piecewise polynomials with one break point. *GLIM Newsletter* 1990; 20:18-24.

99. Healy MJR. *GlIM: an introduction*. Oxford: Clarendon Press; 1988.

100. Press WH, Teukolsky SA, Vetterling WT, Flannery BP. *Numerical recipes in C: the art of scientific computing*. 2nd ed. Cambridge: University Press; 1992.

101. Ulm K. A statistical method for assessing a threshold in epidemiological studies. *Statist In Med* 1991; 10:341-349.

102. Stasinopoulos DM. Rigby RA. Detecting break points in generalised linear models. *Comp statist & data analysis* 1992; 13:461-471.

103. Efron B, Tibshirani RJ. *An introduction to the bootstrap*. New York: Chapman & Hall; 1993.

104. Chernoff H, Zacks S. Estimating the current mean of a normal distribution which is subjected to the changes in time. *Ann Math Statist* 1964; 35:999-1018.

105. Gardner LA. Jr. On detecting changes in the mean of normal variates. *Ann Math Statist* 1969; 40:116-126.

106. MacNeill IB. Tests for change of parameters at unknown times and distributions of some related functionals on Brownian motion. *Ann statist* 1974; 2:950-962.

107. Sen A, Srivastava MS. Some one-sided tests for change in level. *Technometrics* 1975; 17:61-64.

108. Jandhyala VK, MacNeill IB. Residual partial sum limit process for regression models with applications to detecting parameter changes at unknown times. *Stoch. Proc. Appl.* 1989; 33:309-323.

109. Bacon DW, Watts DG. Estimating the transition between two intersecting straight lines. *Biometrika* 1971; 58:525-534.

110. Ferreira PE. A Bayesian analysis of a switching regression model: a known number of regimes. *J Amer Statist Assoc* 1975; 70:370-374.

111. Chin Choy J, Broemeling LD. Some Bayesian inferences for a changing linear model. *Technometrics* 1980; 22:71-78.

112. Moen DH, Salazar D, Broemeling LD. Structural changes in multivariate regression models. *Communs Statist A* 1985; 14:1757-1768.

113. Jandhyala VK, Minogue CD. Distribution of Bayes-type change-point statistics under polynomial regression. *J Statist Planning Inference* 1993; 37:271-290.

114. Martin AT. *Tools for statistical inference.* New York: Springer-Verlag; 1996: Chapter 6.

115. Suissa S, Blais L, Ernst P. Patterns of increasing beta-agonists use and the risk of fatal or near fatal asthma. *Eur Respir J* 1994; 7:1602-1609.

116. Pettitt AN. A non-parametric approach to the change-point problem. *Appl Statist* 1989; 28:126-135.

117. Lombard F. Asymptotic distributions of rank statistics in the change-point problems. *South African Statist J* 1983; 17:83-105.

118. Sen PK. *Theory and applications of sequential non-parametrics*. Philadelphia: Society for industrial and applied mathematics; 1984.

119. Wolfe DA, Schechtman E. Non-parametric statistical procedures for the change-point problem. *J statist Planning Inference* 1984; 9:389-396.

120. Lombard F. Some recent developments in the analysis of change-point data. *South African Statist J* 1989; 23:1-21.

121. Hastie TJ, Tibshirani RJ. *Generalized additive models*. New York: Chapman & Hall; 1990.

122. Aitkin M, Anderson D, Francis B, Hinde J. *Statistical modelling in GLIM*. Oxford: Clarendon Press; 1992.

123. Nelson HS. Beta-adrenergic bronchodilators. *N Engl J Med* 1995; 333:499-506.

124. Heel RC, Brogden RN, Speight TM, Avery GS. Fenoterol: a review of its pharmacological properties and therapeutic efficacy in asthma. *Drugs* 1978; 15:3-32.

125. Newhouse MT, Dolovich MB, Kazim F. Dose-effect relationship of the β-agonists fenoterol and salbutamol in patients with asthma. *Chest* 1994; 105:1738-1742.

126. Wong CS, Pavord ID, Williams J, Britton JR, Tattersfield AE. Bronchodilator, cardiovascular, and hypokalaemic effects of fenoterol, salbutamol, terbutaline in asthma. *Lancet* 1990; 336:1396-99.

127. Higgins RM, Cookson WOCM, Lane DJ, John SM, McCarthy GL, McCarthy ST. Cardiac arrhythmias caused by nebulised beta-agonist therapy. *Lancet* 1987; 2:863.

128. Mickey J, Tate R, Lefkowitz RJ. Subsensitivity of adenylate cyclase and decreased β-adrenergic receptor binding after chronic exposure to isoproterenol *in vitro*. *J Biol Chem* 1975; 250:5727.

129. Samuelson WM, Davies AO. Hydrocortisone-induced reversal of beta-adrenergic receptor uncoupling. *Am Rev Respir Dis* 1984; 130:1023.

130. Conolly ME, Tashkin D, Hui KKP, Wolfe RN. Selective subsensitisation of beta-adrenergic receptors in central airways of asthmatics and normal subjects during long term therapy with inhaled salbutamol. *J Allergy Clin Immunol* 1982; 70:423.

131. Repsher LH, Anderson JA, Bush RK, Falliers CJ, Kass I, Kemp JP, Reed C. et al. Assessment of tachyphylaxis following prolonged therapy of asthma with inhaled albuterol aerosol. *Chest* 1984; 85:34.

Table 1.

Rate of fatal or near fatal asthma attack by ordinal classification of exposure dose

Combined MDI-delivered beta-2 agonists (Canisters/year)	Number of fatal or near fatal asthma attack	Person-months	Rate of fatal or near fatal asthma attack (per10,000 asthmatics per year)
0	4	189,064	2.54
1-6	18	189,719	11.39
7-12	20	82,341	29.15
13-18	23	48,158	57.32
19-24	14	26,900	62.45
25-30	16	15,552	123.46
31-36	6	9,144	78.74
37-42	7	5,298	158.55
43-48	10	3,173	378,59
49-60	3	3,020	119.21
61-72	5	931	644.47
73-100	2	630	380.95
>100	1	178	674.16

Table 2.

Parameter estimates: analysis of the asthmatic cohort data

Model	Link function	Scale deviance	Parameters		Estimates
Change-point model	Identity (Model 3.4)	97.51	Change-point		21 canisters/year
			Excessive rates for one additional canister /month	Before the change-point β_{11}	3.7 (2.7, 4.7) per 1,000 asthmatics pre year
				After the change-point β_{21}	7.0 (3.5, 10.4) per 1,000 asthmatics pre year
Change-point model	Log (Model 3.8)	111.64	Change-point		13 canisters/year
			Relative risks for one additional canister /month	Before the change-point $\text{Exp}(\beta_{11})$	4.9 (3.7, 6.6)
				After the change-point $\text{Exp}(\beta_{21})$	1.3 (1.1, 1.5)
Log-linear model	Log (Model 3.6)	176.16	Relative risks for one additional canister /month		1.6 (1.5, 1.7)

Table 3.

Bootstrap confidence intervals (unit: canisters/year)

			Non-parametric bootstrap		Parametric bootstrap	
Applying missing value criteria (i.e. the numbers of missing likelihood less than 14)			No	Yes	No	Yes
The number of bootstrap replications			1500	1492	1500	1485
Distributions of the bootstrap estimates	Mean		25.2	25.3	24.9	25.0
	Standard deviation		17.5	17.5	17.7	17.7
	Quantiles	Minimum	2	2	2	2
		2.5%	2	2	2	2
		25% (Q1)	9	9	13	14
		Median	22	22	21	21
		75% (Q3)	35	35	32	32
		97.5%	64	64	71	71
		Maximum	71	71	71	71
95% percentile intervals (2.5%, 97.5%)			(2, 64)	(2, 64)	(2, 71)	(2, 71)
Normal based intervals (mean±1.96*std.)			(-9, 60)	(-9, 60)	(-10, 60)	(-10, 60)

Table 4

Parameter estimates: simulated data*

Model	Parameters		True values	Estimated values (95% confidence intervals)
Change-point model with identity link (model 3.4)	Change-point (canisters /year)		21	**23 (17,26)
	Base-line rate β_{10} (per 1,000 asthmatics per year)		4	4.86 (3.18,6.52)
	Excessive rates by one additional canister /month (per 1,000 asthmatics per year)	Before the change-point β_{11}	0	0.08 (-1.31,1.47)
		After the change-point β_{21}	8	9.76 (8.85,10.66)

* The data is simulated according to model 3.9 with seed=10 in SAS.
** 95- percentile interval based on 1371 reliable non-parametric bootstrap samples

FIGURE 1

ASTHMA DEATHS OF PERSONS AGED 5-34 COMPARED WITH SALES AND PRESCRIPTIONS OF ASTHMA PREPARATIONS IN ENGLAND AND WALES FOR 1959-1968.

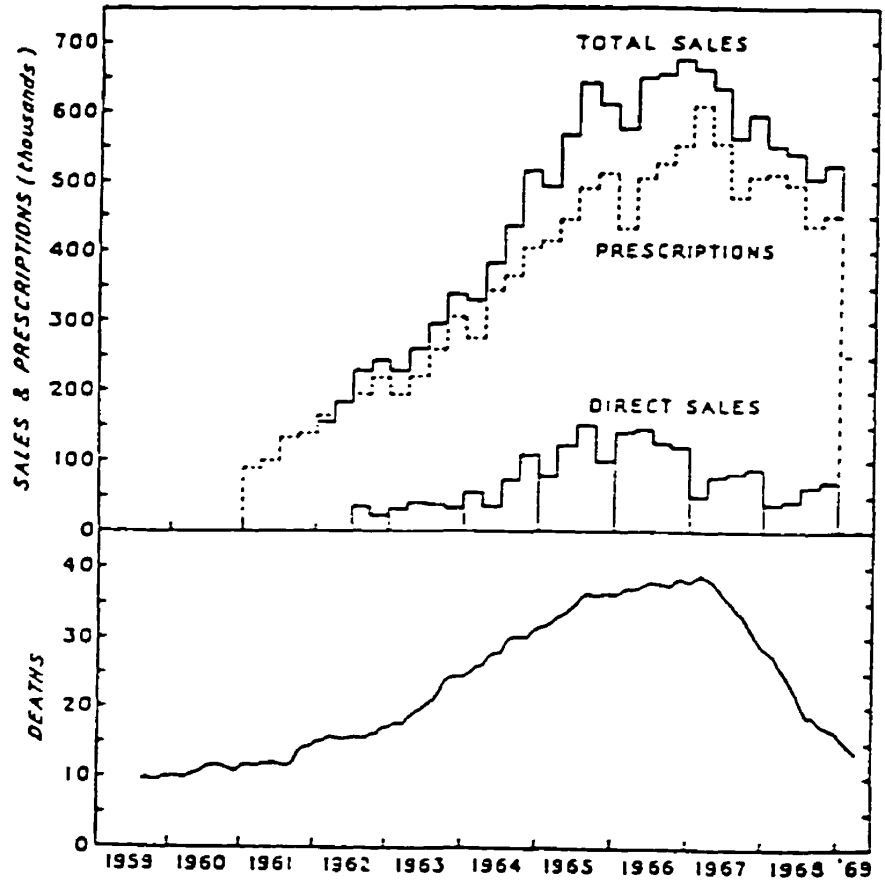

FIGURE 2

ASTHMA MORTALITY (5-34YEAR OLDS) AND FENOTEROL MARKET SHARE (%) IN New Zealand 1974-1990

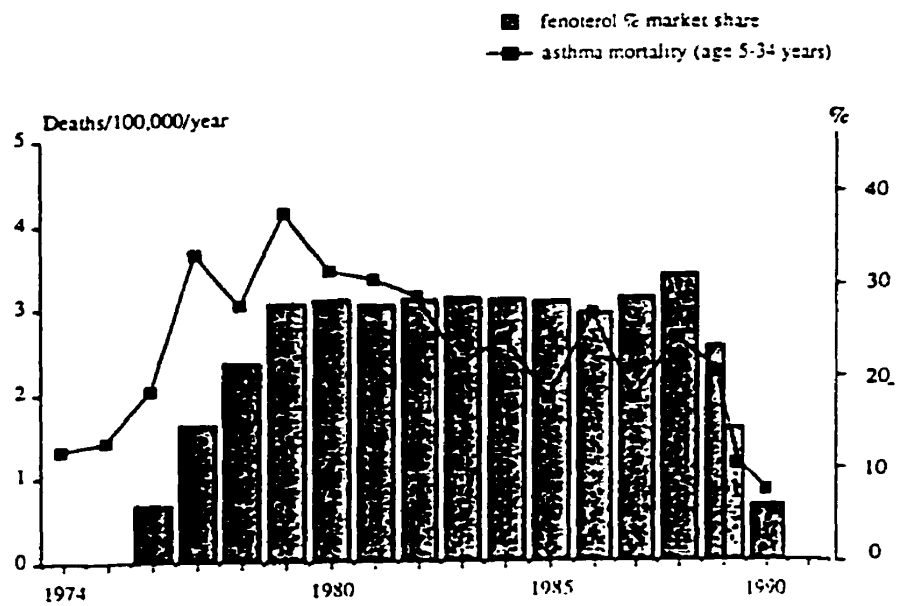

FIGURE 3

PLOT OF POLLEN CONCENTRATION VS.DEPTH OF A LAKE SEDIMENT CORE

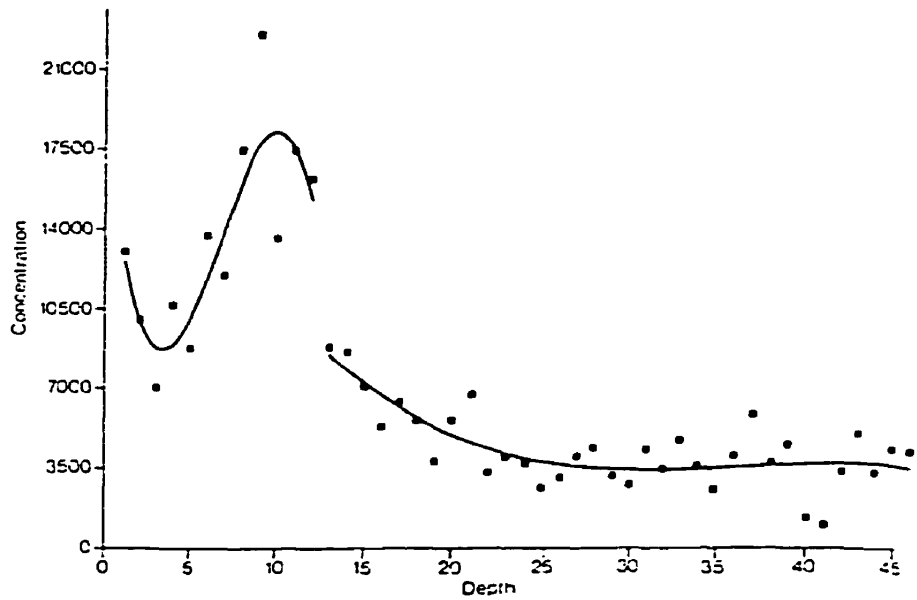

FIGURE 4

PLOT OF (SERUM-CREATININE)$^{-1}$ VS. TIME FOR ONE PATIENT

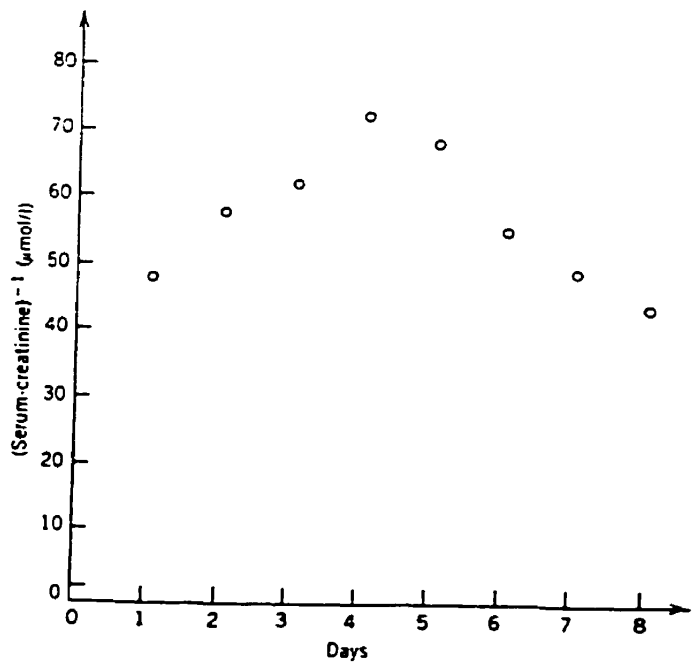

FIGURE 5.

OBSERVED DOSE-RESPONSE PLOT

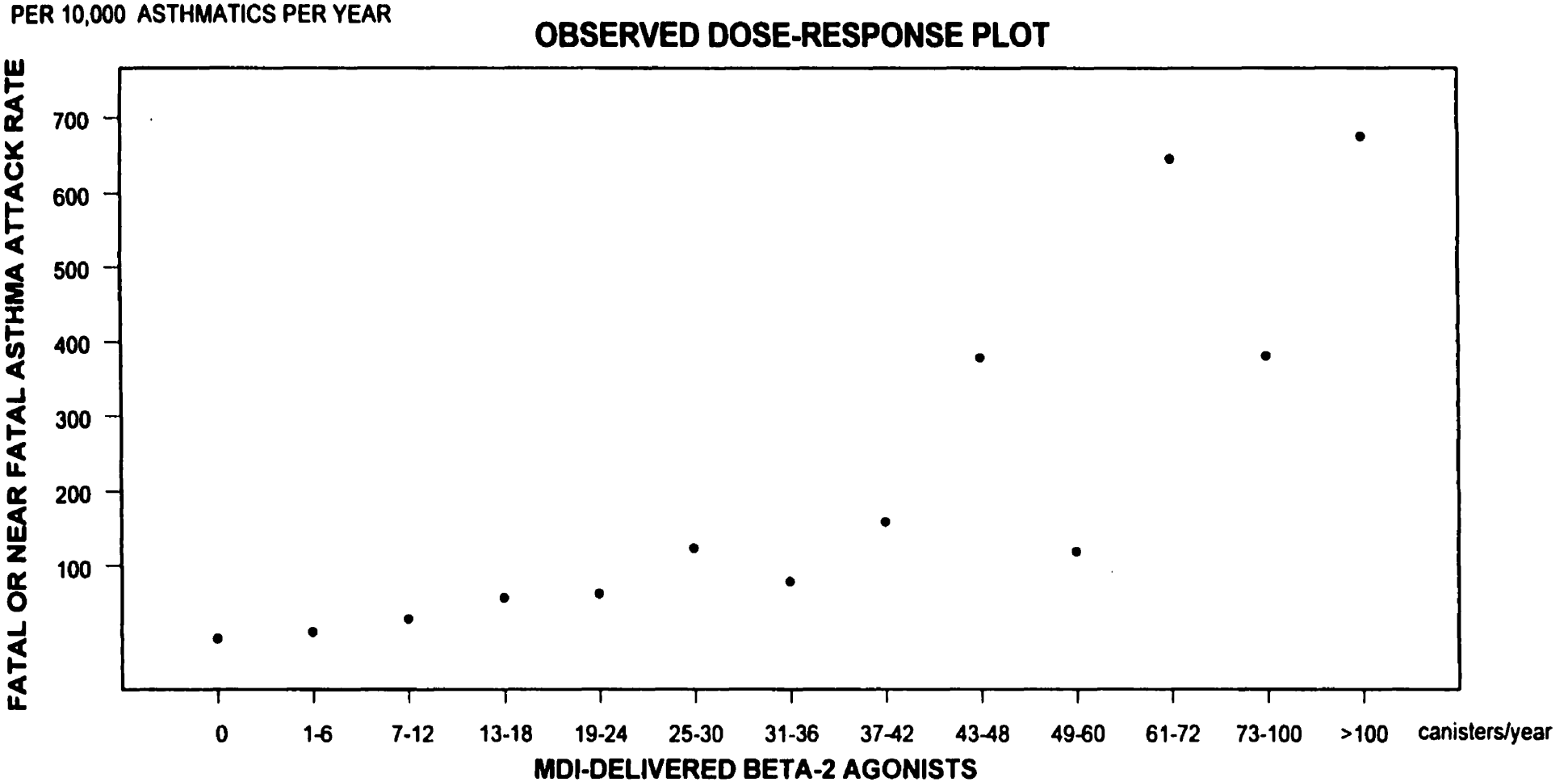

FIGURE 6.

PROFILE LOG LIKELIHOOD PLOT OF CHANGE-POINT MODEL WITH IDENTITY LINK (MODEL 3.4)

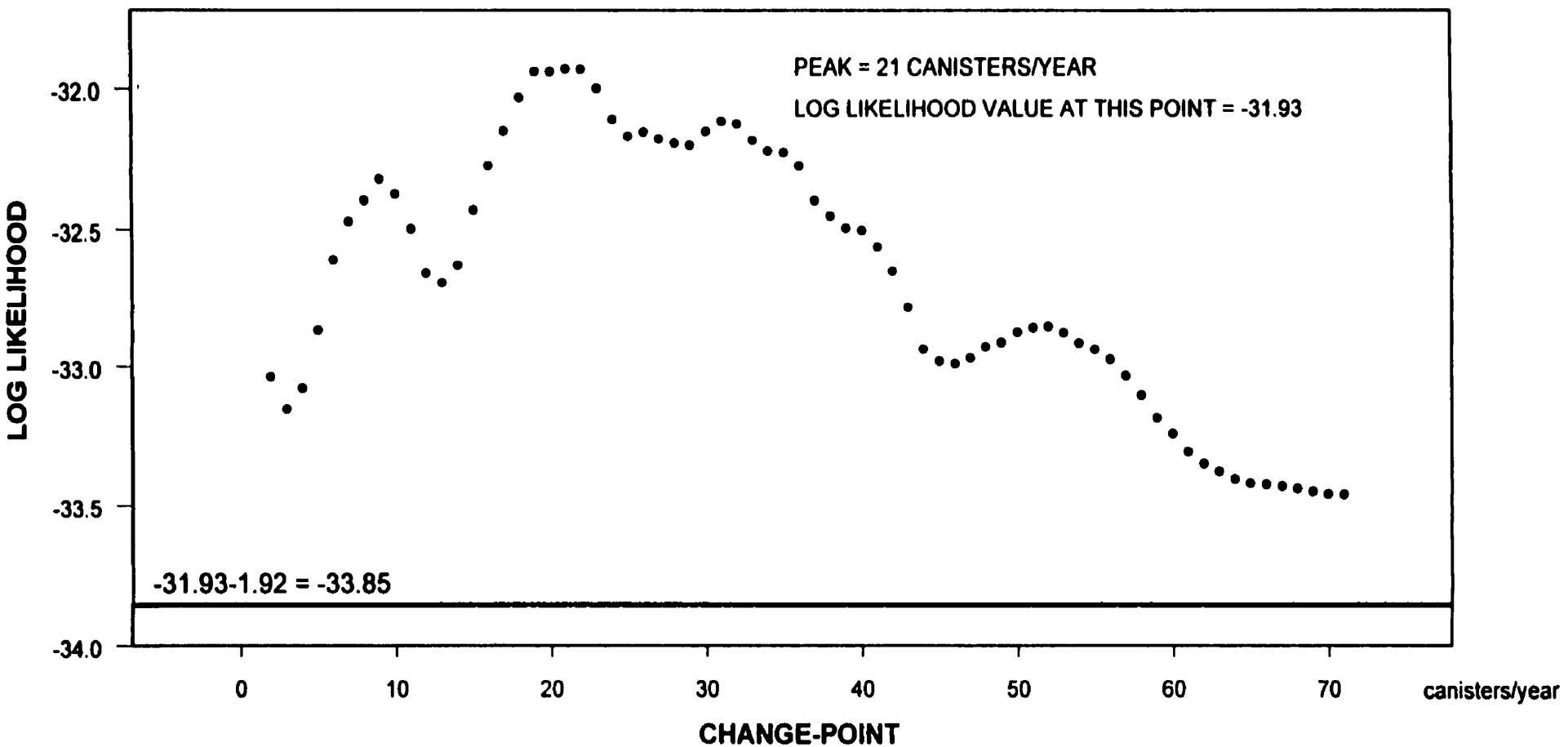

FIGURE 7.

PROFILE LOG LIKELIHOOD OF CHANGE-POINT MODEL WITH IDENTITY LINK (MODEL 3.4) WITH CONSTANT RATE BEFORE THE CHANGE-PONIT

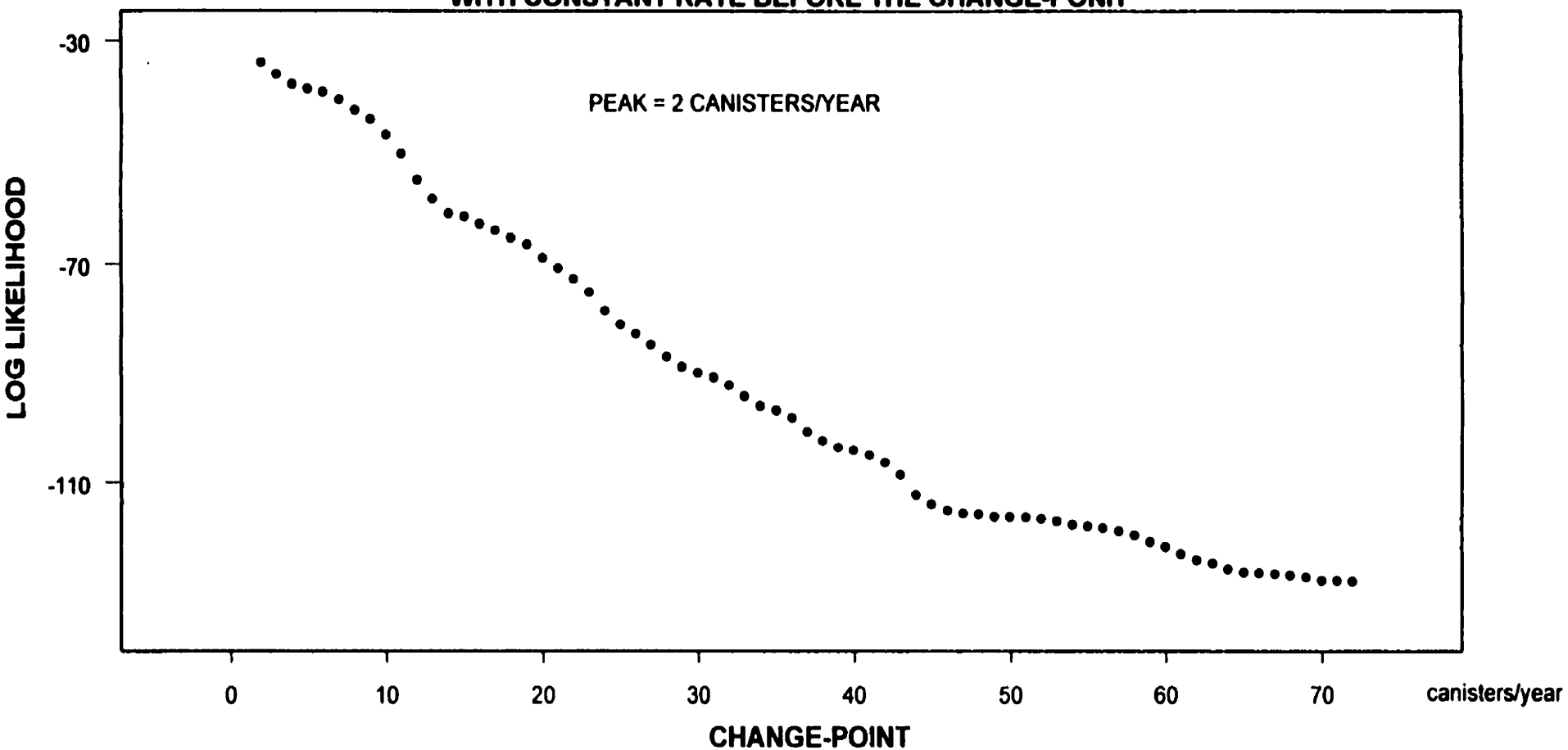

FIGURE 8.

OBSERVED AND PREDICTED DOSE-RESPONSE CURVES

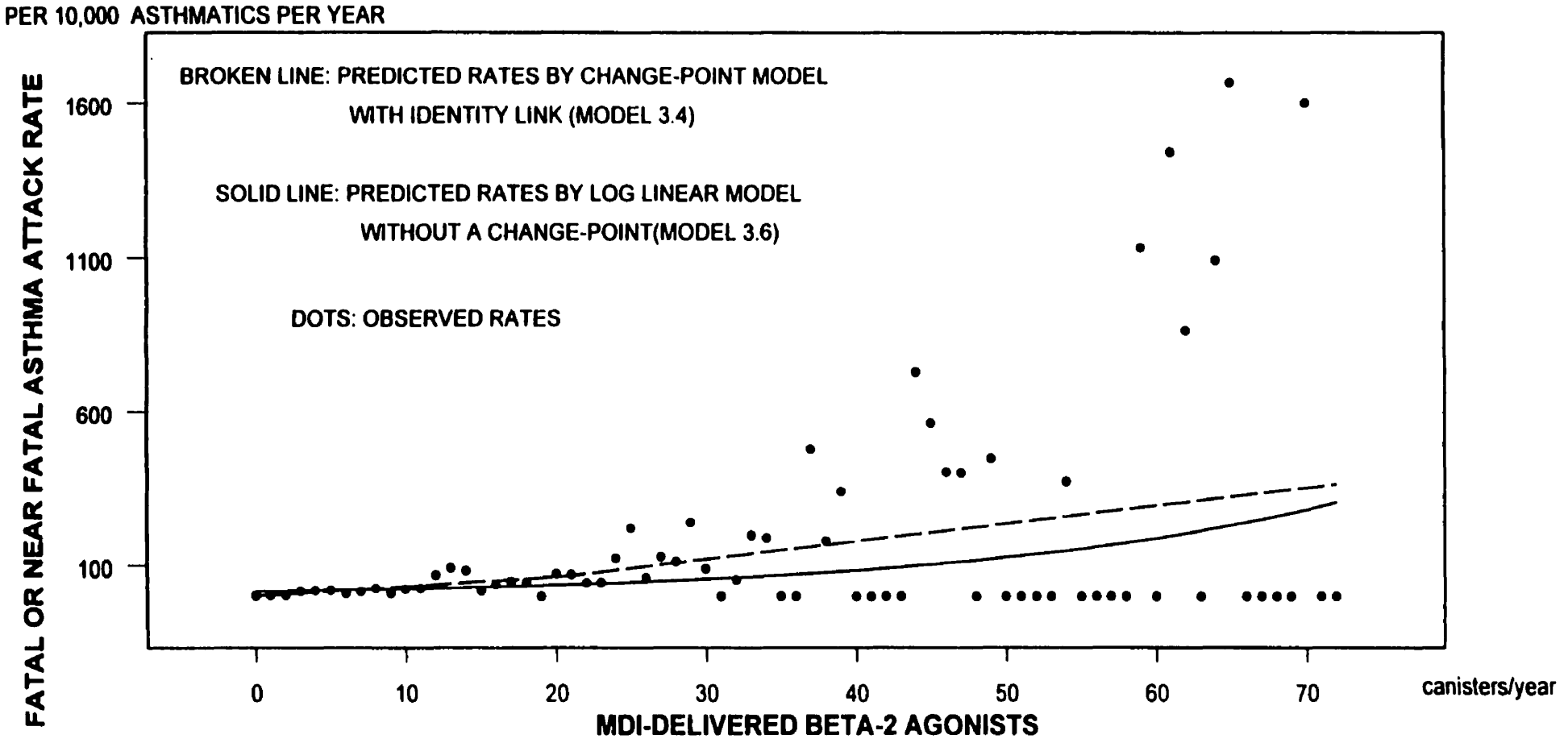

FIGURE 9.

PROFILE LOG LIKELIHOOD OF CHANGE-POINT MODEL WITH LOG LINK (MODEL 3.8)

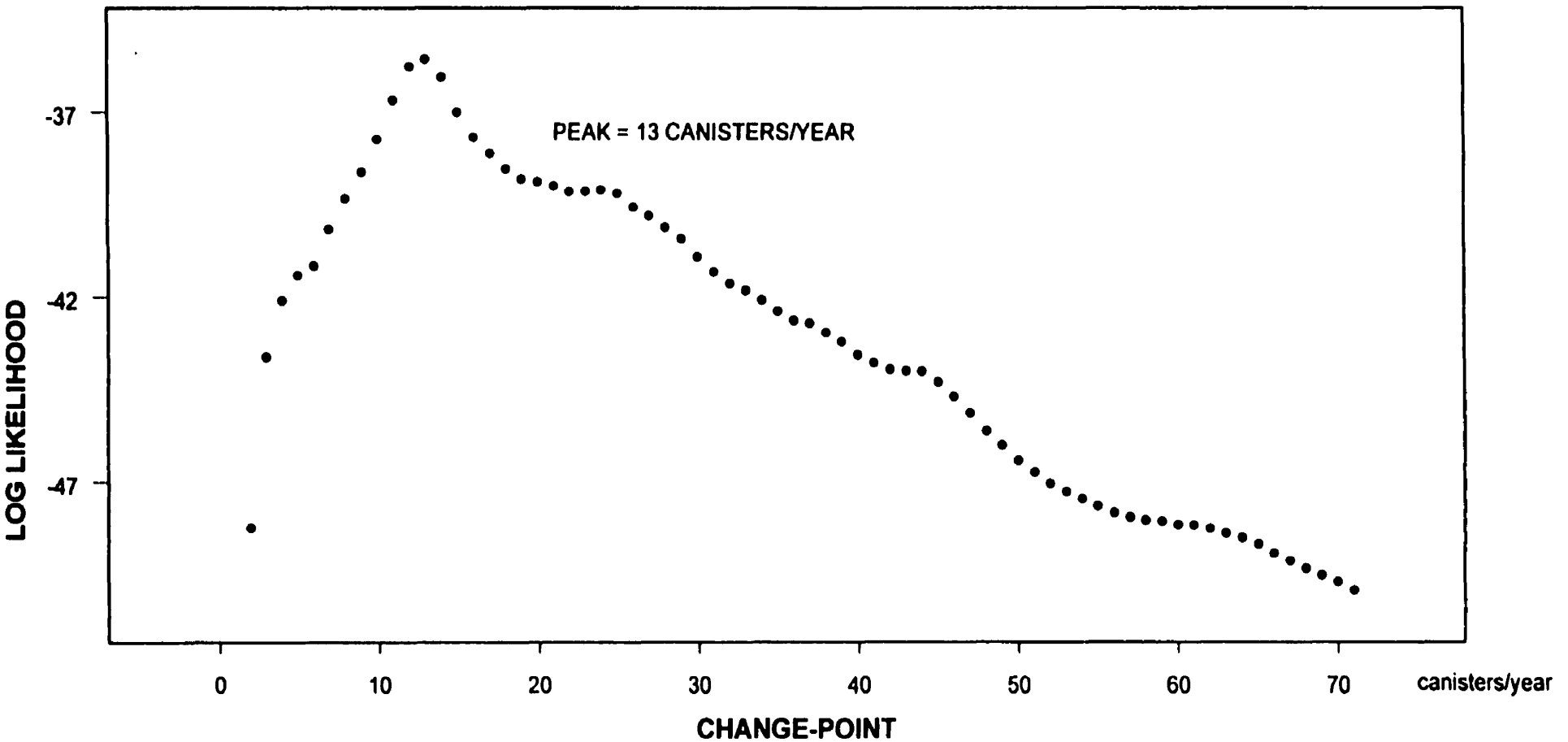

FIGURE 10.

OBSERVED DOSE-RESPONSE PLOT ON LOG SCALE

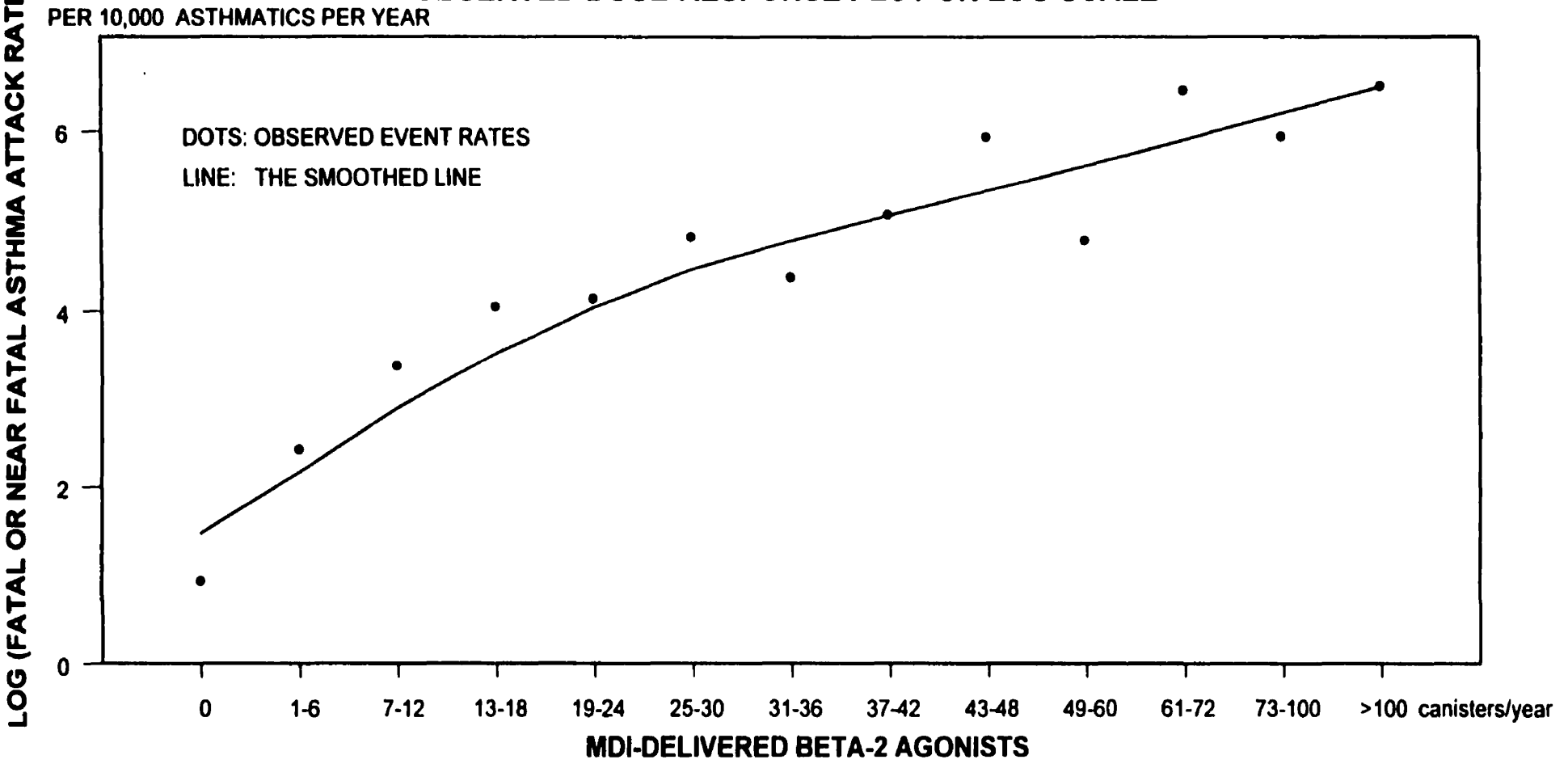

FIGURE 11.
DOSE-RESPONSE CURVE FROM ONE SIMULATED DATA

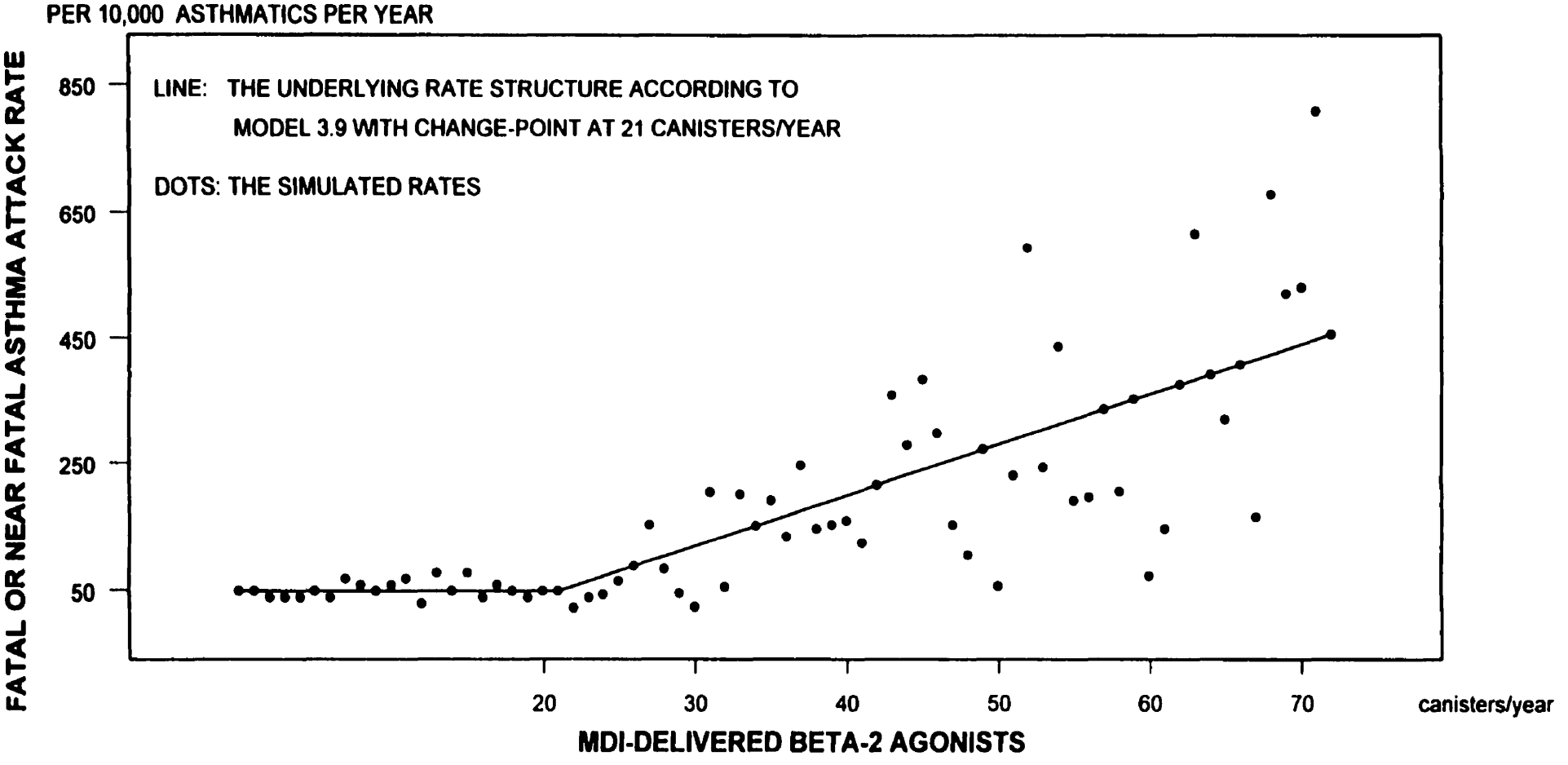

FIGURE12.

PROFILE LOG LIKELIHOOD PLOTS FOR THE CHANGE-POINT- ONE SIMULATED DATA

NOTE:THE DATA IS SIMULATED ACCORDING TO MODEL 3.9 WITH CHANGE-POINT AT 21 CANISTERS/YEAR (SEED=10 IN SAS)

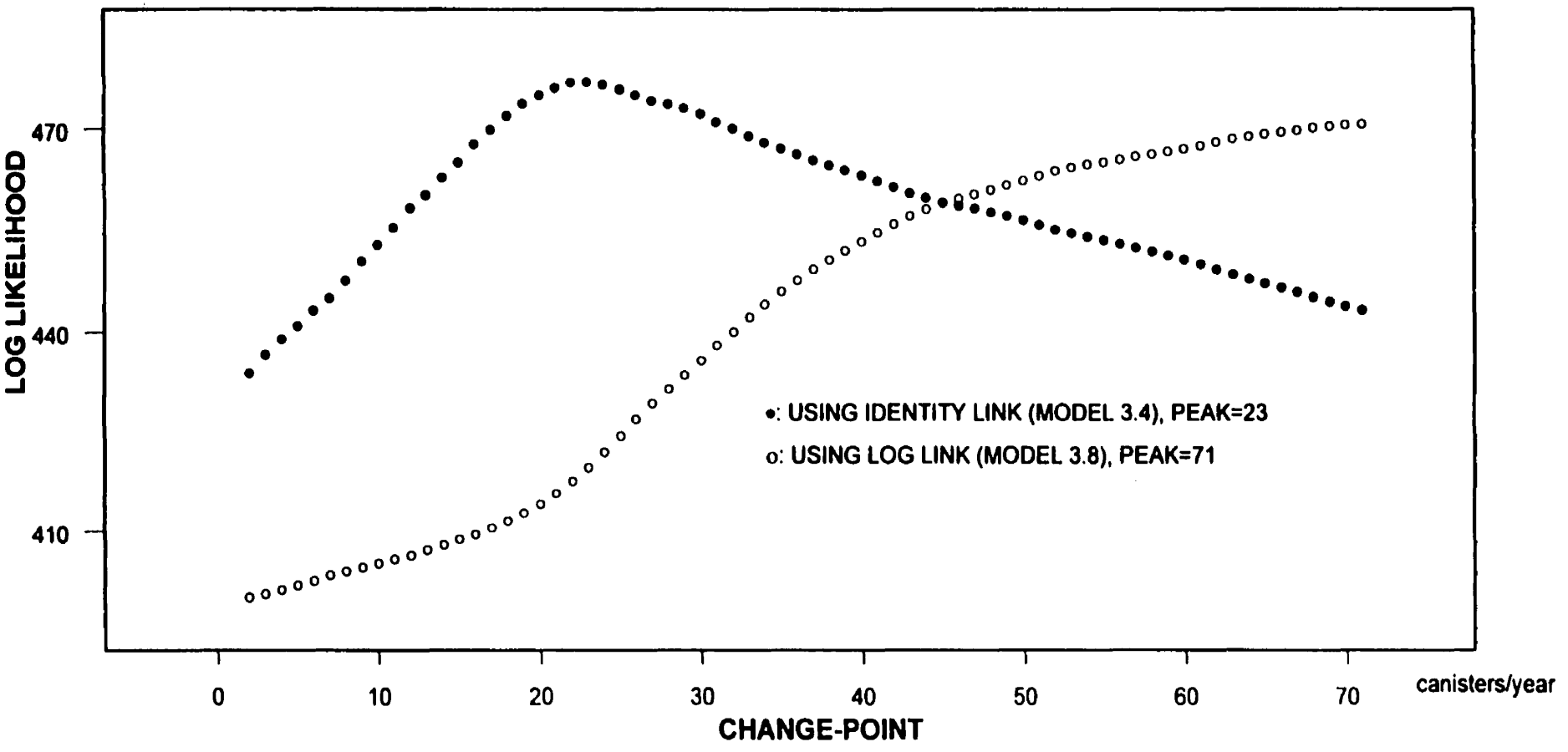

FIGURE 13.
DOSE-RESPONSE CURVE ON LOG SCALE FROM ONE SIMULATED DATA

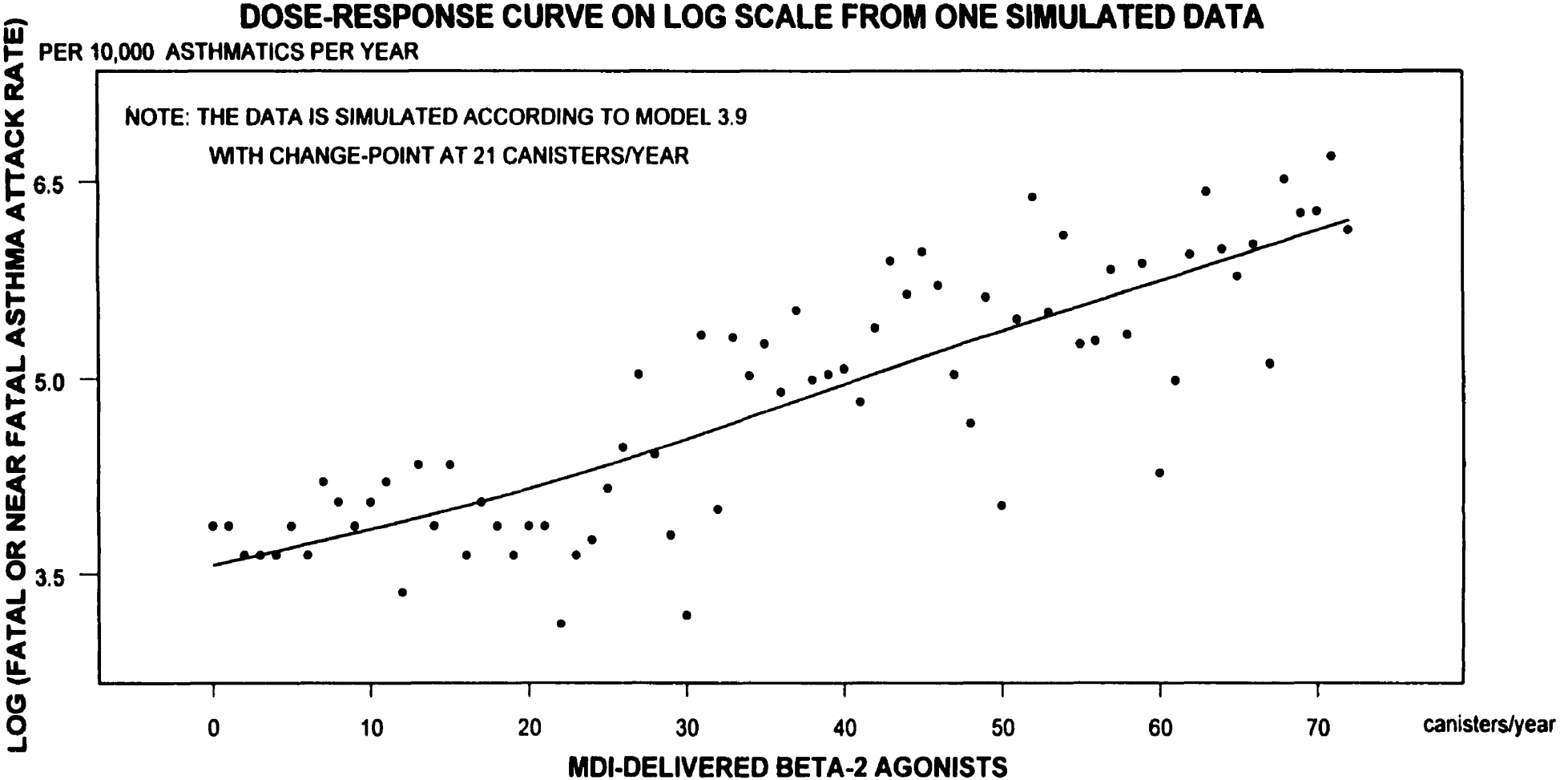

FIGURE 14.

HISTOGRAM OF MLE'S FOR THE CHANGE-POINT ESTIMATED FROM SIMULATED DATA

NOTE:THE DATA IS SIMULATED ACCORDING TO MODEL 3.9 WITH CHANGE-POINT AT 21 CANISTERS/YEAR (SEED=10 IN SAS)

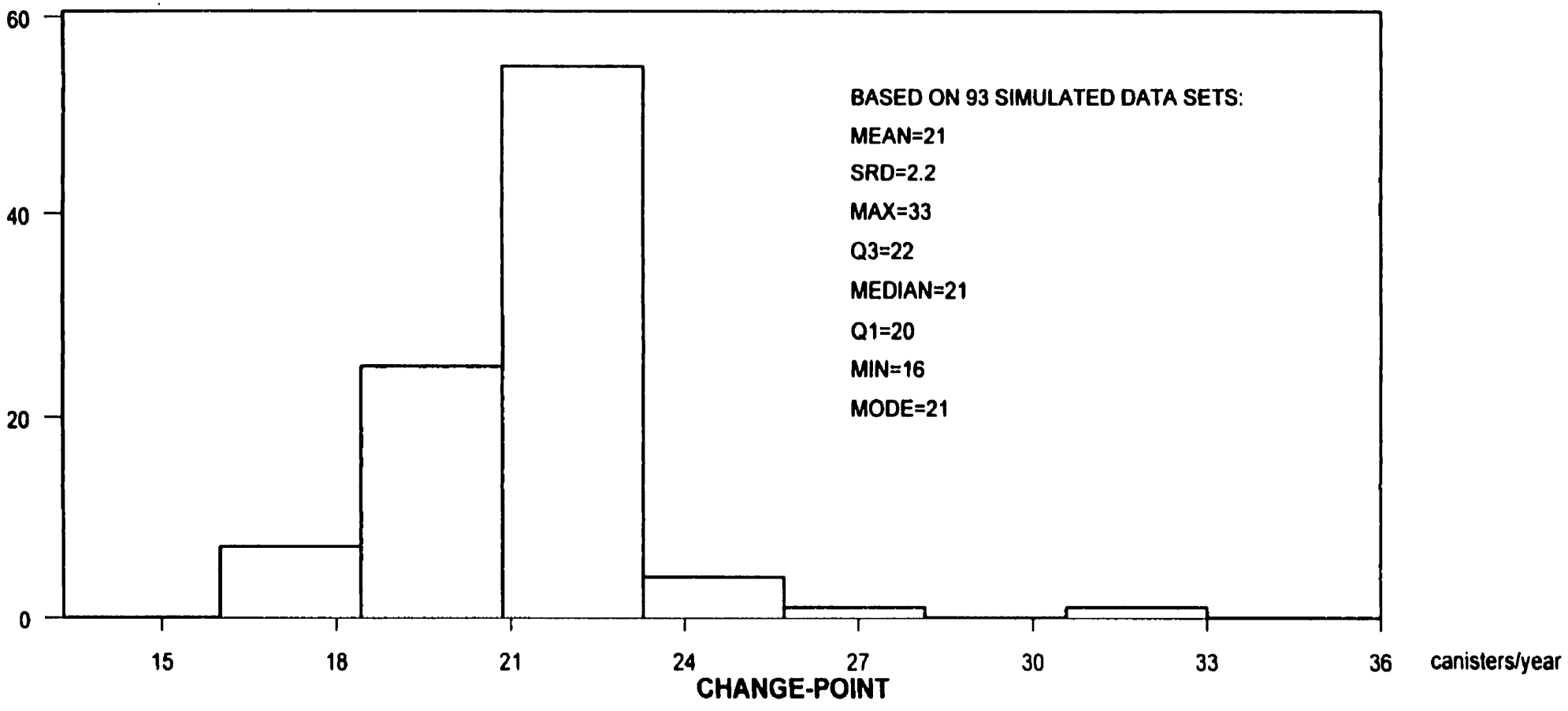

FIGURE 15.

HISTOGRAM OF NON-PARAMETRIC BOOTSTRAP ESTIMATES FOR THE CHANGE-POINT

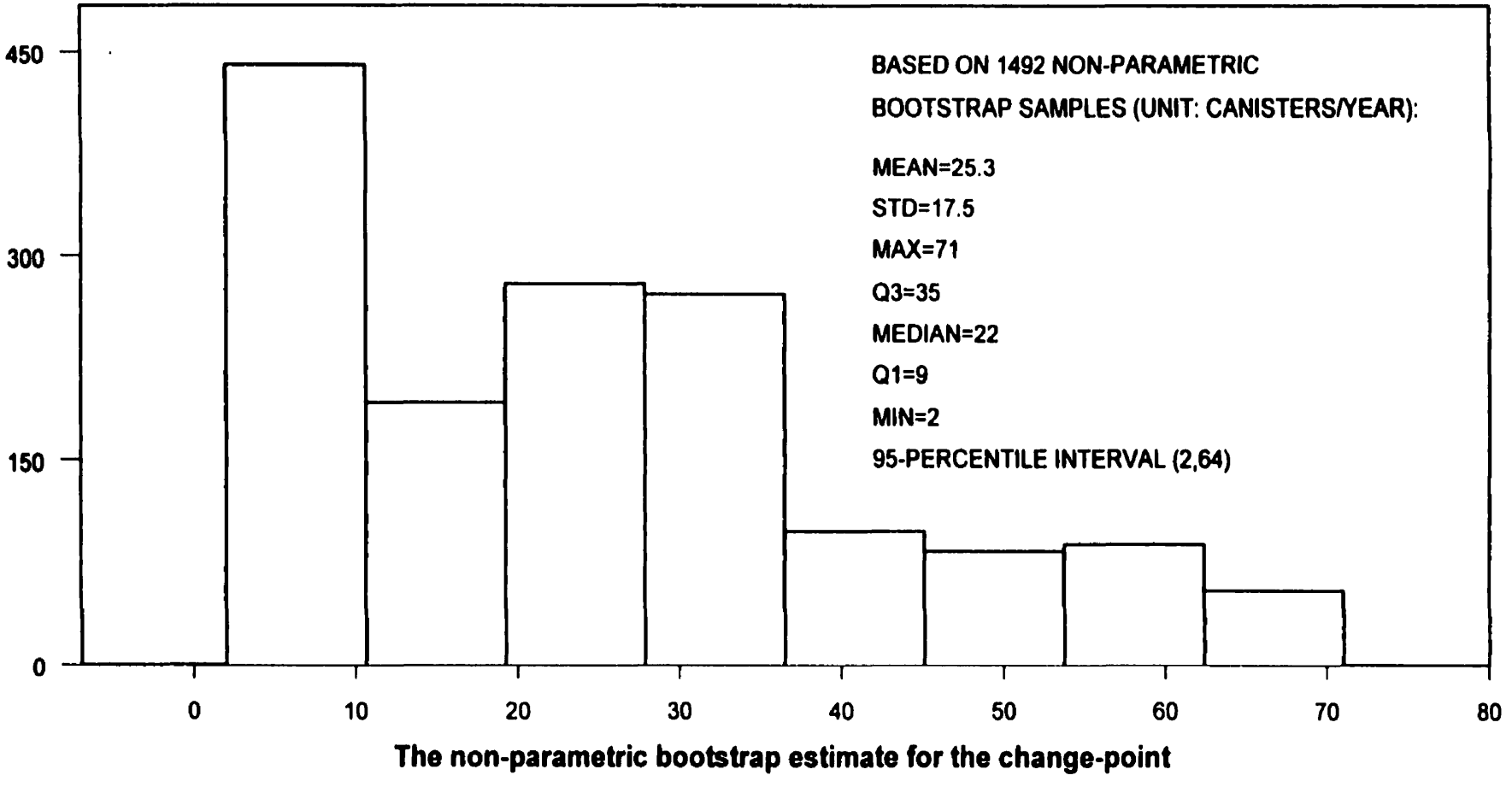

FIGURE 16.

HISTOGRAM OF PARAMETRIC BOOTSTRAP ESTIMATES FOR THE CHANGE-POINT

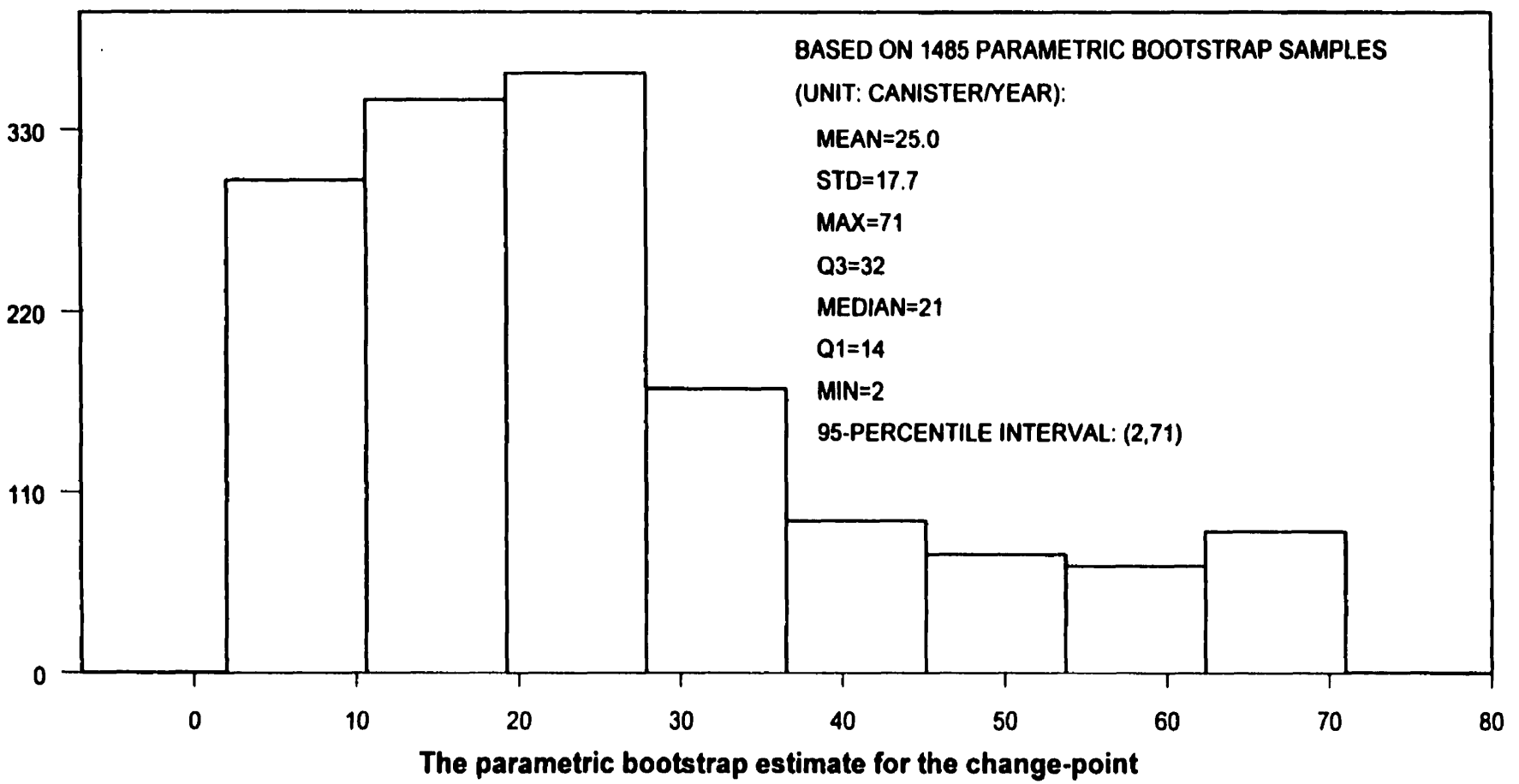

FIGURE 17.

HISTOGRAM OF NON-PARAMETRIC BOOTSTRAP ESTIMATES OF THE CHANGE-POINT IN ONE SIMULATED DATA

NOTE:THE DATA IS SIMULATED ACCORDING TO MODEL 3.9 WITH CHANGE-POINT AT 21 CANISTERS/YEAR (SEED=10 IN SAS)