Ludwig-Maximilians-Universität München
Institut für Statistik
Ludwigstr. 33
80539 München

Statistics with Focus on Economic and Social Science (M.Sc)
Master Thesis:

Non-Linearity of Radiation Dose Related Cancer Incidence Risk in the Atomic Bomb Survivor Cohort

Jonathan Gessendorfer

June 30, 2015

Correspondence: jonathan.gessendorfer@gmail.com

Supervisor: Prof. Dr. Helmut Küchenhoff

Project Partner: Dr. Jan Christian Kaiser
Helmholtz Zentrum München
Institut für Strahlenschutz
Ingolstädter Landstraße 1
85764 Neuherberg

Executive Summary

For purposes of establishing radiation protection guidelines, a linear relation of ionizing radiation dose and cancer risk is assumed. This thesis tries to detect signs for a non-linear dose-response and tests for their significance. The Life Span Study provides data on cancer incidence of the atomic bomb survivors in Hiroshima and Nagasaki. In this thesis, dose related increases in cancer risk are analysed with excess relative risk models after excluding cohort members with high doses (above 2 Gray). In addition to simple parametric dose-response functions more flexible spline approaches are employed. The likelihood-ratio test is used to test for significance of the dose-effect and its non-linearity. Monte Carlo simulations show, however, that the likelihood-ratio test statistic under H_0 does not always follow its asymptotic distribution.
The dose-effect is significant for all types of cancer analysed in this thesis.
There are no signs of non-linearity of the dose-response related increases in risk for chronic myeloid leukemia, female breast cancer and basal cell carcinoma type non-melanoma skin cancer. The dose-response for excess relative acute myeloid leukemia and squamous cell carcinoma type non-melanoma skin cancer risks is, however, significantly non-linear.

Zusammenfassung

Um Strahlungsschutz Richtlinien aufzustellen wird ein linearer Dosis-Wirkungs-Zusammenhang zwischen ionisierender Strahlung und Krebsrisiko angenommen. Ziel dieser Analyse ist es, Anzeichen für Nicht-Linearität in der Dosis-Wirkung zu finden und diese auf ihre Signifikanz zu überprüfen. Die Life Span Study ist eine Kohortenstudie über die Atombombenüberlebenden von Hiroshima und Nagasaki - sie beinhaltet unter anderem Daten zur Krebsinzidenz. Nicht in die Analyse eingeschlossen werden in dieser Arbeit Personen aus der Kohorte, die hoher Strahlung (über 2 Gray) ausgesetzt waren. Zur Analyse der Dosis-Wirkung bei verschiedenen Krebsarten werden sogenannte Excess Relative Risk Modelle verwendet. Für eine möglichst flexible Dosis-Wirkungs Funktion wird diese unter anderem mit B-Splines modelliert. Mithilfe von Likelihood-Ratio Tests wird die Signifikanz des Effekts ionisierender Strahlung und ob diese nicht-linear ist überprüft. Monte Carlo Simulationen zeigen, dass die Likelihood-Ratio Teststatistik unter H_0 nicht immer ihrer asymptotischen Verteilung folgt.
Der Dosiseffekt ist signifikant für alle in dieser Arbeit untersuchten Krebsarten.
Es gibt keine Anzeichen für Nicht-Linearität in der Dosis-Wirkung auf das Risiko für chronische myeloische Leukämie, Brustkrebs bei Frauen und Weissen Hautkrebs vom Typ Basalzellkarzinom. Bei akuter myeloischer Leukämie und Weissen Hautkrebs vom Typ Plattenepithelkarzinom ist die Dosis-Wirkung signifikant Nicht-Linear.

Contents

List of Tables

List of Figures

List of Abbreviations

LNT	Linear No-Threshold
QNT	Quadratic No-Threshold
LQNT	Linear-Quadratic No-Threshold
LT	Linear Threshold
LNT-noEM	Linear No-Threshold without Effect Modifiers
QNT-noEM	Quadratic No-Threshold without Effect Modifiers

Examples for spline model abbreviations:

lBS_M	Spline of degree 1 (linear) with a knot at the median
cBS_Q	Spline of degree 3 (cubic) with knots at the quartiles
qBS_f	Spline of degree 2 (quadratic) with a free knot

AML	Acute Myeloid Leukemia
CML	Chronic Myeloid Leukemia
FBC	Female Breast Cancer
BCC	Non-melanoma Skin Cancer - Basal Cell Carcinoma Type
SCC	Non-melanoma Skin Cancer - Squamous Cell Carcinoma Type

ERR	Excess Relative Risk
EM	Effect Modifiers
TP-Series	Truncated Power Series
BS	B-Spline
GLM	Generalized Linear Model
Gy	Gray
mGy	milli Gray
AIC	Akaike/s Information Criterion
lr	Likelihood-Ratio Test Statistic
KS-test	Kolmogorov-Smirnov Test
CI	Confidence Interval

1 Introduction

The effects of ionizing radiation on various health risks are of major public and scientific interest. Low radiation doses in particular are received by a great number of people in numerous situations - for example during computer assisted tomography or in nuclear medicine. Radiation effects to leukemia and solid cancer risks are probably the most popular field of study in that regard. A multitude of studies are indicating radiation effects on leukemia and various site specific solid cancers.[1]
There is, however, dissent in the scientific community in what functional way risks depend on the level of radiation. The most popular idea is that risks increase in a linear way with radiation dose - this model is used to asses radiation risks for purposes of radiation protection. It could, however, lack in complexity - for example a threshold dose below which radiation does not alter cancer risk could exist for certain types of cancer. Another example for non-linearity in dose-response could be that radiation effects become saturated at high doses. The main goal of this thesis is to look for signs and test for non-linearity in the dose-response function. As the effects differ greatly for different types of leukemia and site specific cancers it is necessary to analyse them separately. In this thesis the effects on acute and chronic myeloid leukemia, female breast cancer and the two important sub-types of non-melanoma skin cancer - basal and squamous cell carcinoma - are analysed. This selection is made due to the importance of these cancer types for the field of radiation research.

Leukemia particularly in children is strongly associated with radiation exposure. In Hiroshima and Nagasaki an increase in risk has been observed soon after the atomic bombings (Folley/Borges/Yamawaki 1952) and has since then been analysed in this cohort (Hsu et al. 2013). Radiation associated increases in leukemia risks were also found in uranium miners (Rericha et al. 2006), Chornobyl cleanup workers (Romanenko et al. 2008), radiation workers (Muirhead et al. 2009) and other populations (e.g. Krestinina et al. 2010).

The mammary gland is one of the organs most sensitive to radiation. Carcinogenics in the female breast associated with radiation is, therefore, thoroughly researched (an overview is given in Ronckers/Erdmann/Land 2005). The association can be observed in various populations (an overview is given in Preston et al. 2002) including in the Hiroshima and Nagasaki atomic bomb survivors (Preston et al. 2007).

The carcinogenic effect of ionizing radiation on non-melanoma skin cancers has long been observed in different populations exposed to radiation. For example in uranium miners (Sevcova/Sevc/Thomas 1978), radiologists (Matanoski et al. 1975) and also in the atomic bomb survivors of Hiroshima and Nagasaki (Preston et al. 2007).
Furthermore basal and squamous cell carcinoma skin cancer risks were investigated separately in an experimental setting while prior radiation therapy was a control variable (Karagas et al. 1996). An association between radiation and basal but not squamous cell carcinoma has

[1] A large number of papers were published in this field. Examples of important studies with relevance to this thesis are given below.

been found. These findings are in line with the results in Preston et al. (2007). The nature of the radiation dose-response seems to differ strongly, therefore basal and squamous cell carcinoma risks will be analysed separately.

This thesis uses - like Preston et al. (2007) and Hsu et al. (2013) - data of the cancer incidence in the Hiroshima and Nagasaki atomic bomb survivor cohort obtained by the Radiation Effects Research Foundation (RERF). In section 2 this dataset is described.
The relation of radiation and cancer risks can be characterized by excess relative risk models: radiation increases the risk at zero radiation exposure in a multiplicative way. In section 3 excess relative risk modelling is explained more thoroughly.
Model parameters are estimated via maximum likelihood estimation in R, a popular open source program in statistics (see section 4).
Models are selected via Akaike's Information Criterion and likelihood-ratio tests are used to test if signs of non-linearity are significant. Monte-Carlo simulations (see section 5 for theoretical explanations) are used to obtain test statistic distributions for non-standard test problems and to check if assumptions about the likelihood-ratio test statistic for nested models are valid.
Results are presented in sections 6-10 and a closing summary will be drawn in section 11.

2 Data

As stated in the introduction, this thesis uses a dataset provided by the Radiation Effects Research Foundation in Hiroshima and Nagasaki: The Life Span Study (LSS) - a cohort study of the atomic bomb survivors in Hiroshima and Nagasaki and probably the most important dataset in radioepidemiology.
It consists of a major part of the survivors who were near the hypocenters (up to 2.5 kilometres away) at the time of the bombings and an equivalently strong random sample of Hiroshima and Nagasaki residents who were between 3 and 10 kilometres away and act - because of their negligible exposition to radiation - as a control group. For about 85 percent of the first and almost all of the latter the individually received ionizing radiation dose could be estimated. For the purpose of this estimation a survey with questions regarding the whereabouts and shelter at the time of the bombings was conducted between 1950 and 1953. In the current datasets provided by the Radiation Effects Research Foundation the improved dosimetry system DS02 (cf. Young/Kerr 2005) is used. Better estimates of the received radiation dose by different organs of each individual were achieved by altering how different types of ionizing radiation are weighted to get the complete dose estimate.
After this initial study, mortality and incidence data on various diseases was collected (e.g. Preston et al. 2007, Hsu et al 2013, Preston et al. 2003). This follow-up started 1950 (for leukemia incidence and later for other diseases) and still continues. In this analysis the LSS datasets on leukemia, lymphoma and multiple myeloma incidence and on solid cancer incidence (follow up started in 1958) are used. The most recent follow-up of the former dataset collected incidence data up to 2001, of the latter up to 1998.

Both datasets[2] are available as grouped survival data (cf. for grouped survival data Armitage/Colton 1998). In these tabulations similar individuals are observed together, the events of interest that occurred in this group - this stratum - are counted and the person-years at risk in this group are observed. The person-years at risk in a stratum is the sum of the length of time that each individual was at risk for an event to happen to it.
The events of interest for this thesis in the leukemia, lymphoma and multiple myeloma incidence dataset are acute myeloid leukemia and chronic myeloid leukemia. Female breast cancer, basal cell carcinoma and squamous cell carcinoma are of interest in the solid cancer incidence dataset.

The former dataset is stratified into 38,578 strata by the variables (i.e. individuals are considered similar if they are equal in - or in the same category of - all of the following characteristics): city (Hiroshima or Nagasaki), sex, age at the time of the bombings, attained age, the calender time, the radiation dose to the bone marrow, whether the shielded kerma estimate was greater than 4 Gy or not and distance from the hypocenter (<3km, 3-10 km and not in city at the time of the bombings). (cf. Hsu et al. 2013: p.363)

[2] A documentation of each dataset can be found in the supplementary material of Preston et al. (2007) and Hsu et al. (2013) respectively. Datasets and documentations are available at:
`http://www.rerf.or.jp/library/dl_e/lssinc07.html` (Last called upon: June 25, 2015)
`http://www.rerf.or.jp/library/dl_e/lsshempy.html` (Last called upon: June 25, 2015)

The stratifying factors in the solid cancer incidence dataset were: city, gender, radiation dose to the colon, the follow-up period, attained age, age at exposure and distance from the hypocenter. For the analysis of non-melanoma skin cancers - this includes basal and squamous cell carcinoma types - the dataset is additionally stratified by whether or not the individual was a participant at the Adult Health Study (cf. Yamada et al. 2004) - a clinical research program on the longterm effects of radiation exposure resulting in potential health benefits to participants. The resulting number of strata is about 45,000 - without this additional stratifying variable the dataset is only stratified into about 27,000 groups. (cf. Preston 2007: p.5)

Besides case counts and person-years at risk, the person-year weighted means of age at the time of the bombings, attained age and of the radiation dose are available in each dataset. Values near the center of each category are more frequent than values near the edges because this categorization was used to stratify the dataset. For example if a stratum is comprised of five individuals and each individual has a radiation dose to the colon between 300 and 500 mGy the mean colon dose has a high chance to be around 400 mGy. This characteristic is especially important to note when using very flexible models to estimate the effect of the radiation dose. Patterns in the dose-response are more meaningful, when they happen at a dose level with data support (i.e. near the center of a category) than without support. In both LSS datasets this problem is relatively small, because on the one hand the intervals that are used for stratification are very small at lower doses - it follows that there are no large dose intervals without data support - and on the other hand only a small number of individuals are in the strata at high doses - what means that taking the mean of the doses has only a small impact.
Even though it is likely that individuals with a similar dose to the colon also have a similar dose to other sites (for example to the breast), it follows that this characteristic of the dose distribution is even less pronounced for doses that were not used to stratify the data. This is the case for female breast cancer and both non-melanoma skin cancer types because in the models for the former the radiation dose to the bone marrow is used and in the latter to the breast and skin.

In this thesis all strata are excluded from the analysis with the respective doses above 2 Gy (as done in Pierce/Preston 2000). There are two reasons for this exclusion: Dose estimates for higher doses are less precise than for lower doses and - as stated in the introduction - this thesis is especially interested in the risks of low radiation doses. The effects of higher doses are often very pronounced. Therefore, estimates when adapting to these effects cannot - especially if non-flexible dose-response models are used - model the dose-response for lower doses correctly.

3 The Excess Relative Risk Model

The relation between cancer risk and ionizing radiation dose can be described using an excess relative risk (ERR) model (cf. Thomas 2009: p.70sq). This model assumes a multiplicative relation between the risk at zero dose conditioned on given covariates - the baseline hazard rate[3] - and the excess relative risk:

$$\lambda = \lambda_0(X_{BL}) \cdot \big(1 + ERR(d, X_{EM})\big)$$

Where λ denotes the total hazardrate, $\lambda_0(X_{BL})$ the baseline hazardrate (given covariates X_{BL}) and $ERR(d, X_{EM})$ the excess relative risk: A function of the dose d and certain effect modifying covariates X_{EM}. The $ERR(d, X_{EM})$ model can be divided into two parts: First the functional form of the dose effect $err(d)$ - where the focus of interest in this thesis lies - and second a function $\epsilon(\cdot)$ of effect modifying variables, which can intensify or weaken the dose-response but not change its underlying functional form. The relation between these two parts is multiplicative:

$$ERR(d, X_{EM}) = err(d) \cdot \epsilon(X_{EM})$$

The number of cases Y_i in each strata group $i = 1, \ldots,$ n is assumed to be poisson distributed with the person-years at risk PYR_i in each strata group i as an offset (for poisson distribution with offset cf. Tutz 2012: p.160):

$$Y_i \sim Po(\lambda_i \cdot PYR_i)$$

If the Y_i are independent and conditioned on PYR_i identically distributed - which is also assumed - one can write:

$$Y_i \sim Po(\lambda \cdot PYR_i)$$

The parameter of the poisson distribution λ just depends on the covariates in strata group i but is independent of PYR_i:

$$\lambda = \exp(\beta X_{BL}) \cdot \big(1 + err(d) \cdot \epsilon(X_{EM})\big)$$

To ensure no negative numbers are expected $\big(1 + \hat{err}(d) \cdot \hat{\epsilon}(X_{EM})\big) > 0$ has to hold.

The following subsections will describe all components of this model and how they are used for the analysis in this thesis in detail.

[3] No to be confused with the baseline hazard rate in a Cox Proportional Hazard Model.

3.1 Baseline Hazard Model

The baseline hazard model is used to describe the risk with no exposure to ionizing radiation. That risk, of course, depends on various covariates. The independent variables in the baseline are similar to control variables in generalized linear models - the main purpose of the baseline model is to ensure that dose-response estimates are free of the effects of other variables. In epidemiology the number of observed variables is usually limited. The Life Span Study provides data on year of birth (as age at exposure), attained age, city and sex for the estimation of baseline risks. Additionally, an indicator for individuals who were in the city at the time of the bombings or not was included in the baseline models in Hsu et al. (2013). Differences in unobserved but important variables between individuals in and not in each city at the time of the bombings can thereby be controlled. Baseline risks are commonly modelled with these variables including interactions between city and sex, and various functions of age at exposure and age attained for each sex. Variables for the baseline model for each site specific cancer risk are chosen via AIC (see section 5) by Preston et al. (2007) for solid cancers and Hsu et al.(2013) for leukemia, lymphoma and multiple myeloma after allowing for a linear dose effect (see LNT in subsection 3.3). In this thesis the baseline modelling of Preston et al. (2007) and Hsu et al. (2013) is used.

3.2 Effect Modification

Effect Modification describes interactions between the dose-response and covariates. If certain traits could affect the intensity of the impact of ionizing radiation on excess relative risk these traits should be considered in the model by using them as effect modifying variables. It is, for example, often very plausible that the age at which an individual was exposed to radiation has great impact on the excess relative cancer risk a dose inflicts.
In exactly the same way as for the baseline hazard model, this analysis is using the effect modifying variables as chosen by Preston et al. (2007) and Hsu et al. (2013). In Hsu (2013) effect modification is described using log-linear functions of city, sex, age at exposure, time since exposure or age attained. In Preston (cf. 2007: p.6) multiplicative relations of the variables are also possible.

3.3 Dose-Response

The functional form of the excess relative cancer risk - the shape of the ionizing radiation dose-response - is of main interest in this thesis. Therefore, $err(d)$ is modelled in different ways:

1. Linear No-Threshold (LNT):

   $err(d) = \alpha d$

The simplest model is the so called linear no-threshold (LNT) model. Modelling the dose-response like this, means assuming radiation has starting from the smallest possible dose an impact on risk and a ten times higher radiation dose equals a ten times higher excess relative risk. Even though it is the standard way to model excess relative cancer and non-cancer risks for purposes of radiation protection (cf. Little et al. 2012) there has been ongoing critique on this very - and possibly too - simple way of modelling dose-response of ionizing radiation for some site specific cancers (cf. Preston et al. 2007: p.5).

2. Quadratic No-Threshold (QNT):

   $err(d) = \alpha d^2 \cdot 1.12$

Only a quadratic effect is used in this model. Due to the lack of a linear dose effect the dose-response is forced to be convex. This means that it is not possible to identify dose effect saturation. But due to the small number of required parameters QNT is often the model of choice for possible non-linear relations. Comparatively small effects of low doses can be fitted very well with QNT modelling.
The factor 1.12 is used to correct standard errors in parameter estimates for uncertainty in the radiation dose estimates. It is proposed by Pierce et al. (1990 in Schöllnberger et al. 2012: 167) as an additional correction for quadratic terms - general random measurement error correction is already included in the data. Pierce et al. corrected it to 1.15 in 2008. In this thesis - as in most published articles - the factor 1.12 is still used for better comparability. After estimating the model the maximum likelihood estimate of α needs to be divided by 1.12.

3. Linear-Quadratic No-Threshold (LQNT):

   $err(d) = \alpha_1 d + \alpha_2 d^2 \cdot 1.12$

Using a linear and a quadratic term, LQNT is - compared to LNT and QNT - a more flexible dose-response model. It can estimate concave functions and even protective effects of lower doses. As in QNT the factor 1.12 is used to correct for random measurement error.

4. Linear Threshold (LT):

   $err(d) = \alpha(d - d_1) \cdot I[d \geq d_1](d)$

While $I[d \geq d_1](d)$ denotes the indicator function - if the interval defined in the square brackets contains d the function value is 1, if not $I[\cdot](d) = 0$.
The linear threshold model assumes a linear effect after a threshold. Excess relatives risks are zero for doses smaller than the estimated threshold d_1. Like LQNT, LT needs two parameters - one for the linear term and one for the threshold. (cf. for the above explained models Schöllnberger et al. 2012: 167)

The goal of this thesis is to show that LNT is often not complex enough to estimate the true dose-response functions. To allow for more flexibility spline modelling (cf. Fahrmeir et al. 2013: p.415seqq.) is used in the dose-response.

5. Truncated Power Series with m knots d_j, degree k:

   $$err(d) = \alpha_1 d + \cdots + \alpha_k d^k + \sum_{j=1}^{m} \alpha_{k+j}(d - d_j)^k \cdot I[d \geq d_j](d)$$

Even though only B-Splines are used, it is worthwhile to explain the flexibility of splines by introducing the Truncated Power Series (TP-Series). The model is a polynomial of grade k, but the k'th parameter is changing at every knot d_j. Truncated Power Series are k-1 times constantly differentiable at the knots (cf. Fahrmeir et al. 2013: p.418).
For example a TP-Series of degree 1 and a single knot is a piecewise linear function with a change in the slope at the knot. It is constant but not constantly differentiable. A TP-Series of degree 2 is a linear-quadratic function with a different quadratic term after every knot.

6. B-Splines l-th degree with m inner knots d_j:

   $$err(d) = \sum_{j=1}^{m+l+1} \alpha_j B_{j,l}(d)$$

   The basis functions $B_{j,l}(d)$ are defined recursively:

   $$B_{j,l+1}(d) = \frac{d - d_j}{d_{j+1} - d_j} B_{j,l}(d) + \frac{d_{j+l+1} - d}{d_{j+l+1} - d_{j+1}} B_{j+1,l}(d)$$

   While degree l=0 basis function are defined as follows:

   $$B_{j,0} = I[d_j \leq d \leq d_{j+1}](d)$$

B-Splines are weighted sums of basis functions. These functions are defined recursively and take effect on intervals defined by the knots. With a higher degree of the spline follow wider intervals on which each basis function acts (more knots are overlapped by each function). The higher the spline degree, the more basis functions take effect at each knot.
To visualize how these basis functions are defined, figures 3.1 and 3.2 show the degree 1-3 B-Spline basis functions for the dose to the bone marrow on the restricted dataset used for the analysis of acute and chronic myeloid leukemia risks.
B-Splines and TP-Series of the same degree and with the same knots are equivalent in the estimated functions. Even though the parameter estimates of the former are not as easy to interpret, they are used in this thesis because B-Splines are numerically much more stable (cf. Fahrmeir et al. 2013: 426f).
The number of knots and their positions are essential for the estimated dose-response function. Therefore, models with one free knot as well as models with a fixed knot at the median and knots at the quartiles of the dose-distribution are estimated.[4]

[4]In figures and tables splines will be denoted as shown in the list of abbreviations.

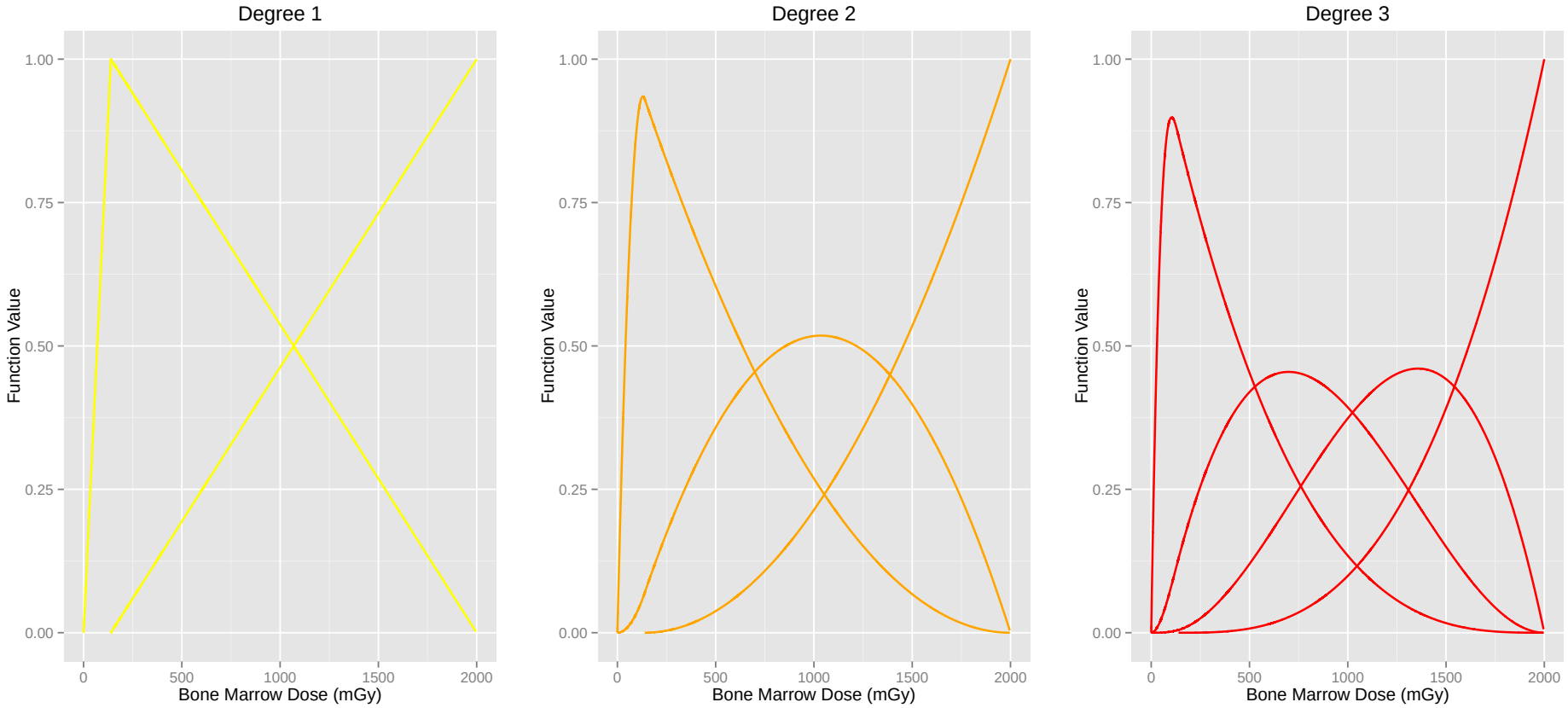

Figure 3.1: CML & AML - B-Spline Basis Functions, Knot at Median

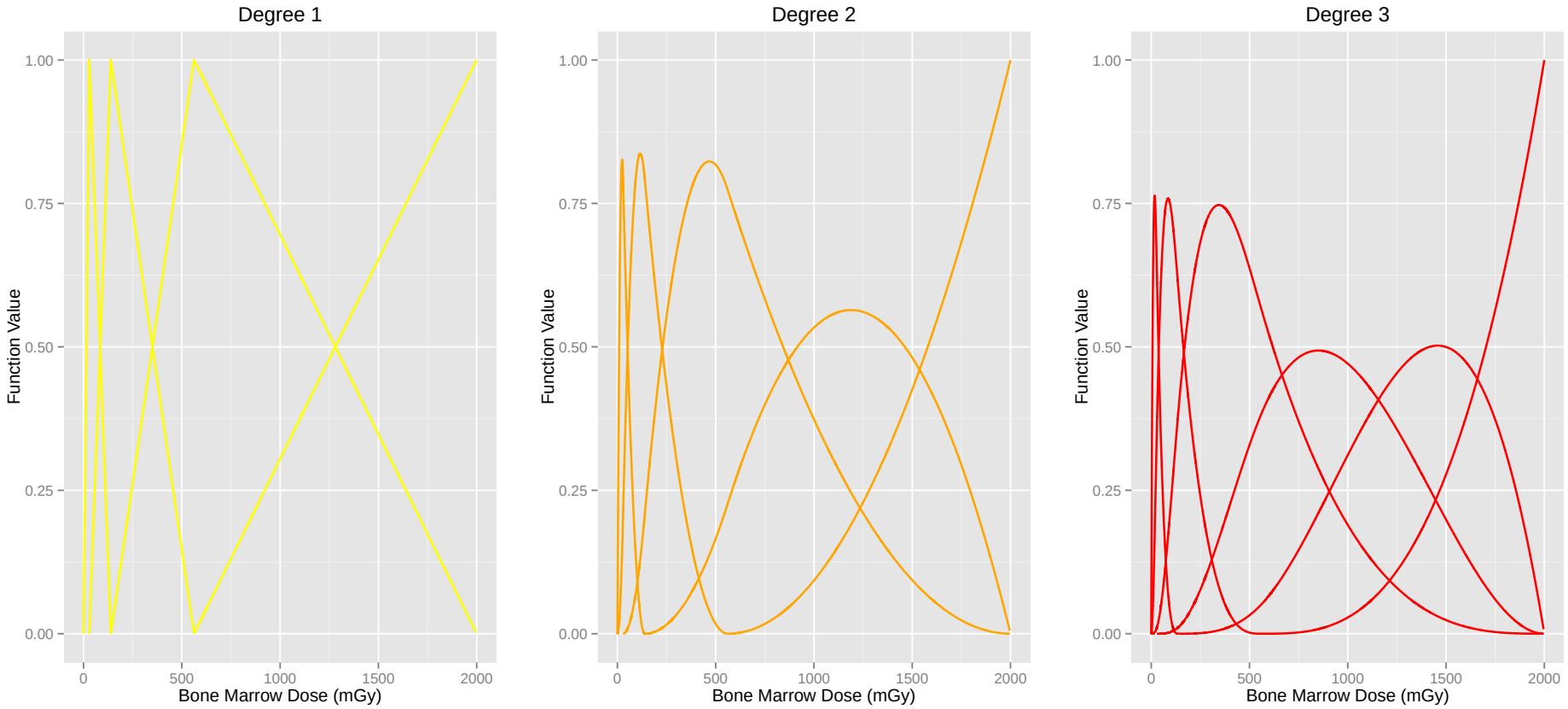

Figure 3.2: CML & AML - B-Spline Basis Functions, Knots at Quartiles

4 Maximum Likelihood Estimation

The parameter vector of each model is estimated via maximum likelihood. The likelihood function can be derived as follows:
As stated in section 3 a poisson distribution $Y_i \sim Po(\lambda \cdot PYR_i)$ is assumed. While:

$$\lambda = \exp(\beta X_{BL}) \cdot \big(1 + err(d) \cdot \epsilon(X_{EM})\big)$$

With the poisson distribution follows that the probability for y_i events in stratum i is:

$$\mathbb{P}_\lambda(Y_i = y_i) = \frac{(\lambda \cdot PYR_i)^{y_i}}{y_i!} \exp(-\lambda \cdot PYR_i)$$

Assuming all Y_i are independent of each other, the complete likelihood is the product of the individual densities:

$$L(\lambda | X_{BL}, d, X_{EM}) = \prod_{i=1}^{n} \frac{(\lambda \cdot PYR_i)^{y_i}}{y_i!} \exp(-\lambda \cdot PYR_i)$$

The corresponding log-likelihood is:

$$l(\lambda | X_{BL}, d, X_{EM}) = \log L(\lambda | X_{BL}, d, X_{EM}) = \sum_{i=1}^{n} \big(y_i \log(\lambda \cdot PYR_i) - \log(y_i!) - \lambda \cdot PYR_i\big)$$

Maximizing the log-likelihood yields $\hat{\lambda}$ (and of course $\hat{\alpha}$, $\hat{\beta}$ and $\hat{\gamma}$). For the vector of expected number of cases $\hat{y}$ follows (with PYR denoting the vector of person-years at risk):

$$\hat{y} = \exp(\hat{\beta} X_{BL}) \cdot \big(1 + \hat{err}(d) \cdot \hat{\epsilon}(X_{EM})\big) \cdot PYR$$

Excess relative risk models are rather uncommon non-linear models, therefore very specialized programs like Epicure (a software for risk analysis developed by Hirosoft International Corporation; for a description see Preston et al. (1993)), are commonly used for analysis. These programs are, however, rather expensive. Therefore, R (R Core Team 2014) a very popular open source programming language in statistic is used in this thesis. For the most common statistical models very well working and efficient R packages have been developed by members of the R community. However, for excess relative risk models no generally usable package has yet been written.[5]
The negative log-likelihood function can, however, be set up manually and minimized with R inherent tools - this, of course, is equivalent to maximizing the log-likelihood.

[5]For some simple excess relative risk models it is possible to use the package `gnm` - generalized non-linear models (Turner/Firth 2012). It allows to program self written functions. It is, however, not possible to include the constraint $\big(1 + err(d) \cdot exp(\gamma X_{EM})\big) > 0$, which, as described in section 3 on page 5, is essential to excess relative risk models. The underlying optimization is using the Newton algorithm (cf. Deuflhard 2011: p.11) with a preceding method to choose good starting values. The Newton algorithm uses the gradient of the non-linear function to linearise the optimization problem. The model under the constraint is, however, not differentiable. Constraining the non-linear part of the function like this is therefore not possible in `gnm`. However, simple models with few parameters in dose-response and effect modification can often be optimized with `gnm` because the constraint often holds without explicitly modelling it. With more complex models as used in this analysis a different approach has to be used.

A multitude of optimization tools are provided by R. In this thesis mainly the function `nlminb(·)` - non-linear minimization in bounds - is used. This optimizer is using PORT routines (cf. Fox/Hall/Schreyer 1977) to minimize functions. Besides PORT-Routines, a Nelder-Mead (cf. Nelder/Mead 1965) and the Quasi-Newton method of Broyden, Fletcher, Goldfarb and Shanno (all four were published independently: cf. Broyden 1970, Fletcher 1970, Goldfarb 1970 and Shanno 1970) were used. Both are available in the R function `optim(·)`. Nelder-Mead is a simplex algorithm. A simplex is the simplest possible volume in a p-dimensional space (a volume that is using all p dimensions) - it is defined by p+1 vertices. For example in a 2 dimensional space a simplex is a triangle, in 3 dimensional space a tetrahedon. Each vertice is defined by a set of parameter values. Evaluating the function at this argument yields a function value. By replacing the worst vertice (with the highest function value) with another vertice the simplex is shifted towards the minimum. After a first convergence criterion - if replacing vertices does not yield considerable change in the position of the simplex - the p worst vertices are drawn in towards the best vertice by reducing their distance to it. In this second part of the algorithm the simplex is tightened around the minimum. (cf. Nelder/Mead 1965: p.309)
The method of Broyden, Fletcher, Goldfarb and Shanno is based on the idea of Newton's method in optimization (cf. Deuflhard 2011): linearising (i.e. using the derivative of) the non-linear function. It is, however, numerically more efficient because the hessian matrix is only approximately estimated.
In this thesis `nlminb(·)` is primarily used because it was able to replicate the estimates of Preston et al. (2007) and Hsu et al. (2013), it is rather efficient and the possibility to define bounds for certain parameter estimates is crucial for simulations (see section 5). Only if convergence with `nlminb(·)` is questionable the other optimization tools are used to look for better estimates.

Whenever a knot is free - i.e. when it is included in the model as a parameter - a grid search (cf. Fahrmeir et al. 2013: p.481) is used to estimate its position. This manual and direct approach to optimization is used because knot positions very often have local maxima. Therefore when using a conventional optimizer the starting value can potentially have a huge impact on the outcome of an estimation by letting the optimization converge to a local optimum.
The grid-search is a direct optimization procedure. The main idea is to estimate multiple models with the parameter fixed at different values. In the first step parameter values at the whole range of possible (or plausible) outcomes are used. The parameter value, whose model yields the best function value is used as the center for the smaller grid of the next step. The grid is tightened around the best parameter value until the parameter is estimated sufficiently accurate.

5 Model Selection and Testing for Non-Linearity

In general the Akaike Information Criterion (AIC) is used for model selection in this thesis. The AIC rewards a good fit to the data but penalizes the number of parameters used to achieve the fit:

$$\begin{aligned} AIC &= -2 \cdot l(estimated\,model) + 2p \\ &= Deviance + 2p \end{aligned}$$

While $l(\cdot)$ denotes the log-likelihood and p the number of parameters. Models with lower AIC are preferred (cf. Fahrmeir et al. 2013: 288). However, a model selected this way does not have to produce a *significantly* better fit than the model not selected by the AIC.

To test for significance of parameters Hsu et al. (2013) and Preston et al. (2007) use the likelihood-ratio test (Fahrmeir et al. 2013: 662) and approximate the test statistic distribution with a $\mathcal{X}^2$ distribution:
The likelihood-ratio test can be used to formally evaluate the significance of log-likelihood differences between nested models. A model is nested in another model if the first model (the H_0-model) can be generated by restricting parameters of the second model (the H_1-model). Of course, the nested model always has a lower likelihood than the model it is nested in. The likelihood-ratio test answers the following question: Is the difference in log-likelihood so large that one can be certain enough to assume that the nested model is not the true data generating process.
Denoting the vector of parameters which need to be restricted to zero in the H_1-model to get the H_0-model as β_r, yields the formal hypotheses:

$$H_0 : \beta_r = 0 \quad \text{vs.} \quad H_1 : \beta_r \neq 0$$

The likelihood-ratio test statistic[6]

$$lr = 2(l(\hat{\theta}) - l(\tilde{\theta}))$$

compares the log-likelihood of the unrestricted model $l(\hat{\theta})$ with the log-likelihood of the restricted model $l(\tilde{\theta})$.
Under H_0 - that is if H_0 is true; if a model in H_0 is the true data generating process - lr *asymptotically* follows the $\mathcal{X}^2$ distribution with r degrees of freedom, while r denotes the number of restricted parameters. If the estimated lr is so high, that it would be improbable to assume that it is a random variable from the $\mathcal{X}^2$ distribution with r degrees of freedom, then it is equally improbable that H_0 is true. If it is improbable enough (depending on the confidence level) one can assume that H_0 is not true, what would mean that at least one parameter in β_r is unequal to zero.

[6] In literature often denoted as Λ.

The assumption that lr actually follows its asymptotic distribution could, however, not be generally true. Another main goal of this thesis is to check if these assumptions are appropriate. If that is not the case, significance testing and confidence intervals based on the profile likelihood - that also utilize the $\mathcal{X}^2$ assumption - will be numerically much more expensive. Though they are commonly used for interval estimation (e.g. Preston et al. (2007) and Hsu et al. (2013)) they are out of scope for this thesis (simulating the respective lr distribution - to avoid false assumptions - would dramatically increase computing time).
Through Monte Carlo simulations (for a simple introduction in the philosophy of Monte Carlo simulations cf. Smith 1973: p.1seqq; for a more current textbook cf. Robert/Casella 2004) one can draw random variables from the true distribution of the likelihood-ratio test statistic under H_0:

1. Draw new event counts for each stratification group out of $Po(\hat{\lambda} \cdot PYR_i)$ while $\hat{\lambda} \cdot PYR_i$ denotes the expected number of events of stratum i provided by the estimated model under H_0. The expected number of events in each stratum of course depends on the covariates in this group.

2. Estimate both models with the new case counts and the original covariate values and calculate lr. This is a realisation of lr under H_0.

If this simulation is completed multiple times (in this thesis 1000 times) one can approximate the distribution of lr under H_0 with the resulting empirical distribution.
Comparing the quantiles of the $\mathcal{X}^2$ distribution with the simulated distribution can be a first indicator if the $\mathcal{X}^2$ assumption is appropriate or not. The most interesting is the 1-α quantile $x_{1-\alpha}$ while α denotes the confidence level because the likelihood-ratio test decision changes at $x_{1-\alpha}$. With a Kolmogorov-Smirnov test (KS-test) one can test if the simulated lr's are random variables from the corresponding $\mathcal{X}^2$ distribution or not. If not then lr does not follow a $\mathcal{X}^2$ distribution under H_0.

The linear no-threshold model is nested in every other model considered in this thesis with the exception of the quadratic no-threshold model. Restricting the quadratic term for LNT and the threshold for LT to zero yields a linear no-threshold model. Considering the TP-Series it follows that LNT is also nested in every spline model. If every parameter in the spline is restricted to zero except for the first linear term the result is a simple LNT model. From the equivalence of the TP-Series and the B-Spline follows that LNT is also nested in the latter. Testing if excess relative risks are non-linear can, therefore, in most cases be rephrased by testing if β_r is unequal to zero. As explained above this can be tested either with the help of the above mentioned simulation procedure or (if asymptotic properties are appropriate) the $\mathcal{X}^2$ assumption.
Contrary to that, the quadratic no-threshold model is only nested in LQNT and all splines of degree 2 and above. The baseline model is nested in every model used in this analysis.

However, for some nested models it is not clear what distribution lr would follow asymptotically.
When restricting the second linear parameter of a degree 1 TP-Series with a free knot to zero the knot position is irrelevant for the estimated dose-response. The dose response will not differ before and after the knot. This means that only one parameter needs to be restricted and with the above mentioned definition of the likelihood-ratio test for nested models it is assumed that lr follows under H_0 a $\mathcal{X}^2$ distribution with one degree of freedom. Considering the same distribution is also assumed of lr under H_0 when testing LNT against a degree 1 spline with a fixed knot and that this spline definitely has a lower (or equal) maximum likelihood than the respective spline with a free knot, likelihood differences compared to LNT for the former should, therefore, be higher than for the latter.
A similar problem exists when testing the baseline model against any excess relative risk model with dose modifiers. Restricting all parameters of the dose-response to zero results in meaningless effect modifiers. Therefore, LNT without effect modifiers is often used as an H_1 model, when testing for a significant dose effect.
This problem can, however, be avoided by obtaining the distribution of the likelihood-ratio test statistic under H_0 via simulation. Restricting the linear parameter in a linear threshold model to zero also results in an irrelevant knot position.

It seems that testing non-nested models against each other could work in a similar way. The test idea - explained for the case of LNT (H_0) versus QNT (H_1) - stays the same: Is the maximum likelihood of QNT so high that it is improbable for LNT to be the true model. However, rejecting the null-hypothesis in this case does not mean that QNT is chosen because the hypotheses are not composite. Table 8.4 in section 8.2 shows that LNT in this case would be improbable if the likelihood-ratio of LNT and QNT would be higher than -0.51. This means LNT would still have a higher likelihood than QNT while using the same number of parameters to achieve that likelihood. Therefore, QNT is definitely not the better model. However, if QNT is almost as good as LNT is, it is improbable that LNT is the true data generating process.

6 Acute Myeloid Leukemia

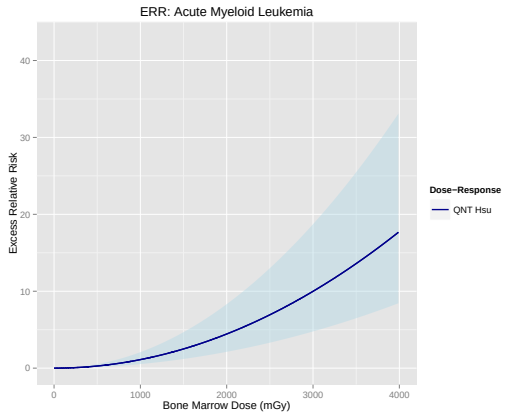

Figure 6.1: AML - in Hsu et al. (2013)

Dose	PYR	AML Cases
-0.005	2,039,095	77
-0.1	957,889	36
-0.2	201,935	9
-0.5	206,749	12
-1	117,855	11
-2	64,123	18
2+	25,761	13
Total	3,613,406	176

Table 6.1: PYR & Cases by Dose

There were 176 cases of Acute Myeloid Leukemia (AML) in the LSS cohort until 2001. By restricting the dose range to 0-2 Gy only about 0.7 percent of the available data in person-years and 13 AML cases are lost (see table 6.1[7]). Whether this number is high or low cannot, however, be a reason for restricting the analysed dose range because this would lead to bias in the estimated models.

According to Hsu et al. (2013: p.367) the baseline risk is lower for women than for men. It increases with age but the increase rate is also stronger for men. There is a birth cohort effect and a (non-significant) effect of the city on AML risk. The preferred baseline model equation in Hsu et al. (ibid.) is therefore:

$$\begin{aligned}\lambda_0(X_{BL})(\text{c, s, a, e}) = \exp \quad & (\beta_0 \\ & + \beta_1\text{female} && (1)\\ & + \beta_2\text{bcohort} + \beta_3\text{bcsq} && (2)\\ & + \beta_4\text{lage70} + \beta_5\text{lage70sq} && (3)\\ & + \beta_6\text{flage70} + \beta_7\text{flage70sq} && (4)\\ & + \beta_8\text{nic·hiro} + \beta_9\text{nic·naga}) && (5)\end{aligned}$$

While (1) denotes the effect of sex (β_1 is the difference in risk of women compared to men), (2) the linear quadratic effect of birth cohort, (3) the linear quadratic effect of log age for men and (3)+(4) for women. Not-in-city indicators are included in the baseline model by (5).[8] They are factored in to make use of cohort members who were far away from the epicenter at the time of the bombings and, therefore, potentially differ from individuals who were near the center in risk altering ways. This makes a better estimation of baseline risks and thereby a better estimation of the complete model possible.

[7] In the dose column, -0.1 for example denotes the dose interval from 0.005 to 0.1 Gy.
Tables showing the number of diagnosed cancers and person-years at risk by radiation dose and age attained and by radiation dose and age at exposure can be found in the appendix for every site specific cancer type in this analysis.

[8] Notation as in the Hsu et al. (2013) supplementary material AMFIT Code.
Online at: `http://www.rrjournal.org/doi/suppl/10.1667/RR2892.1/suppl_file/10.1667_rr2892.1.s3.pdf` (Last called upon: June 25, 2015)

The preferred dose-response model in Hsu et al. (2013) is a purely quadratic model with a linear-quadratic effect of age at exposure and a linear effect of the logarithm of age attained as effect modifiers:

$$\epsilon(X_{EM}) = exp(\gamma_1 \cdot e30 + \gamma_2 \cdot e30sq + \gamma_3 \cdot lage70 + \gamma_4 \cdot over4gy)$$

Figure 6.1 shows the dose-effect and its 95% (profile likelihood) confidence interval. For a 70 year old person, who was 30 at the time of the bombings the excess relative risk is 0.99 (95% confidence interval = [0.47-1.86]) at 1 Gy bone marrow dose.[9] The excess risk is increasing in a linear-quadratic way for an increased age at exposure (with linear term $\gamma_1 = 0.17$ and quadratic term $\gamma_2 = 0.25$) and decreasing for an increasing age attained (γ_3 = -0.89) (cf. Hsu et al. 2013).
To control for dose uncertainties in high and, therefore, uncertain dose estimates Hsu et al. include a dummy denoted by *over4gy* which is equal to 1 if the shielded kerma estimate is over 4 Gy and 0 if not (cf. Hsu et al. 2013: p.363).
In the following subsection the estimated models for acute myeloid leukemia in this thesis are discussed.

6.1 Dose-Effect Estimation

Just considering doses up to 2 Gy does not yield immediate differences in dose-response compared to the results in Hsu et al. (2013).
When comparing the AICs of simple models with dose-response functions 1-3 (see section 3.3) the best model is still QNT. The likelihood of the quadratic model is considerably higher as that of the linear model and adding a linear term has almost no effect on the estimated dose-response function and yields almost no increase in likelihood. Parameter estimates are similar to QNT in Hsu et al. (2013)[10]. At 1 Gy bone marrow dose the excess relative risk is 0.95 (symmetric 95% CI = [0.20, 1.71]) for 70 year old cohort members that were 30 at the time of the bombings. The increase in estimated risk for higher doses seems to be very high. A 2 Gy Dose almost quadruples ($0.95 \cdot 2^2$=3.8) it. Compared to that, a dose at 100 mGy has almost no effect ($0.95 \cdot 0.1^2$=0.0095). This is of course due to the QNT modelling. The excess acute myeloid leukemia risk increases with age at exposure (γ_1=0.21 and γ_2=0.24) and decreases with age attained (γ_3=-1.32).
The best threshold estimate is at 469 mGy. The likelihood of QNT is higher even though it needs one parameter less to fit.
As described in section 4 a grid search is used to estimate the threshold. The first grid covers the whole dose range used in this analysis with 100 mGy accuracy. In every following step the grid is tightened around the maximum of the likelihood. The whole grid search can be retraced in figure 6.4. The deviance of every model is plotted for each fixed knot position.

[9] In Hsu et al. (2013) the parameter estimate was not corrected, after multiplying the quadratic dose by 1.12 for measurement error corrections. See the QNT model in subsection 3.3 on page 6 as well as the AMFIT code supplement to the paper of Hsu et al. (2013).

[10] Summaries of the results of LNT and the preferred model by AIC can be found in the appendix for every cancer type analysed in this thesis.

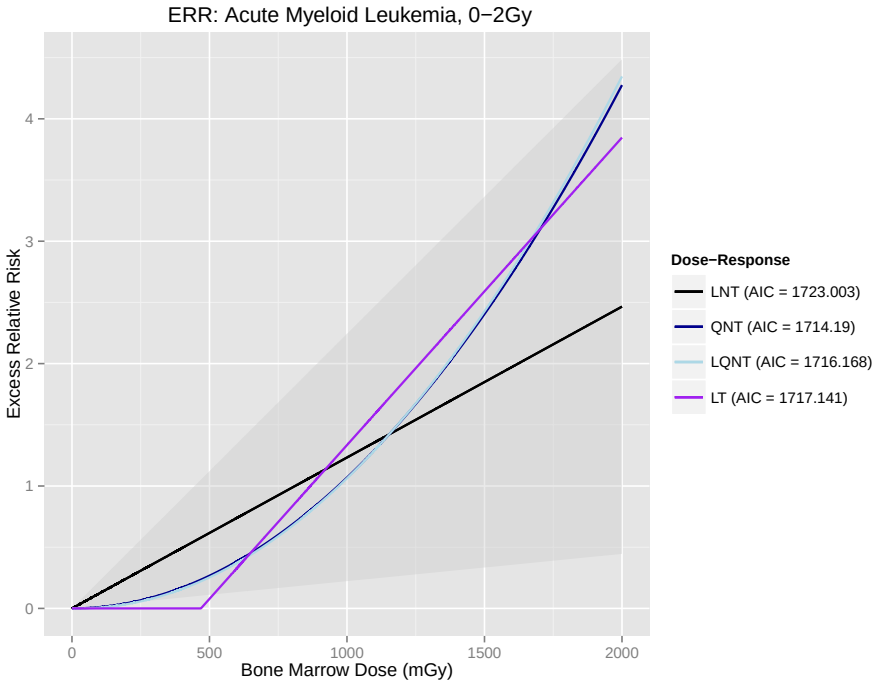

Figure 6.2: AML - LNT, QNT, LQNT, LT

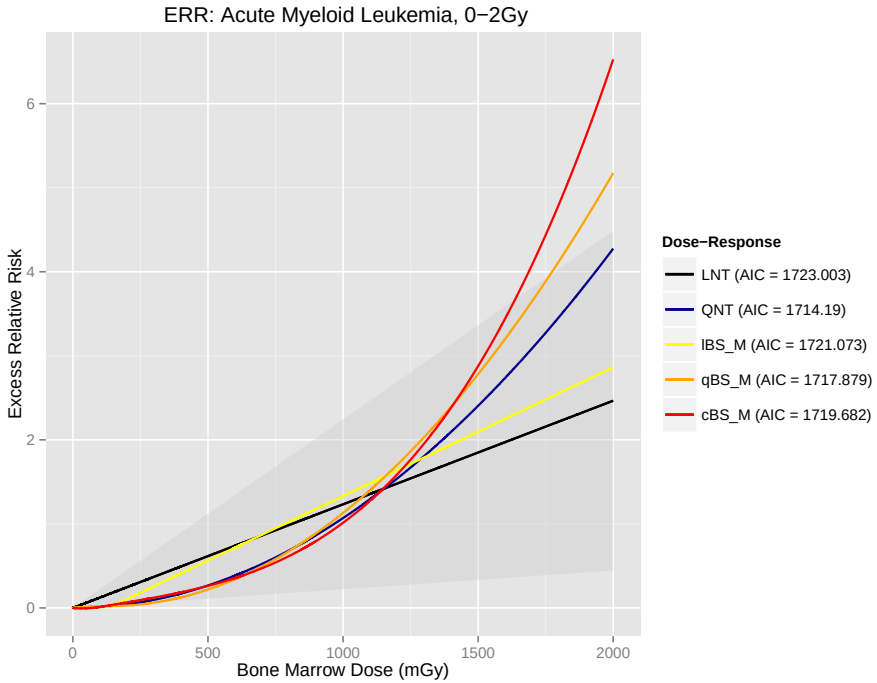

Figure 6.3: AML - Splines - Knot at Median

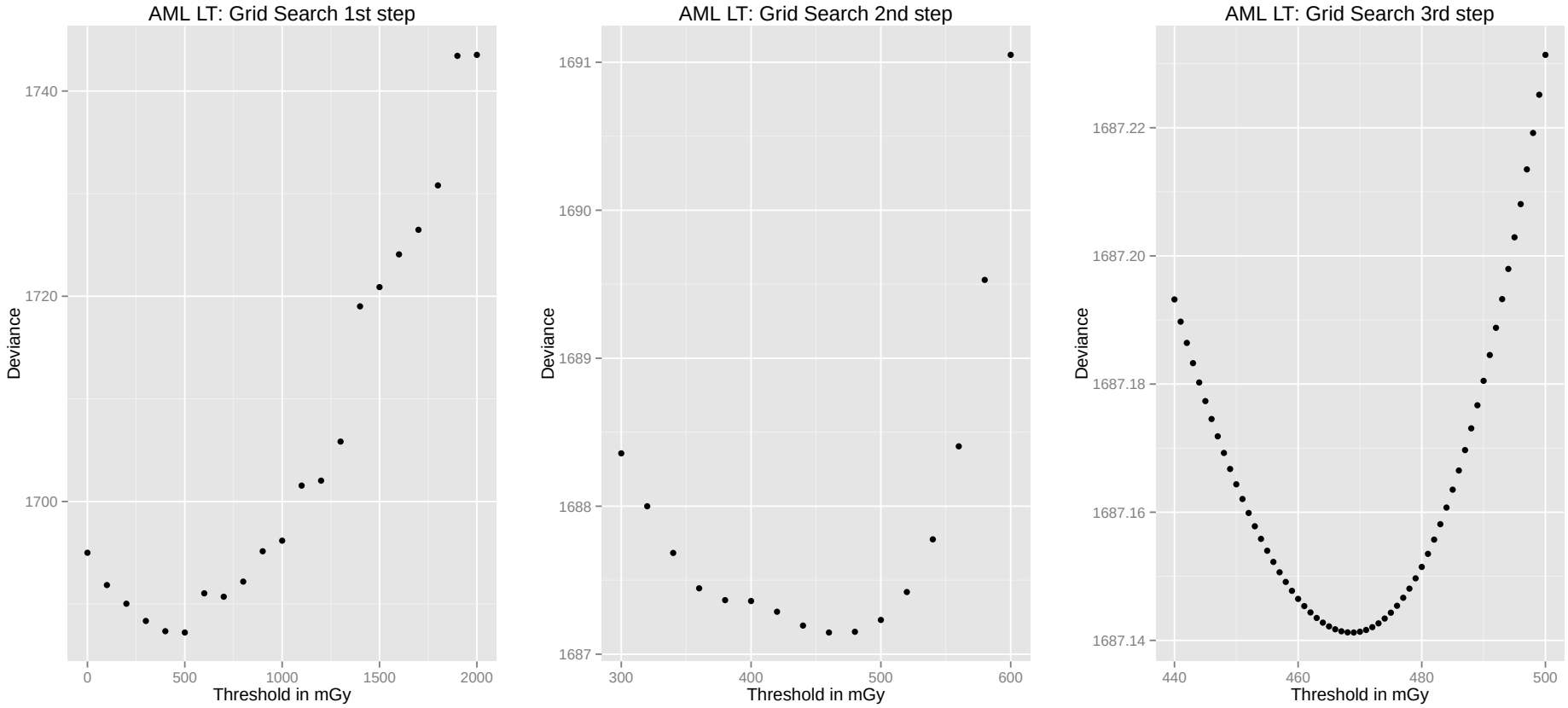

Figure 6.4: AML - Grid Search Linear Threshold

As stated in subsection 3.3 spline modelling is used to allow a flexible dose-response function. Figure 6.3 shows dose-response functions and AICs for splines of degree one to three with one fixed knot at the median (139 mGy).
The spline of degree 1 with a knot at the median is not able to fit the data well. Second and third degree splines do have a slightly better fit than QNT but need 2 and 3 parameters more to achieve that fit respectively. When comparing these models via AIC QNT is still best.
Splines with additional knots at the quartiles (29 and 562 mGy) of the bone marrow dose distribution do not change the form of the estimated dose-response function (figure 6.5). The necessary additional parameters are not - according to AIC - justified by the gains in likelihood.
The spline of degree 1 with one free knot (figure 6.6) yields a lower likelihood than the quadratic no-threshold model but is also not preferred by AIC. The complete grid search procedure can again be retraced at figure 6.7. Trying to estimate the knot position directly could have led to convergence at a local deviance minimum at around 500 Gy - a problem that can in general be avoided by grid-searching carefully for maximum likelihood estimates of knot positions.
As depicted in figures 6.3 and 6.5 all higher degree splines show a dose-response function very similar to QNT. It seems that QNT describes the data very well even after restricting the dose range to 2 Gy.

Via simulation (see section 5) it was tested whether or not a significant dose effect exists. The likelihood-ratio (lr) when testing the baseline model (H_0) against the column model (H_1), the 95% quantile of the $\mathcal{X}^2$ and the simulated distribution as well as the p-value of the (in section 5 discussed) KS-test are given in table 6.2. The differences in deviance of QNT - the preferred model by AIC - and the baseline model is larger than the 95% quantile of both the simulated and the $\mathcal{X}^2$ distribution. The dose effect is significant. Using LNT without effect modifiers (LNT-noEM) to test for significant differences in likelihood - like Hsu et al. (2013) and Preston et al. (2007) - does not change this result.
The 95% quantiles of the simulated dose distributions and the $\mathcal{X}^2$ distributions differ strongly. The Kolmogorov-Smirnov test confirms that the simulated likelihood-ratios are probably not random values from the corresponding $\mathcal{X}^2$ distribution even for the models without effect modifiers. This is an indicator that the $\mathcal{X}^2$ assumption is not always appropriate even for nested models.

	LNT	LNT-noEM	QNT	QNT-noEM
lr	48.52	35.85	57.33	42.82
$\mathcal{X}^2$ 95%	9.49	3.84	9.49	3.84
Simulated 95%	14.24	4.30	14.31	4.03
KS-test p=	<0.001	0.016	<0.001	0.004

Table 6.2: AML - Likelihood-Ratio Tests vs. Baseline

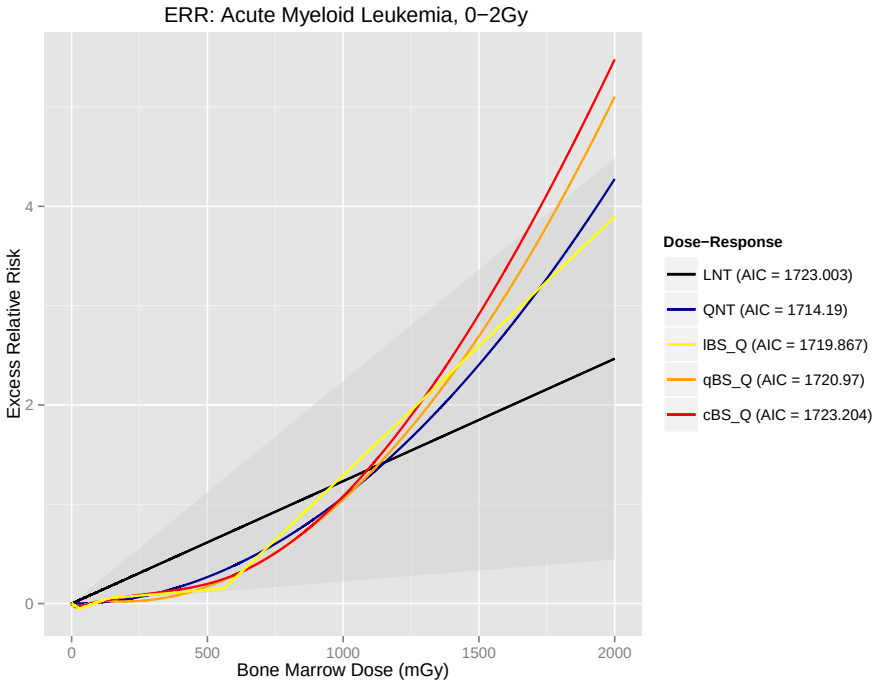

Figure 6.5: AML - Splines - Knots at Quartiles

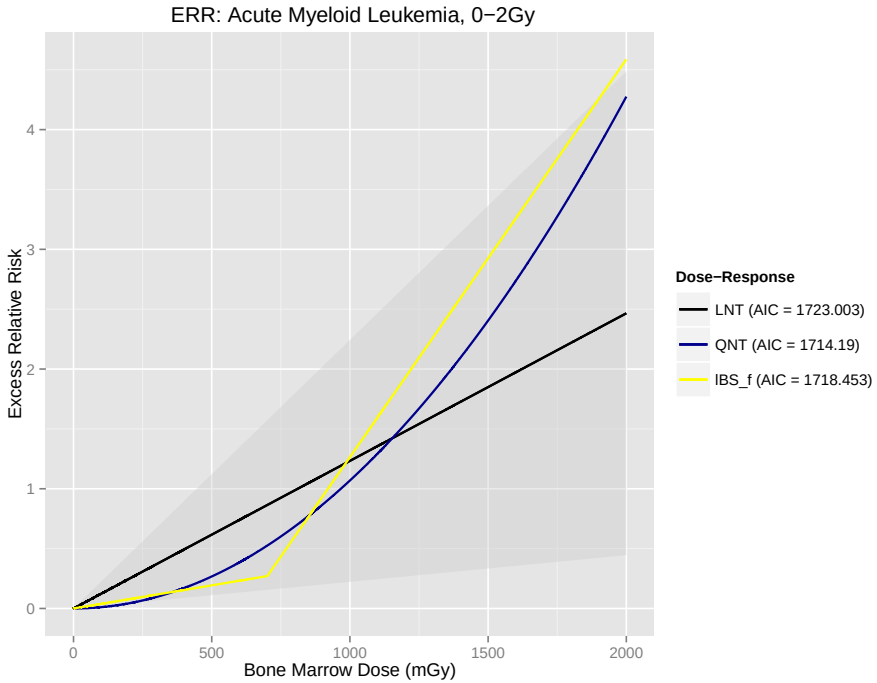

Figure 6.6: AML - Splines - Free Knot

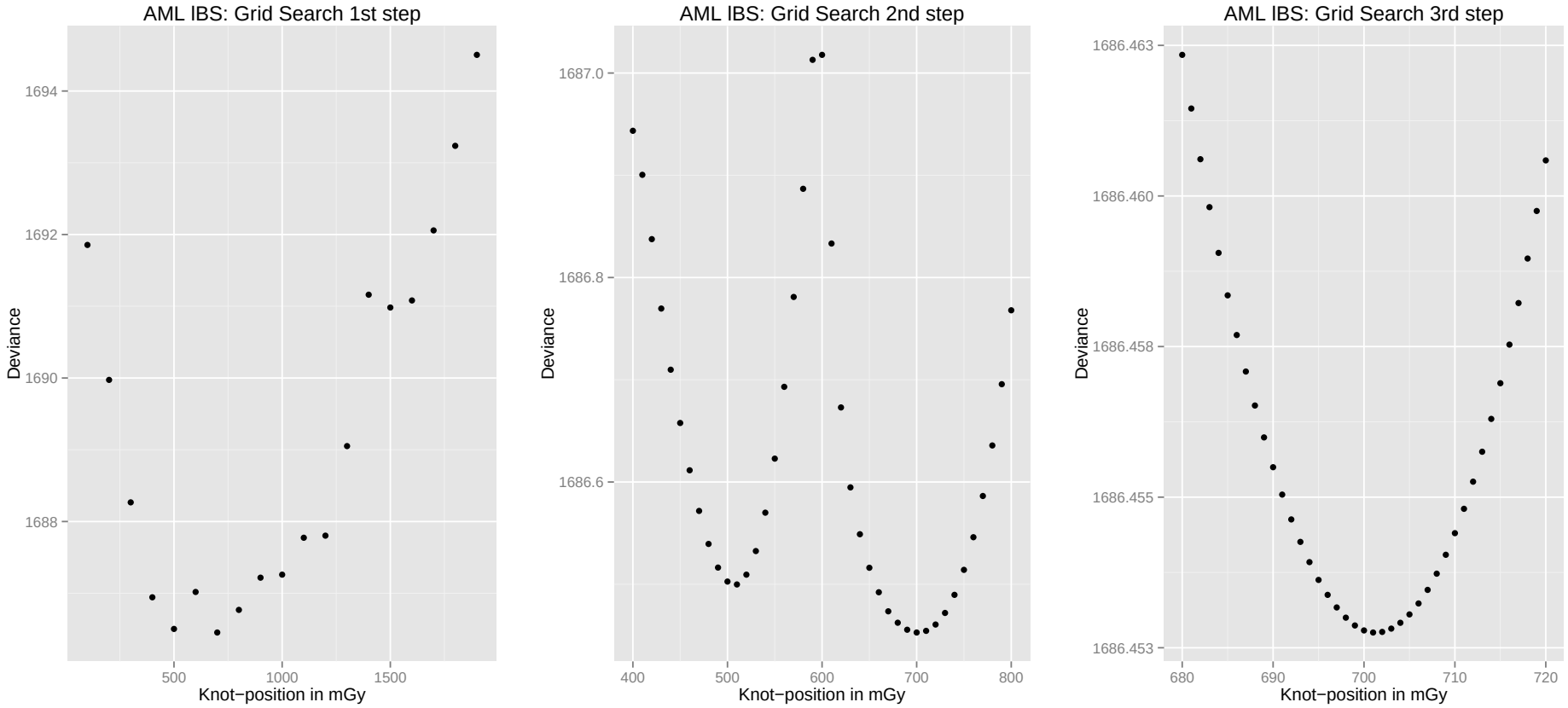

Figure 6.7: AML - Grid Search B-Spline degree 1 (1 knot)

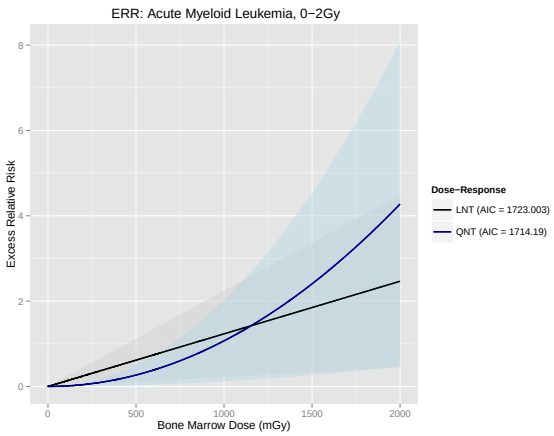

Figure 6.8: AML - 95% CI LNT QNT

	Deviance	#Par.	AIC
Baseline	1743.524	10	1763.524
LNT noEM	1707.676	11	1729.676
QNT noEM	1700.704	11	1722.704
LNT	1695.003	14	1723.003
QNT	1686.190	14	1714.190
LQNT	1686.168	15	1716.168
LT	1687.141	15	1717.141
lBS_M	1691.073	15	1721.073
qBS_M	1685.879	16	1717.879
cBS_M	1685.682	17	1719.682
lBS_Q	1685.867	17	1719.867
qBS_Q	1684.970	18	1720.970
cBS_Q	1685.204	19	1723.204
lBS_f	1686.453	16	1718.453

Table 6.3: AML - Model Comparison

6.2 Testing for Non-Linearity

When comparing the different models via AIC (see table 6.3) the best model is easily the quadratic no-threshold model (with effect modification).
Figure 6.8 depicts the symmetric 95% standard confidence intervals for LNT and QNT. For doses below 123 mGy the confidence intervals do not overlap. This is again a sign for non-linearity in the dose-response.[11]

A lot of the other more complex models still are preferred by AIC compared to LNT. Even though the low likelihood of LNT suggests that it is not the true model this does not lead to the conclusion that the other models are *significantly* better than LNT. When comparing LNT with the models it is nested in and assuming that asymptotic properties hold one can use the $\mathcal{X}^2$ distribution to see if differences in likelihood are significant. Considering the problematic presented in section 5 and the results above (see table 6.2) testing if the asymptotic properties hold is recommended. Therefore, the simulated distribution of the likelihood-ratio test statistic under H_0 is compared with the commonly assumed $\mathcal{X}^2$ distribution.

	QNT	LQNT	LT	lBS_M	qBS_M	cBS_M	lBS_Q	qBS_Q	cBS_Q	lBS_f
lr	8.81	8.83	7.86	3.93	9.12	9.32	9.14	10.03	9.80	8.45
$\mathcal{X}^2$ 95%	—	3.84	3.84	3.84	5.99	7.81	7.81	9.49	11.1	5.99
Simulated 95%	2.64	4.59	3.79	4.55	7.63	9.64	9.59	12.62	14.19	8.76
KS-test p=	—	0.09	<0.001	0.004	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001

Table 6.4: AML - Likelihood-Ratio Tests vs. LNT

There is strong evidence that the $\mathcal{X}^2$ assumption is not appropriate. The 95% quantiles of the simulated distribution and the corresponding $\mathcal{X}^2$ distribution are with a few exception rather far away from each other. The Kolmogorov-Smirnov test p-values show that it is highly improbable for the simulated lrs under H_0 to be random variables from a $\mathcal{X}^2$ distribution.

[11] The assumption used for symmetric standard confidence intervals that the standard error is normally distributed is - like the $\mathcal{X}^2$ assumption used by confidence intervals based on the profile likelihood - probably not met either. It follows that the estimated confidence intervals should be interpreted cautiously. This applies as well for the confidence interval estimates tabulated in the appendix.

The test decision in most cases would have been different with the $\mathcal{X}^2$ assumption.
In the test of LNT against LQNT the simulated likelihood-ratios could be - according to the KS-test - random variables from the $\mathcal{X}^2$ distribution with 1 degree of freedom. Despite this result the 95% quantiles differ heavily. Given that enough simulation runs are provided the simulated distribution should be used for testing. A more conservative approach would be to use the higher quantile as the critical value.
Linear threshold (LT) and the second degree spline with one knot at the median of the dose distribution (qBSM) still are significantly better than LNT and the likelihood-ratio of QNT and LNT also provides significant evidence against LNT. There is clear evidence that the true functional form of the dose-response on excess relative acute myeloid leukemia risks is not linear without a threshold.

Table 6.5 yields that there is no significantly better model in this analysis than QNT. For most of the tests the $\mathcal{X}^2$ distribution of lr under H_0 can not be assumed because QNT is only nested in LQNT and splines of degree 2 and above. In these cases the asymptotic properties do not hold either.

	LNT	LQNT	LT	lBS_M	qBS_M	cBS_M	lBS_Q	qBS_Q	cBS_Q	lBS_f
lr	-8.81	0.02	-0.95	-4.88	0.31	0.51	0.32	1.22	0.99	-0.26
$\mathcal{X}^2$ 95%	—	3.84	—	—	5.99	7.81	—	9.49	11.1	—
Simulated 95%	1.97	4.77	5.46	5.13	7.05	8.85	8.95	11.01	12.85	7.89
KS-test p=	—	<0.001	—	—	<0.001	<0.001	—	<0.001	<0.001	—

Table 6.5: AML - Likelihood-Ratio Test vs. QNT

A simulation with a 100 times higher expected number of events for each stratum (new case counts are drawn from $\text{Po}(\hat{\lambda} \cdot PYR \cdot 100)$) leads to a high probability that the simulated lrs are from the $\mathcal{X}^2$ distribution. It seems that the asymptotic properties of lr under H_0 depend on the number of events and not only on the number of stratification groups.
This simulation provides strong indication that the likelihood-ratio test statistic for linear threshold (LT) and the first degree spline with a free knot (lBS_f) as H_1 do not asymptotically follow the $\mathcal{X}^2$ distribution.

	LQNT	LT	lBS_M	qBS_M	cBS_M	lBS_Q	qBS_Q	cBS_Q	lBS_f
$\mathcal{X}^2$ 95%	3.84	3.84	3.84	5.99	7.81	7.81	9.49	11.1	5.99
Simulated 95%	3.88	2.85	3.76	5.91	8.91	7.85	9.82	11.30	7.68
KS-test p=	0.08	<0.001	0.55	0.29	0.10	0.37	0.26	0.83	<0.001

Table 6.6: AML - Likelihood-Ratio Test vs. LNT, $\lambda \cdot PYR_i \cdot 100$

7 Chronic Myeloid Leukemia

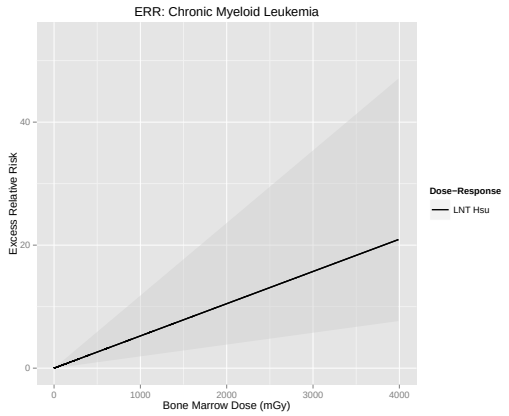

Figure 7.1: CML - in Hsu et al.(2013)

Dose	PYR	CML Cases
-0.005	2,039,095	22
-0.1	957,889	17
-0.2	201,935	2
-0.5	206,749	11
-1	117,855	6
-2	64,123	9
2+	25,761	8
Total	3,613,406	75

Table 7.1: PYR & Cases by Dose

Chronic myeloid leukemia (abbreviated by CML) is more than twice as rare in the LSS cohort as acute myeloid leukemia - only 75 cases were registered. As in the analysis of AML the ionizing radiation dose to the bone marrow is used. Therefore, by constraining the dose to 2 Gy, the same number of person-years at risk is lost: 25.761. After the exclusion 67 cases of chronic myeloid leukemia remain.

According to Hsu et al. (2013: p.370) the baseline risk depends on sex and attained age while the effect of the latter differs for each sex. Baseline risks for men are higher than for women until age 75. After age 75 - due to the stronger effect of attained age for women - the baseline risk is higher for women. The preferred baseline model in Hsu et al. (2013) is:

$$\begin{aligned}\lambda_0(X_{BL})(\text{c, s, a, e}) = \exp \quad & (\beta_0 \\ & + \beta_1 \text{female} && (1) \\ & + \beta_2 \text{lage70} && (2) \\ & + \beta_3 \text{flage70} && (3) \\ & + \beta_4 \text{nic}\cdot\text{hiro} + \beta_5 \text{nic}\cdot\text{naga}) && (4)\end{aligned}$$

While (1) denotes the effect of sex, (2) the effect of log-attained age for men and (2)+(3) the age effect for women. As in every baseline model of Hsu et al. (2013) the not-in-city effects are used to be able to include cohort members who not in the city at the time of the bombings.

The preferred dose-response model in Hsu et al. (2013: p.370) is LNT. The preferred effect modification model is:

$$\epsilon(X_{EM}) = exp(\gamma_1 \cdot naga + \gamma_2 \cdot lage55 + \gamma_3 \cdot ltsx25 + \gamma_4 \cdot over4gy)$$

Figure 7.1 shows the dose effect of the preferred model (LNT) and its 95% confidence interval. The estimated excess relative risk is 5.24 per Gy for Hiroshima residents age 55 and 25 years since exposure. The excess risk is estimated to be less than a fourth (γ_1 = -1.50) for Nagasaki compared to Hiroshima residents. It is lower for both higher attained age (γ_2 = -1.42) and increasing time since exposure (γ_3 = -1.59). (cf. Hsu et al. 2013)

The reasoning behind the fourth parameter γ_4 is equivalent to the reasoning in section 6.

7.1 Dose-Effect Estimation

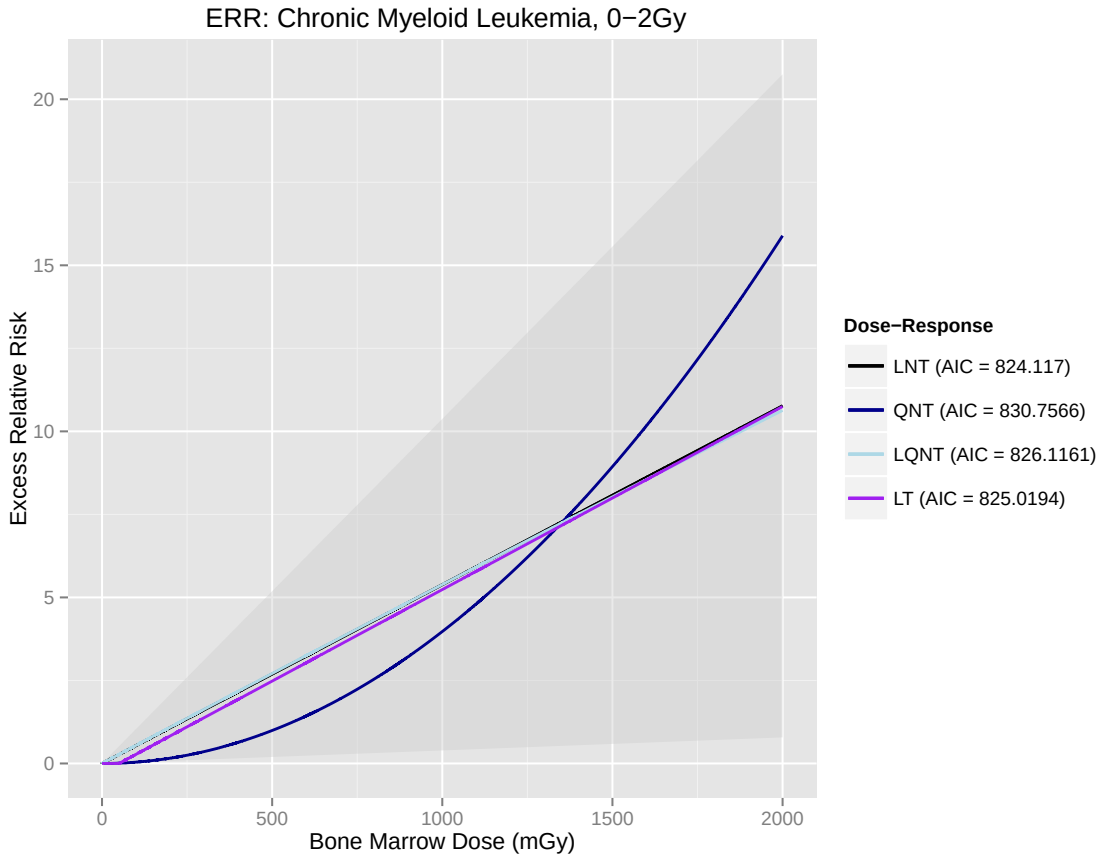

Figure 7.2: CML - LNT, QNT, LQNT, LT

Excluding strata groups with radiation doses higher than 2 Gy again returns no immediate differences compared to the results in Hsu et al. (2013). The AIC still prefers LNT and estimated excess risks stay about the same at 5.3 per Gy (symmetric 95% CI [0.393;10.375]) standardized for Hiroshima residents age 55, 25 years after exposure with similar estimates for the effect modifying variables. Adding a quadratic term yields almost no likelihood gains. Dose-response plots of LNT and LQNT are almost identical.
The best threshold is estimated to be at 50 mGy. This threshold yields a to small increase to be preferred by AIC compared to LNT. The grid search can be retraced with figure 7.3.

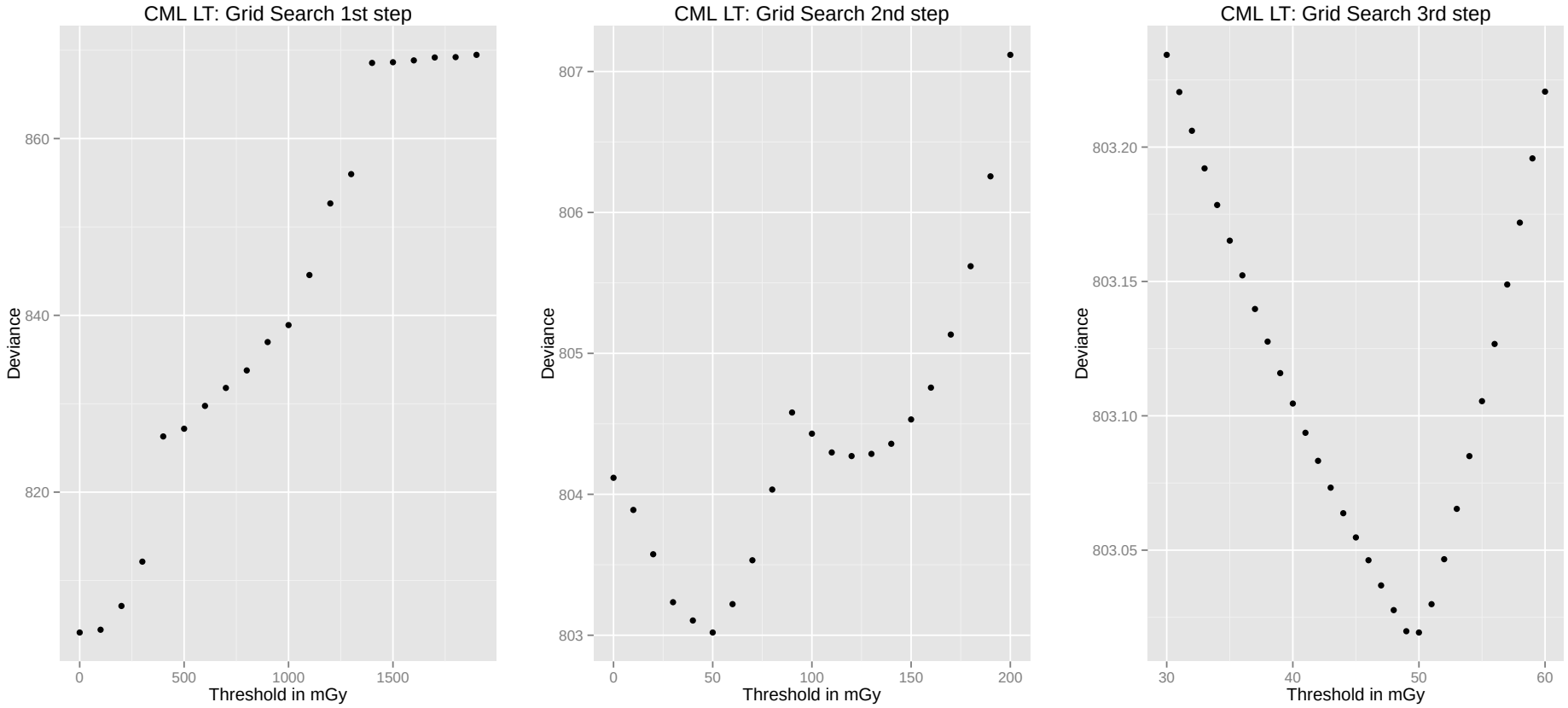

Figure 7.3: CML - Grid Search Linear Threshold

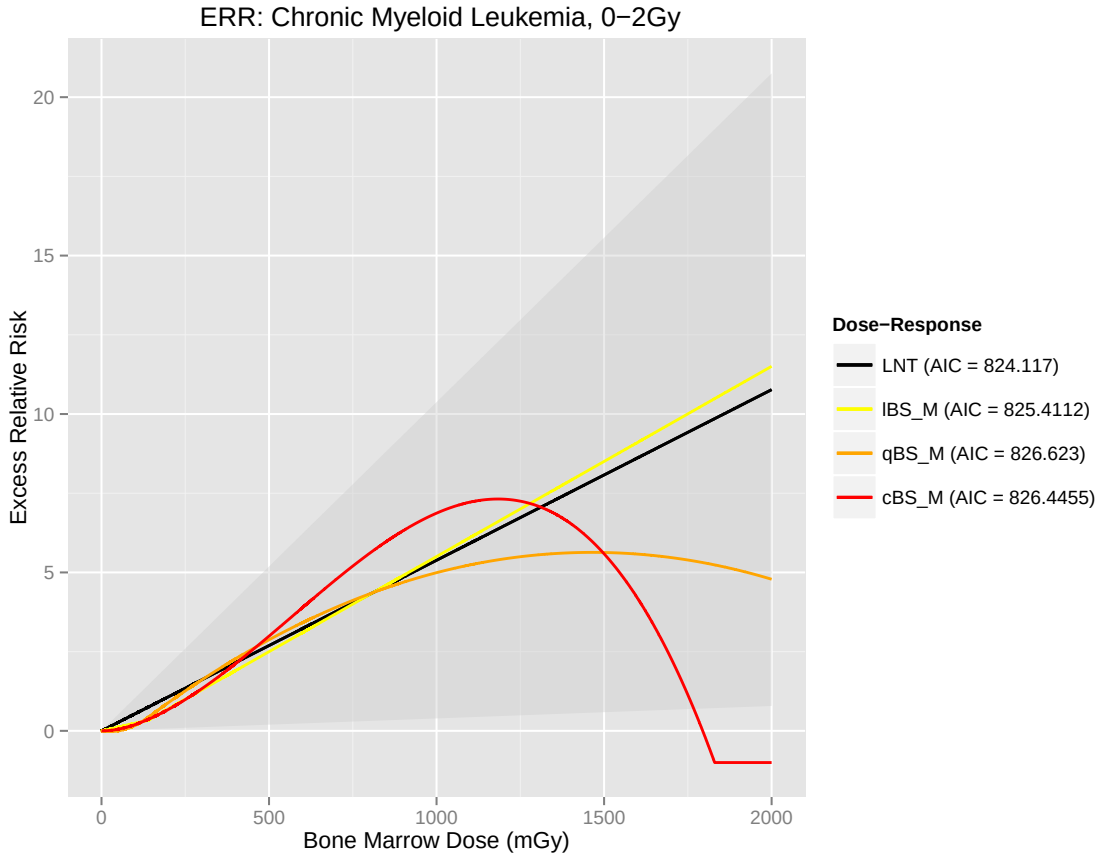

Figure 7.4: CML - Splines - Knot at Median

Figure 7.4 show splines of degree one to three with one knot at the median (139 mGy) of the dose distribution. The third degree spline seems to overfit the data. It is flexible enough to estimate a risk of zero ($err(\cdot) = -1$) for doses higher than about 1.8 Gray (for Hiroshima residents age 55, 25 years after exposure), because there are zero cases of chronic myeloid leukemia for individuals with a dose of 1.4 to 2 Gy. Even though one can safely assume that this is not the true dose-response model - radiation at doses over 1750 Gy surely does not erase the risk for chronic myeloid leukemia completely - it leads to a considerable increase in likelihood compared to LNT.

The same overfitting can be observed in the estimates of the degree 2 and 3 splines with additional knots at the quartiles (29 and 562 mGy) of the bone marrow dose distribution (see figure 7.5) and even more so in the degree one spline with one free knot (see figure 7.6). The likelihood maximizing knot is at 1.4 Gy - the grid search can be retraced in figure 7.7. It is the only model that has a better AIC value than LNT. However, all four models would be discarded for biological reasons.

Additionally the hessian matrices of both the degree 1 spline with a free knot and the degree 3 spline with knots at the quartiles (cBS_Q) are not positive definite. This is generally a sign that either the estimated model is not the maximum likelihood estimate or that the likelihood function value in the proximity of the estimate is equal to the value at the estimate (cf. Fahrmeir et al. 2013: p.637). At least for the spline with a free knot the latter is probably the reason for the problems with the hessian. After 1.4 Gy the dose-response function does

not change if the second parameter of the spline gets even lower - the excess relative risk after 1.4 Gy is estimated to be -1 anyway. The result is that the likelihood will not be different if the estimate changes from α_2=-20,619.52 to α_2=-30,000. Trying to optimize the likelihood function of cBS_Q with varying starting values also yields that similar parameter estimates yield the same function value as the model presented in figure 7.5. The parameter with the most variation for all likelihood maximizing estimates are the 5th and 6th parameter of the spline. Therefore, cautious interpretation of the spline at the higher doses is advised.
A model with no strict optimum achieving a very high (compared to LNT) likelihood nonetheless can still be a sign for - and can still be used for testing for - the existence of a dose-effect or non-linearity in the dose-response.

The difference in likelihood of LNT compared to the baseline model is very high. Likelihood-ratio testing (see table 7.2) yields that there is a significant dose effect, either if LNT with or LNT without effect modifiers (LNT-noEM) is used as the H_1 model. In both tests the $\mathcal{X}^2$ assumption would not have been appropriate - the simulated likelihood-ratios are probably not random values from the respective $\mathcal{X}^2$ distributions.

	LNT	LNT-noEM
lr	65.39	32.68
$\mathcal{X}^2$ 95%	9.49	3.84
Simulated 95%	14.02	4.29
KS-test p=	<0.001	0.022

Table 7.2: CML - Likelihood-Ratio Tests vs. Baseline

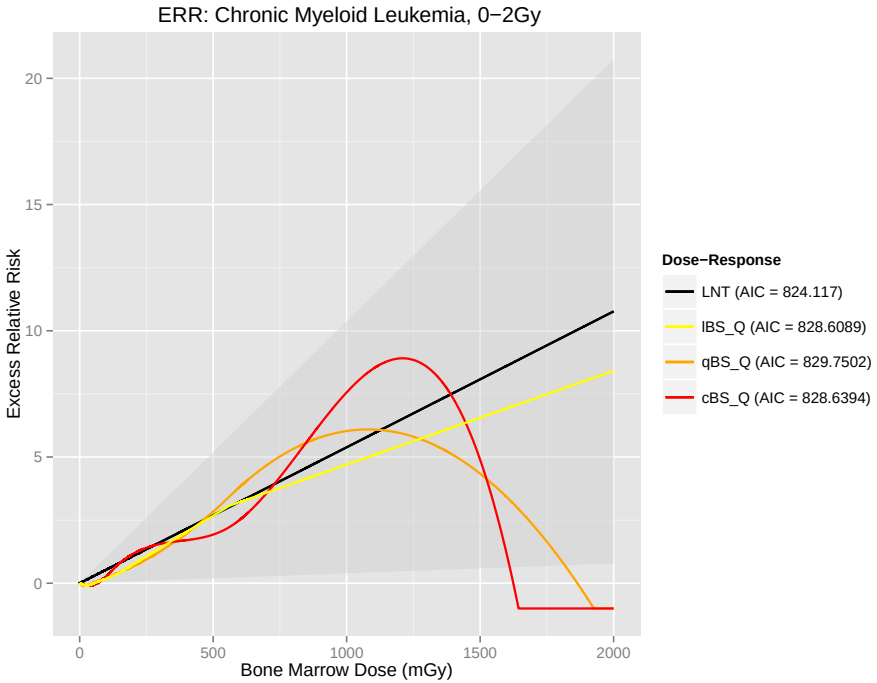

Figure 7.5: CML - Splines - Knots at Quartiles

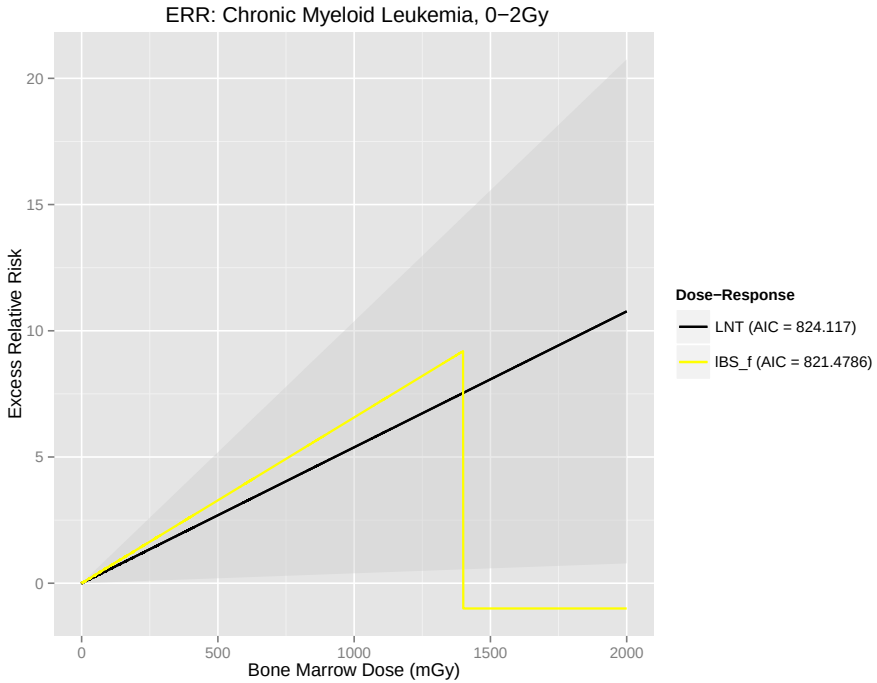

Figure 7.6: CML - Splines - Free Knot

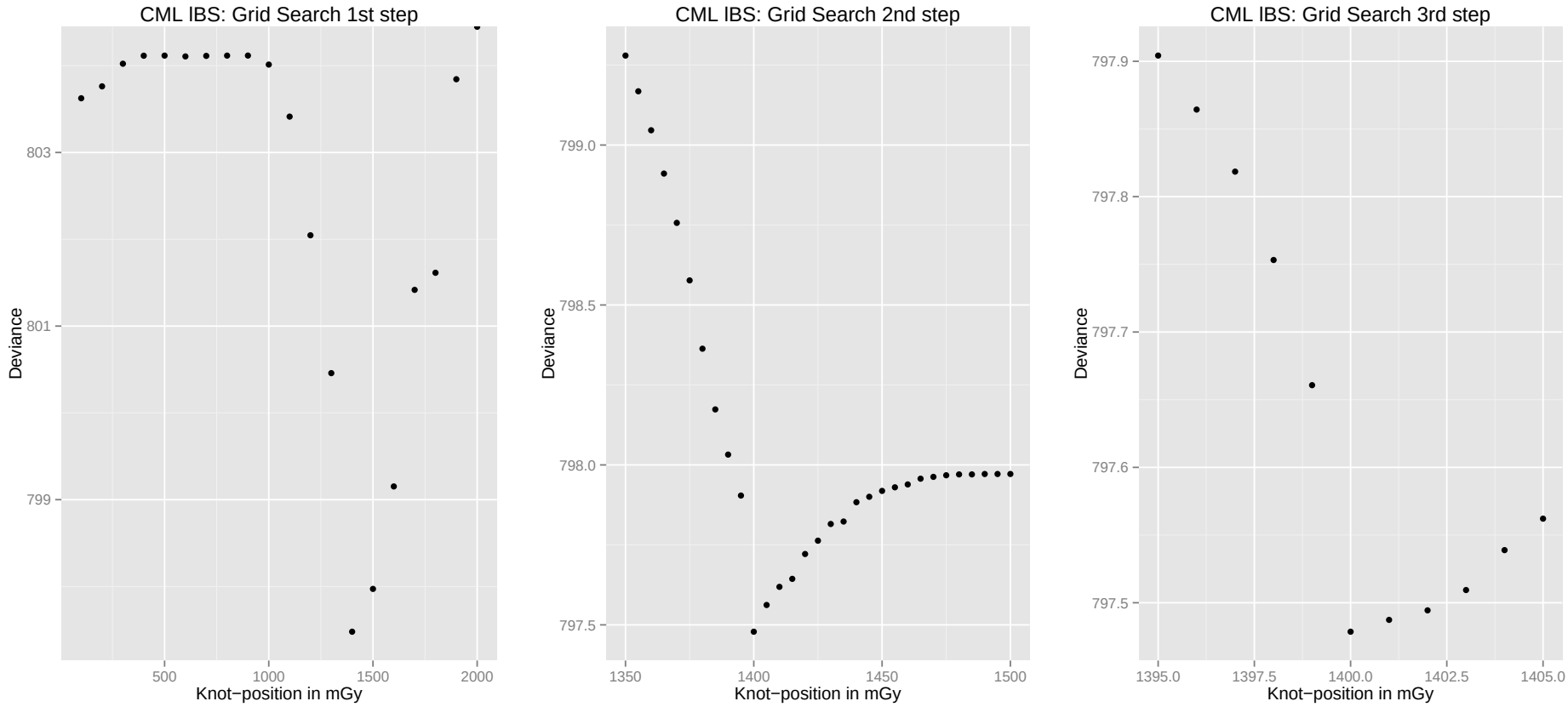

Figure 7.7: CML - Grid Search B-Spline degree 1 (1 knot)

7.2 Testing for Non-Linearity

	Deviance	#Par.	AIC
Baseline	869.5117	6	881.5117
LNT noEM	836.8291	7	850.8291
LNT	804.117	10	824.117
QNT	810.7566	10	830.7566
LQNT	804.1161	11	826.1161
LT	803.0194	11	825.0194
lBS_M	803.4112	11	825.4112
qBS_M	802.623	12	826.623
cBS_M	800.4455	13	826.4455
lBS_Q	802.6089	13	828.6089
qBS_Q	801.7502	14	829.7502
cBS_Q	798.6394	15	828.6394
lBS_f	797.4786	12	821.4786

Table 7.3: CML - Model Comparison

Compared to LNT the only other good models (according to the AIC) are those that seem to overfit the data at high doses. The highest increase in likelihood per extra parameter is achieved in the degree 1 spline with a free knot. Even though the higher degree spline models also lead to a small increase in likelihood, it is the only model that has a better AIC than LNT. The likelihood-ratios of the different models compared to LNT (H_0) can be viewed in table 7.4.

	QNT	LQNT	LT	lBS_M	qBS_M	cBS_M	lBS_Q	qBS_Q	cBS_Q	lBS_f
lr	-6.64	0.00	1.10	0.71	1.49	3.67	1.51	2.37	5.48	6.61
$\mathcal{X}^2$ 95%	—	3.84	3.84	3.84	5.99	7.81	7.81	9.49	11.1	5.99
Simulated 95%	0.96	3.94	3.36	4.22	6.64	8.81	8.67	11.10	13.22	7.11
KS-test p=	—	0.114	<0.001	0.36	0.067	<0.001	<0.001	<0.001	<0.001	<0.001

Table 7.4: CML - Likelihood-Ratio Tests vs. LNT

There are no signs for significant non-linearity in the dose-response. When assuming that the *lr* of LNT and the linear spline with a free knot follows - under H_0 - a $\mathcal{X}^2$ distribution the spline would be significantly better than LNT. Simulating this distribution yields that the $\mathcal{X}^2$ assumption is not correct ($p<0.001$) and the test decision would have been wrong. The spline does not provide a significantly better fit than LNT.
Kolmogorov-Smirnov testing shows again that - at least when testing very flexible models against LNT - the simulated likelihood-ratios are not random values from the opposing $\mathcal{X}^2$ distribution. For LQNT, lBS_M and qBS_M - the decision if the $\mathcal{X}^2$ distribution is appropriate is still disputable because 95% quantiles differ even though the Kolmogorov-Smirnov test is not significant. Therefore, the more careful approach would still be to simulate the distribution - at least if enough simulated likelihood-ratios are estimated to minimize randomness in the approximated 95% quantile.

8 Female Breast Cancer

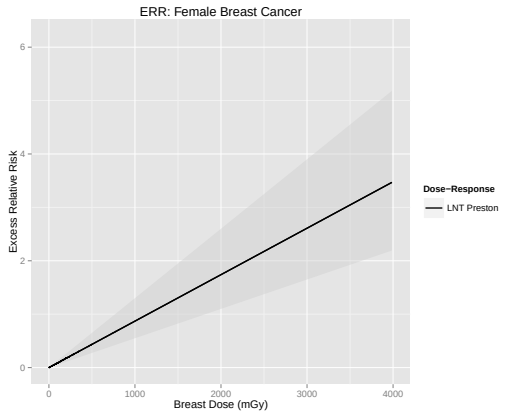

Figure 8.1: FBC - in Preston et al. (2007)

Dose	PYR	FBC Cases
-0.005	994,844	546
-0.1	426,129	217
-0.2	93,719	58
-0.5	111,197	104
-1	50,885	51
-2	34,115	60
2+	13,565	37
Total	1,724,454	1073

Table 8.1: PYR & Cases by Dose

With 1073 cases, female breast cancer (abbreviated by FBC) is one of the most common cancers in the LSS cohort. Even though compared to western populations breast cancer rates in Japan are considerably lower (cf. Preston et al. 2007: p.32). Excluding strata groups with doses above 2 Gy costs 13,565 person-years at risk and still leaves 1036 cases of breast cancer in women.

Preston et al. (2007: p.32sq) show that baseline breast cancer risks in the LSS strongly increase with age, while the increase is strongest before the menopause and the least strong between age 50 and 70. There has also been a large increase in breast cancer rates over the last 50 years in Japan. Therefore, a birth cohort effect on the baseline risk has to be allowed. The complete baseline model (ibid.) which is also used for analysis in this thesis is:

$$
\begin{aligned}
\lambda_0(X_{BL})(\text{c, s, a, e}) = \exp \quad & (\beta_0 \\
& + \beta_1 \text{naga} && (1)\\
& + \beta_2 \text{lage50preqsp} + \beta_3 \text{lage70} && (2)\\
& + \beta_4 \text{lage70sq} + \beta_5 \text{lage70qsp} && (3)\\
& + \beta_6 \text{e30} + \beta_7 \text{e30sq} && (4)\\
& + \beta_8 \text{nic}\cdot\text{hiro} + \beta_9 \text{nic}\cdot\text{naga}) && (5)
\end{aligned}
$$

While (1) is the city effect, (2)+(3) is a second degree truncated power series of the natural logarithm of age attained with two knots at age 50 and 70 and (4) a linear-quadratic function of the birth cohort.

There is strong evidence for a linear relation of ionizing radiation dose and excess relative breast cancer risks among women both in Preston et al. (2007) (0.87 excess risk per Gray; 90% confidence interval [0.55;1.3]; see Figure 8.1) as well as in numerous other data sets (e.g. the meta analysis of Preston et al. 2002). Preston et al. (2007) state that there is no statistically significant non-linearity even when limiting the dose range to 0-2 Gy (cf. Preston et al. 2007: p.33). The tests used were, however, based on $\mathcal{X}^2$ approximations to the likelihood-ratio test statistic distribution (cf. Preston et al. 2007: p.5).

The preferred effect modification model is:

$$\epsilon(X_{EM}) = exp(\gamma_1 \cdot e30 + \gamma_2 \cdot lage70)$$

Scientific consensus was that low age at exposure has a big impact on ionizing radiation related increases in breast cancer risks (cf. Ronckers et al. 2005). Controlling for attained age ($lage70$) leaves, however, only a non-significant effect modification of age at exposure ($e30$) in the LSS cohort. The age at exposure effect is included for theoretical reasons and is not a result of model selection.

8.1 Dose-Effect Estimation

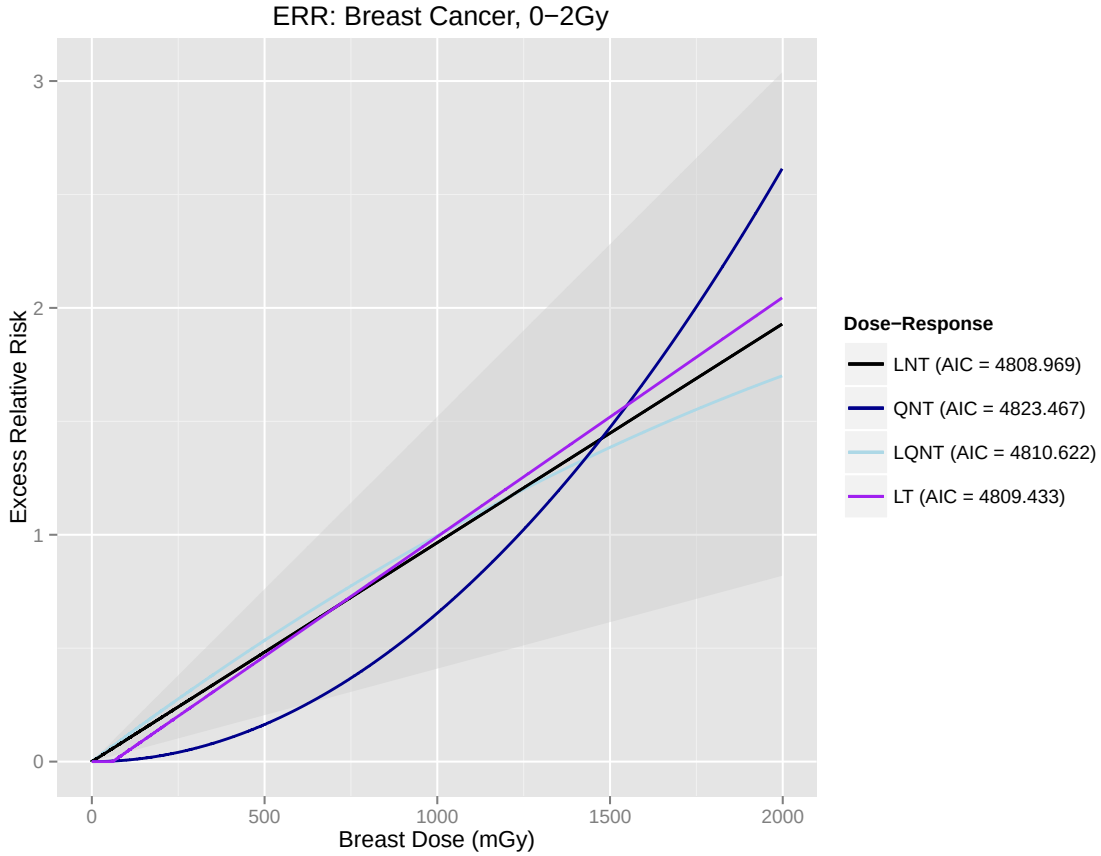

Figure 8.2: FBC - LNT, QNT, LQNT, LT

When compared with the other less complex models via AIC LNT is still best, when data is constrained to 0-2 Gy (cf. Figure 8.2). There is almost no gain in likelihood when adding a quadratic term. The estimated threshold is rather low at 60 mGy. Its inclusion yields no big increase in likelihood. The complete grid search procedure for the best threshold can again be seen in figure 8.3.[12]

[12] A fourth step was done to ensure that the estimate at 60 mGy is not a local optimum. See electronic appendix R-file breast_LT_2G.R

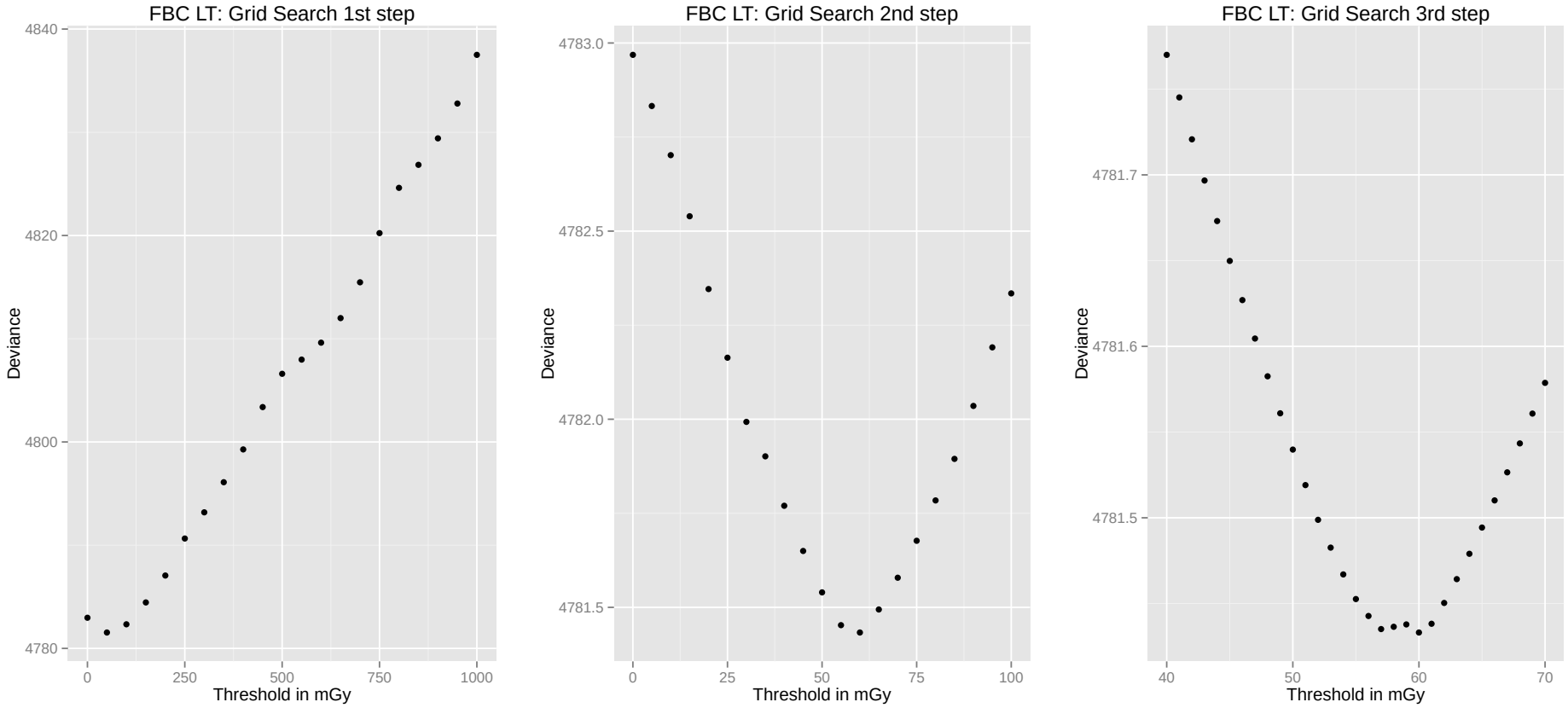

Figure 8.3: FBC - Grid Search Linear Threshold

In LNT the excess relative breast cancer risk is estimated to be 0.96 (symmetric 95% CI: [0.41;1.52]) per Gy for women who attained age 70 and were 30 at the time of the bombings. The age at exposure effect is - as in Preston et al. (2007) - very small. The impact of radiation decreases at increasing attained age.

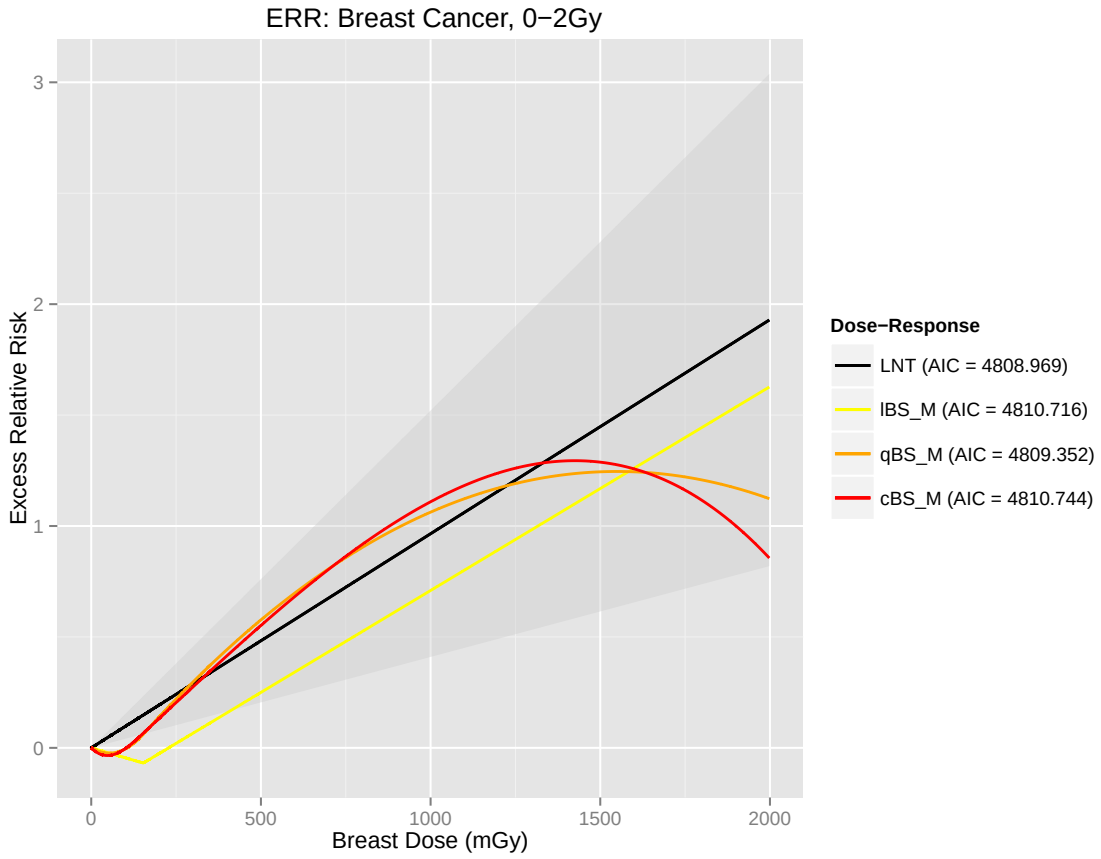

Figure 8.4: FBC - Splines - Knot at Median

As can be seen in figures 8.4, 8.5 and 8.6 all spline models (but the degree one splines with fixed knots) show some kind of effect saturation at doses above 1500 mGy. Excess relative risk estimates for very low doses are even negative in the more flexible models.
All models need too many parameters to compete with LNT when they are compared via AIC. The degree 2 spline with one knot at the median (154 mGy, quartiles are at 33 and 355 mGy) yields the most gains in likelihood (compared with LNT) per parameter. The second best model - and the only model that is AIC-wise rather close to LNT while having a different dose-response - is the second degree spline with one knot at the median.

A second degree spline model is, therefore, also estimated with a free knot. Grid search procedures for the free knots of both the degree one and the degree two splines can be retraced in figures 8.7 and 8.8 respectively. The knot in the degree 1 spline maximises the likelihood at 1395 mGy. When compared to both degree 1 splines with fixed knots via AIC the model with a free knot is preferred. This is not the case for the degree 2 spline because the maximum likelihood estimate of the free knot (83 mGy) is rather close to the median and yields, therefore, a very similar dose-response shape.

Likelihood-ratio testing (see 8.2) again shows a significant dose effect. However, unlike in the case of acute and chronic myeloid leukemia the simulated distribution of lr of the baseline model (H_0) and LNT without effect modification (H_1) does not differ significantly from the $\mathcal{X}^2$ distribution. When testing the more complex model LNT with effect modifiers against the baseline model the simulated distribution does differ significantly from the $\mathcal{X}^2$ distribution. Assuming a specific $\mathcal{X}^2$ distribution for the likelihood-ratio test statistic in this case is - as explained in section 5 - problematic anyway.

	LNT	LNT-noEM
lr	81.54	74.01
$\mathcal{X}^2$ 95%	10.33	3.84
Simulated 95%	14.02	3.76
KS-test p=	<0.001	0.177

Table 8.2: FBC - Likelihood-Ratio Tests vs. Baseline

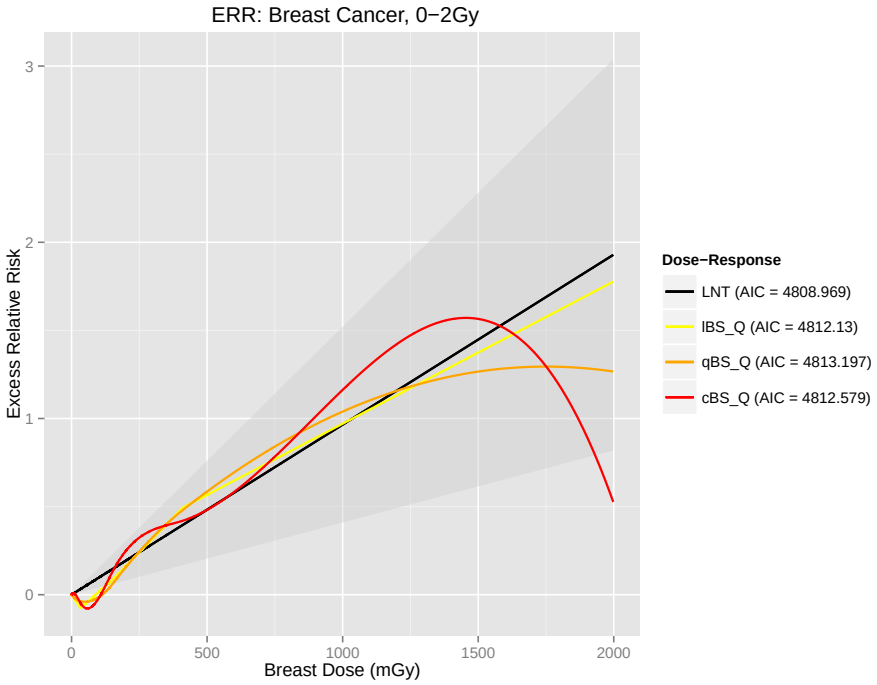

Figure 8.5: FBC - Splines - Knots at Quartiles

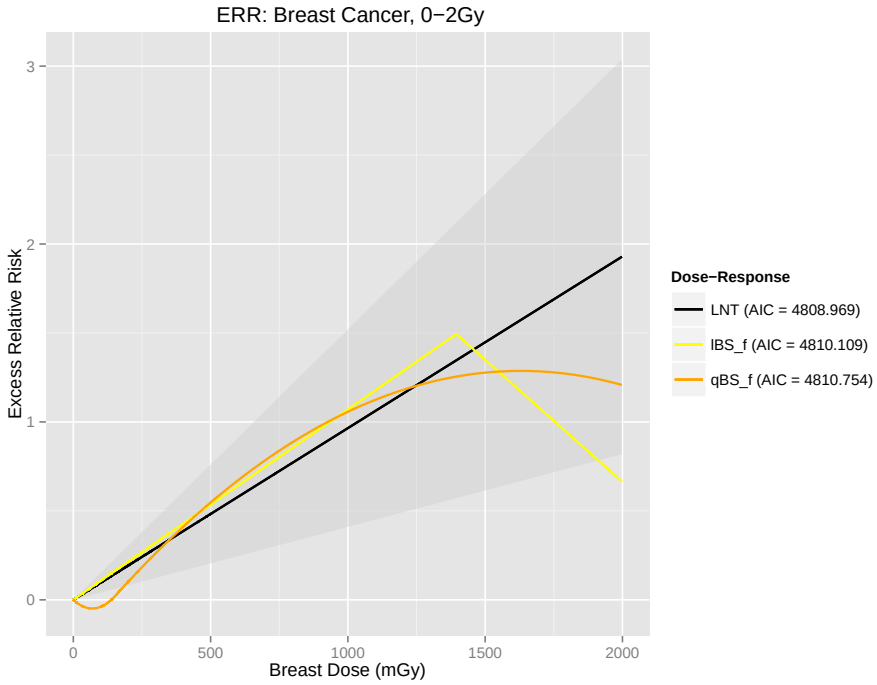

Figure 8.6: FBC - Splines - Free Knot

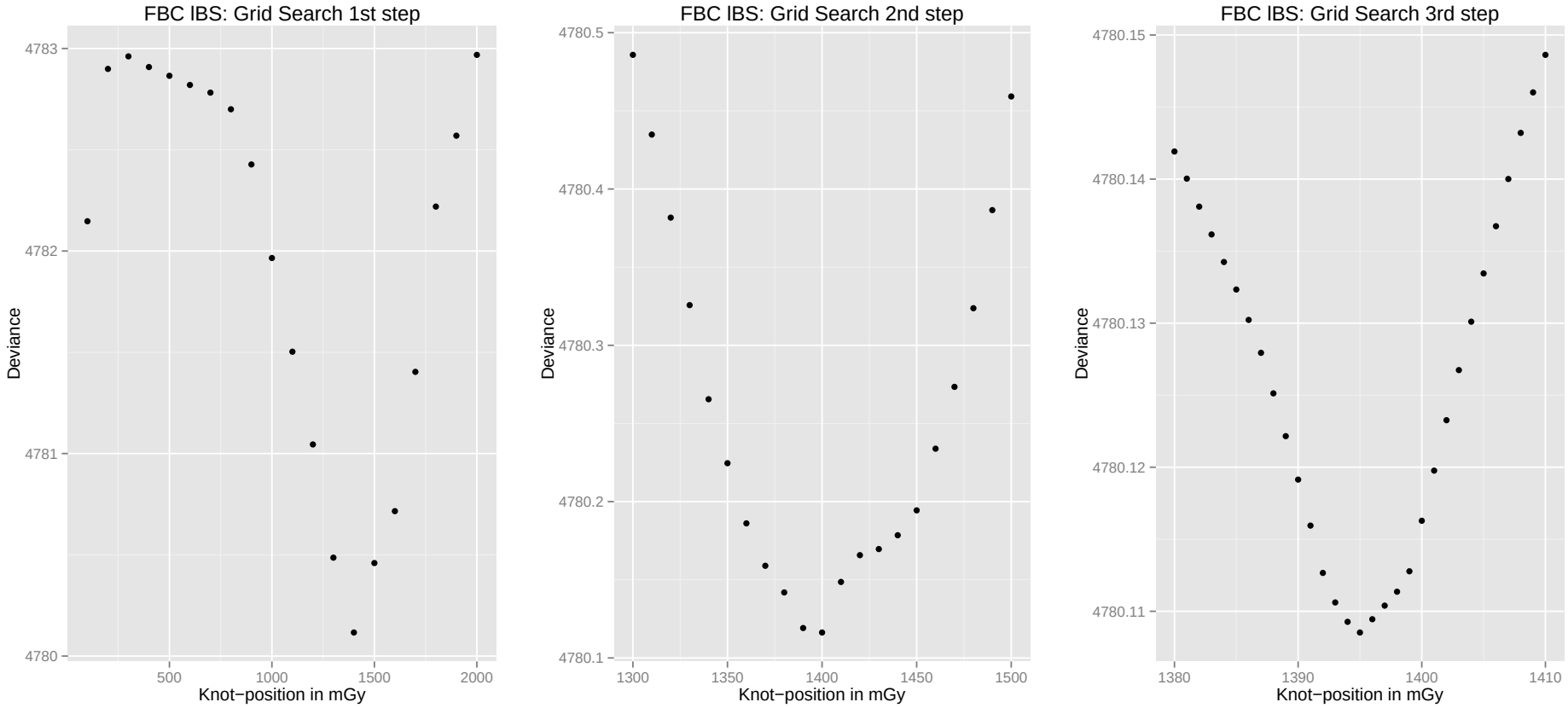

Figure 8.7: FBC - Grid Search B-Spline degree 1 (1 knot)

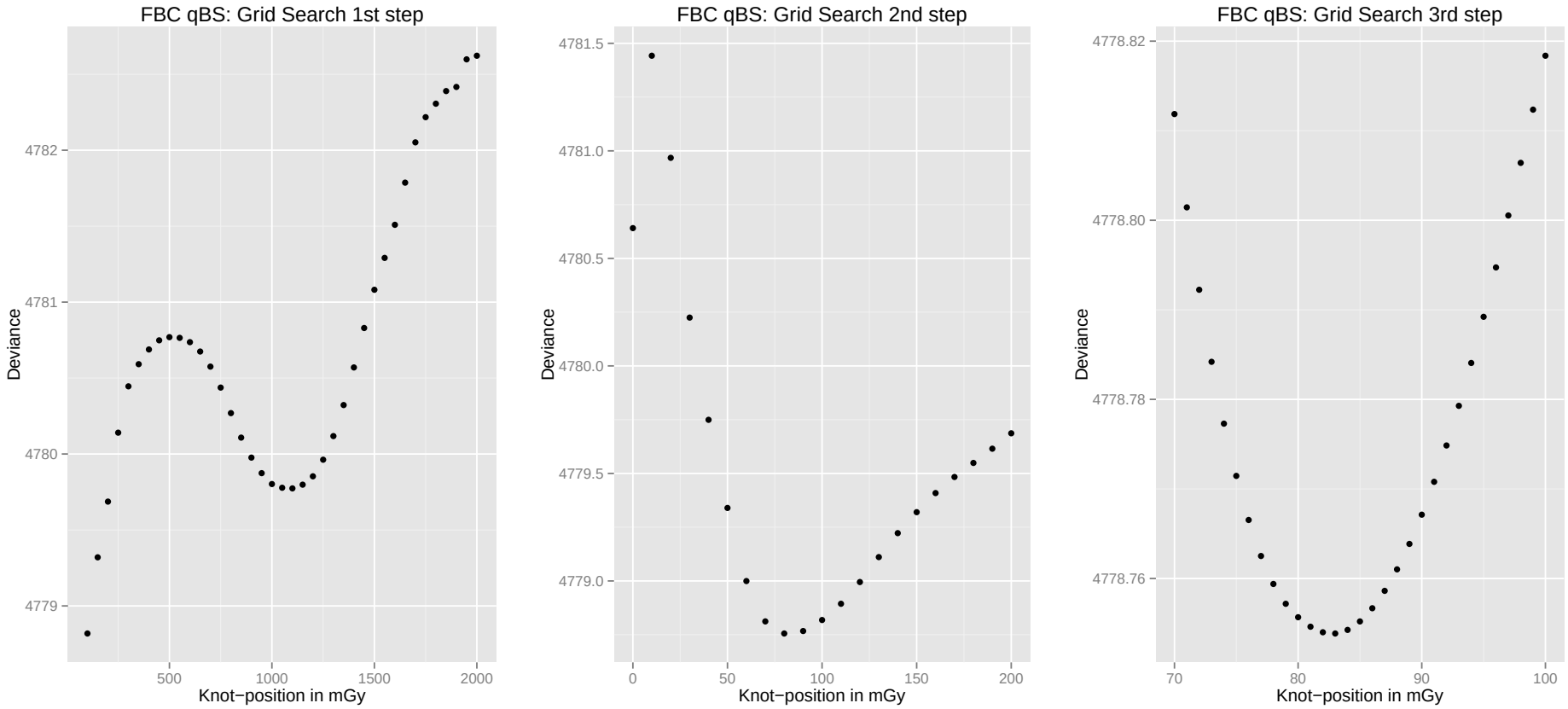

Figure 8.8: FBC - Grid Search B-Spline degree 2 (1 knot)

8.2 Testing for Non-Linearity

	Deviance	#Par.	AIC
Baseline	4864.508	10	4884.508
LNT noEM	4790.503	11	4812.503
LNT	4782.969	13	4808.969
QNT	4797.467	13	4823.467
LQNT	4782.622	14	4810.622
LT	4781.433	14	4809.433
lBS_M	4782.716	14	4810.716
qBS_M	4779.352	15	4809.352
cBS_M	4778.744	16	4810.744
lBS_Q	4780.130	16	4812.130
qBS_Q	4779.197	17	4813.197
cBS_Q	4776.579	18	4812.579
lBS_f	4780.109	19	4810.109
qBS_f	4778.754	15	4810.754

Table 8.3: FBC - Model Comparison

When comparing the different models via AIC, LNT is the best model (see table 8.3). Second best is the degree 2 spline with one knot at the median - naturally it is also not significantly better than LNT on a 95% confidence level either when assuming a $\mathcal{X}^2$ distribution of lr under H_0 (LNT) or when using the simulated distribution (see table 8.4). Simulating lr under H_0 does not yield drastically different 95% quantiles than the appropriate $\mathcal{X}^2$ distribution and the p-values of the Kolmogorov-Smirnov test (see table 8.4) show that it is plausible that the simulated lrs are random values from the $\mathcal{X}^2$ distribution. This holds at least for all models without free knots - which are in the setting of nested models problematic anyway (cf. section 5). These results are coherent with the lrs on simulated AML event counts when the expected counts are increased by a factor of 100 (subsection 6.2 on page 23). It seems that lr follows its asymptotic distribution ($\mathcal{X}^2$ with r degrees of freedom) for unproblematic nested models when event counts are high.

	QNT	LQNT	LT	lBS_M	qBS_M	cBS_M	lBS_Q	qBS_Q	cBS_Q	lBS_f	qBS_f
lr	-14.57	0.35	1.54	0.25	3.62	4.22	2.83	3.77	6.39	2.86	4.21
$\mathcal{X}^2$ 95%	—	3.84	3.84	3.84	5.99	7.81	7.81	9.49	11.07	5.99	7.81
Simulated 95%	-0.51	3.51	3.36	3.93	5.83	7.68	8.18	9.95	11.18	7.19	9.31
KS-test p=	—	0.27	<0.001	0.57	0.62	0.78	0.65	0.58	0.22	<0.001	<0.001

Table 8.4: FBC - Likelihood-Ratio Tests vs. LNT

9 Non-melanoma Skin Cancer: Basal Cell Carcinoma Type

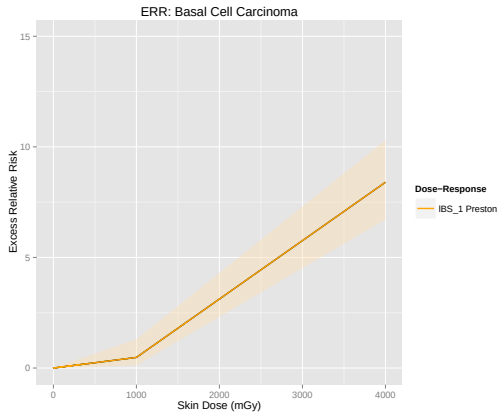

Figure 9.1: BCC - in Preston et al. (2007)

Dose	PYR	Cases
-0.005	1,598,934	77
-0.1	641,494	32
-0.2	149,943	11
-0.5	160,287	8
-1	113,556	8
-2	63,808	10
2+	36,704	20
Total	2,764,726	166

Table 9.1: PYR & Cases by Dose

In the Life Span Study cohort basal cell carcinoma type non-melanoma skin cancer (BCC) was diagnosed 166 times. Constraining the dose range to doses up to 2 Gy costs 36,704 person-years at risk (in 6154 strata groups) and leaves 146 cases. Table 9.1 shows that the rate of basal cell carcinoma cases per person-year at risk is high at doses above 1 Gy and exceptionally high above 2 Gy. Therefore, excess relative risk estimates are expected to be much lower when stratification groups with high doses are excluded.

In Preston et al. (2007) all non-melanoma skin cancers are analysed together (cf. Preston et al. 2007: pp. 30). Results of the separate analysis of BCC and SCC are summarized in two sentences. There is also no supplementary material that describes the complete results of these models. It is unclear how the models for the separate analysis were chosen and what model is used. Trying to replicate the excess relative risk of 0.57 at 1 Gy in a LNT model indicates that this is the result of a baseline model identical to the models for all non-melanoma skin cancers:

$$
\begin{aligned}
\lambda_0(X_{BL})(\text{c, s, a, e}) = \exp \quad & (\beta_1 \text{female} + \beta_2 \text{male} && (1)\\
& + \beta_3 \text{naga} && (2)\\
& + \beta_4 \text{mlage70} + \beta_5 \text{mlage70sq} + \beta_6 \text{mlage70qsp} && (3)\\
& + \beta_7 \text{flage70} + \beta_8 \text{flage70sq} + \beta_9 \text{flage70qsp} && (4)\\
& + \beta_{10} \text{me30} + \beta_{11} \text{me30sq} && (5)\\
& + \beta_{12} \text{fe30} + \beta_{13} \text{fe30sq} && (6)\\
& + \beta_{14} \text{nic·hiro} + \beta_{15} \text{nic·naga} && (7)\\
& + \beta_{16} \text{inahs}) && (8)
\end{aligned}
$$

The baseline risk depends on sex (1), the city (2), a second degree truncated power series of the logarithm of age attained with one knot at age 70 for men (3) and for women (4), a linear quadratic effect of the birth cohort again for men (5) and for women (6), not-in-city effects of Hiroshima and Nagasaki (7) and whether or not the individuals in the strata group were Adult Health Study participants or not (8).

In the models for all types of non-melanoma skin cancers the strength of the dose-response

depends on sex, age at exposure and age attained and is best modelled in the following way:

$$\epsilon(X_{EM}) = exp(\gamma_1 \cdot e30 + \gamma_2 \cdot lage70) \cdot (1 + \gamma_3 \cdot msex)$$

By effect coding the variable sex (msex=-1 for men and msex=1 for women) the estimated α parameter vector in the $err(\cdot)$ function is averaged over both sexes.
With this effect modification model the estimate in Preston et al. (2007) cannot be perfectly replicated.[13] Another possible effect modification model used by Preston et al. (2007) could be:

$$\epsilon(X_{EM}) = exp(\gamma_1 \cdot e30 + \gamma_2 \cdot lage70)$$

Excluding the sex effect has almost no effect on the likelihood and has therefore definitely no significant effect and the model without it is preferred by the AIC. For that reason - even though it is still not possible to exactly replicate the results in Preston et al. (2007) - the latter model for the analysis is used in this thesis.

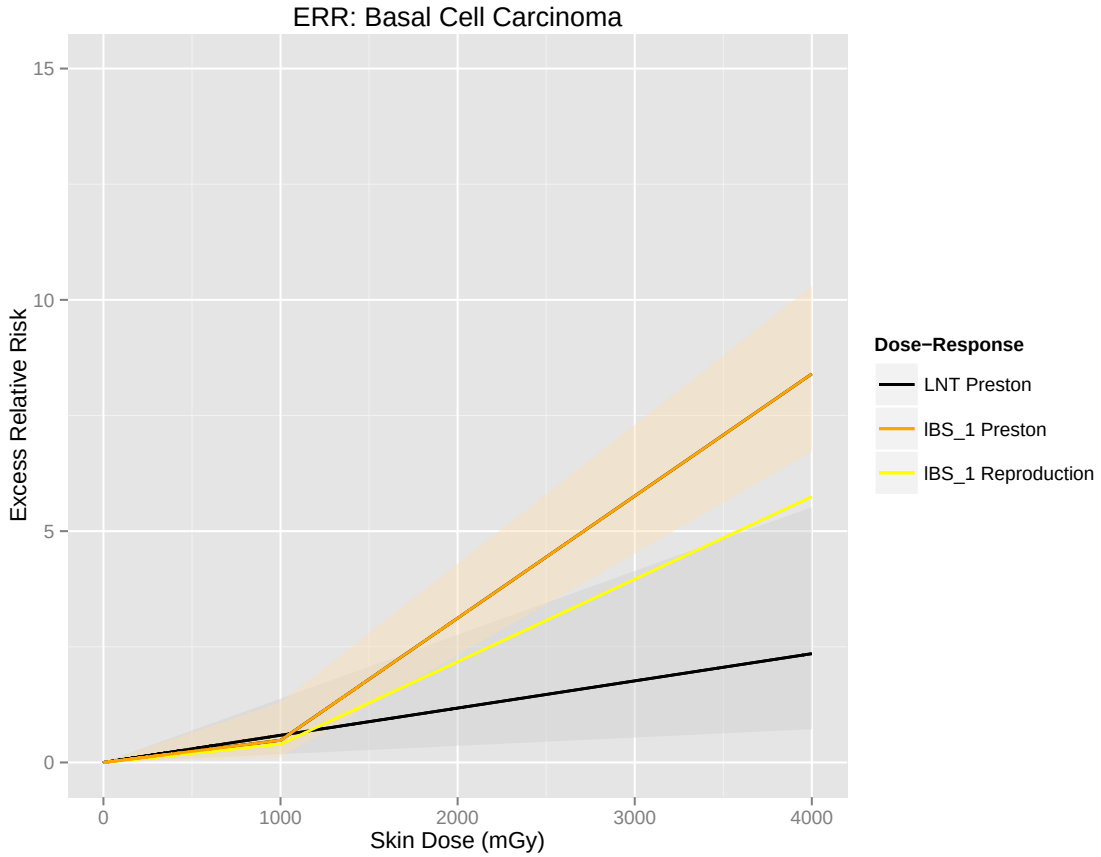

Figure 9.2: BCC - whole dose range

[13]ERR of 0.59 instead of 0.57. However, estimates in the supplementary materials for all non-melanoma skin cancers are also not completely consistent with the results presented in their paper. The best estimates for the spline coefficients are in the former 0.1813 below and 1.2653 (= 1.084 + 0.1813) above 2 Gy and in the latter 0.17 and 1.2 respectively (cf. Preston 2007: p.31 and in its supplement lss07ahs.pdf (in the electronic appendix).

In Preston et al. (2007) the preferred dose-response model is a first degree spline with one knot at 1 Gy. The spline model estimates of the excess relative risks (of 70 year olds who were 30 at the time of the bombings, averaged over sex) are considerably higher than with LNT. However, when trying to replicate these results, much lower estimates for the model are found (0.32 per Gy below and 1.18 above 2 Gy compared to 0.48 and 2.64 in Preston et al. (2007: p. 31), see figure 9.2). Even checking for local minima by minimizing with different starting values (100 different random starting value vectors), using different effect modifying variables and other optimizers (BFGS, Nelder-Mead (see section 4)) do not change the maximum likelihood estimate. It seems highly improbable that the model in this thesis is not the maximum likelihood spline model with one knot at 1 Gy.

9.1 Dose-Effect Estimation

After excluding all strata groups with an ionizing radiation dose above 2 Gy on the skin, the estimates are a lot lower. Excess relative risks in LNT - averaged over 30 years at exposure and 70 years attained - are estimated to be 0.04 per Gy (symmetric 95% CI [-0.14;0.22]). The estimated effect of the ionizing radiation dose completely disappears for cohort members age 30 at exposure and age 70 attained.

The differences in estimated excess relative risk estimates in all models seem to depend strongly on the estimated importance of age attained. When comparing the estimated effect of age attained for LNT on the whole range ($\gamma_{2\,LNT}$=-6.05), for the degree 1 spline with a knot at 1 Gy on the whole dose range ($\gamma_{2\,lBS\_1}$=-2.65), for LNT on the 0-2 Gy dose range ($\gamma_{2\,2Gy\,LNT}$=-13.27) and for QNT on 0-2 Gy ($\gamma_{2\,2Gy\,QNT}$=-4.43).
Plotted excess relative risk estimates (for cohort members with attained age 70 who where 30 at exposure) are lower when age attained is more important. Considering that LNT is - at least as of now - the best model, it is from a theoretical perspective very odd. Life span study cohort members age 20 at exposure and age 40 attained have an estimated excess relative BCC risk of 242 per Gy[14], what is of course incredibly high.

There seems to be no threshold dose below which excess risks are zero - the estimated threshold in the linear threshold model is at 3 mGy what is essentially equal to background radiation. The grid search can again be reconstructed in figure 9.5.

[14]Estimated parameters of the $ERR(\cdot)$ function of LNT are: $\delta_{1\,2Gy\,LNT}$=0.04024, $\gamma_{2\,2Gy\,LNT}$=-13.27 and $\gamma_{1\,2Gy\,LNT}$=-1.27.

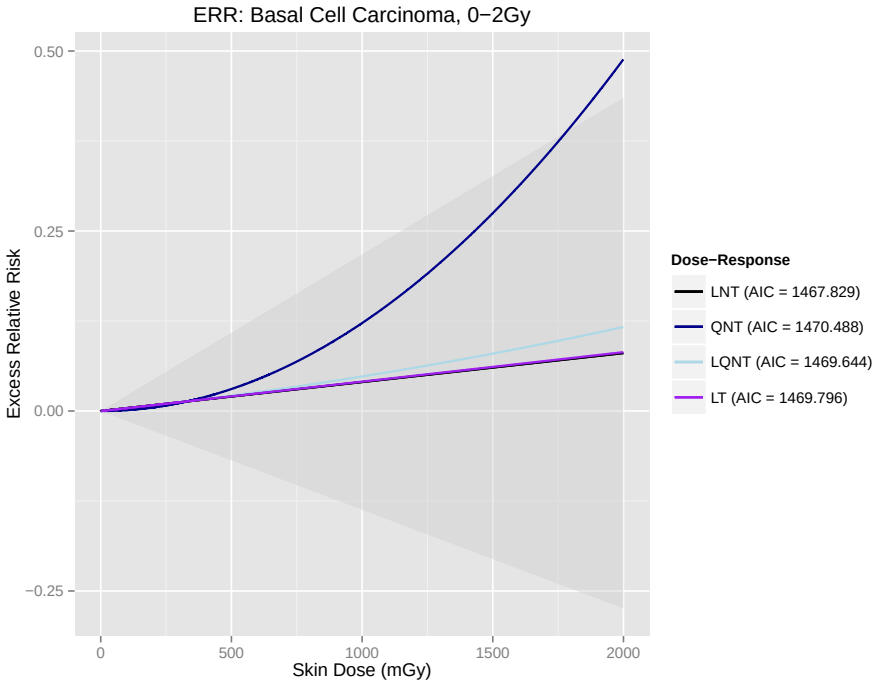

Figure 9.3: BCC - LNT, QNT, LQNT, LT

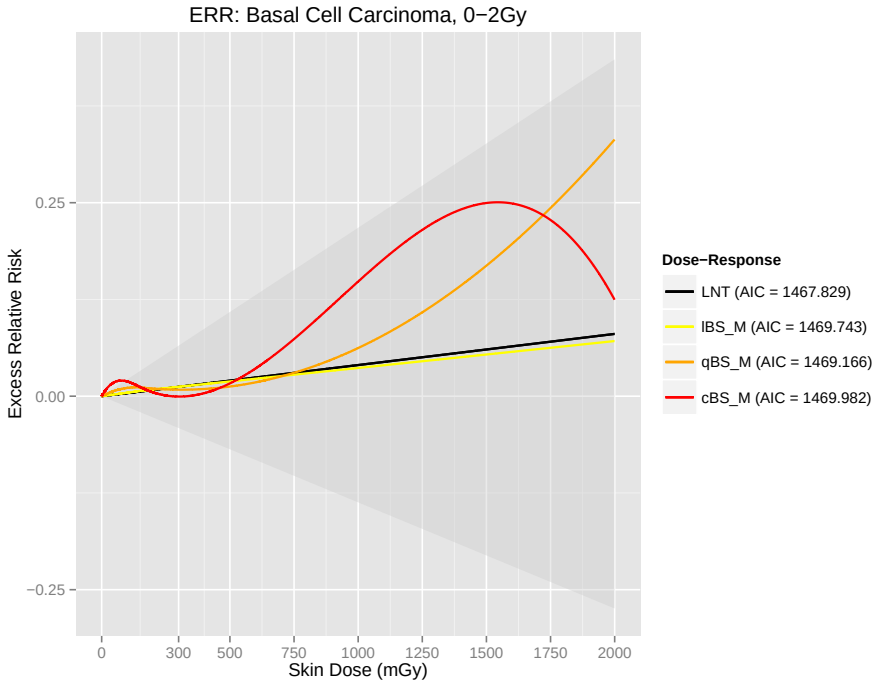

Figure 9.4: BCC - Splines - Knot at Median

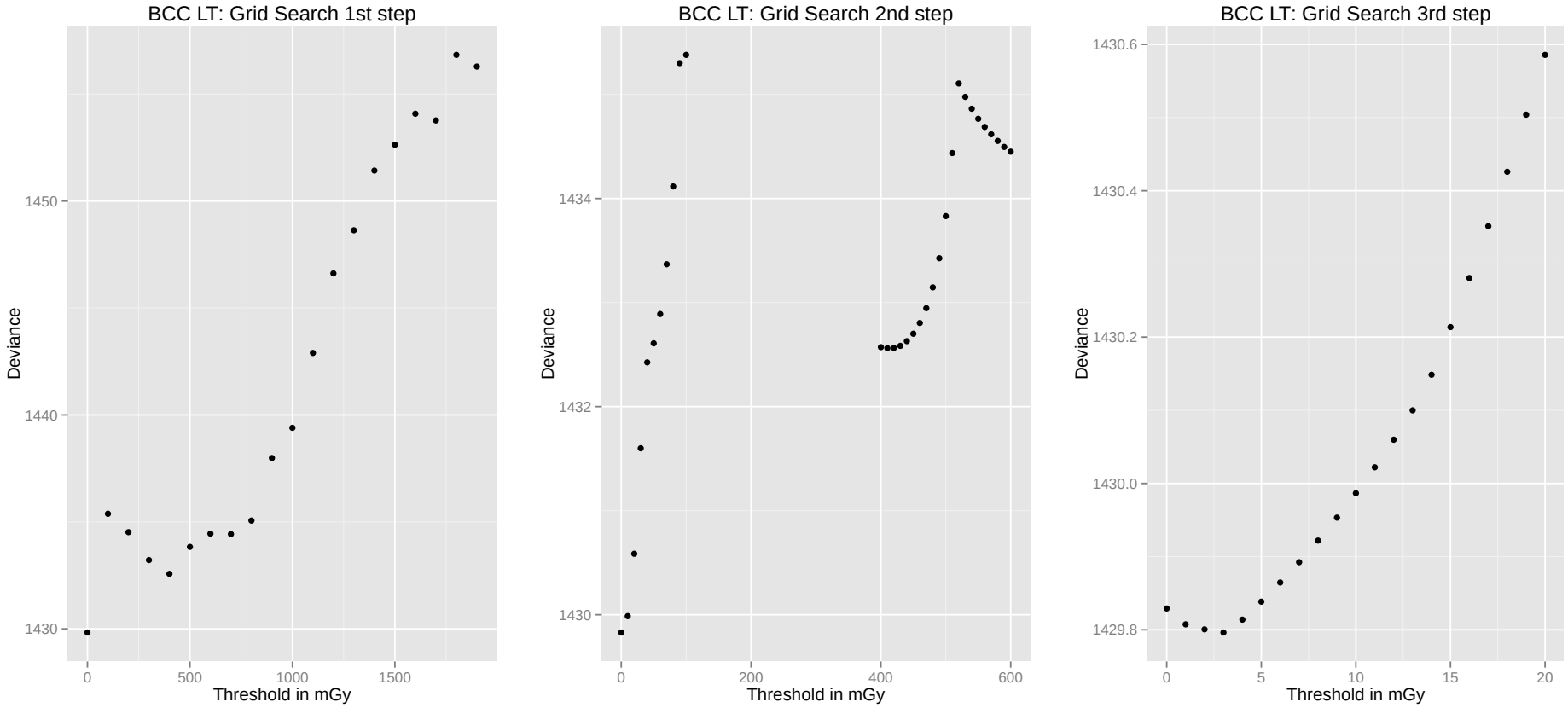

Figure 9.5: BCC - Grid Search Linear Threshold

Excess relative risk spline models (averaged for 70 years attained after being 30 years old at exposure) for doses above 500 mGy are estimated higher than LNT but lower than QNT. Consequently the estimated effect of age attained is lower than in LNT but higher than in QNT respectively. The degree 3 splines both with one knot at the median and with additional knots at the quartiles (the median is at 160 mGy, both quartiles at 18 mGy and 338 mGy respectively) and the linear spline with a free knot - grid searching (see figure 9.8) yields the maximum likelihood knot position at 1816 mGy - show small signs of effect saturation.
An explanation for the - in the degree 3 splines very substantial - reduction of estimated excess relative risks on high doses could be that data support is poor for doses only slightly below 2 Gy - a problem that can arise from stratifying the data (see section 2)[15]. However, as explained in section 2, this problem should be relativized by the fact that strata with high doses only contain data of a small number of individuals. The data support between 1.9 and 2 Gy is about 5626 person-years. Compared to the about 10,430 person-years between 1.75 and 2 Gy one can assume that data support between 1.9 and 2 Gy is more or less equally strong as on the whole range of the former interval.
In consideration of the estimates in Preston et al. (2007) and the high basal cell carcinoma skin cancer rates per person-year at risk at high doses (see table 9.1) the signs of effect saturation are probably an artefact of the choice to restrict the analysed dose-range to 0-2 Gy.
The degree 2 spline with knots at the quartiles has a non-positive definite hessian matrix - this means that (as already explained in section 7) at least one parameter in the model is poorly determined. Using 100 different starting values shows that the likelihood is equally high for different estimates for all parameters of the quadratic effect of the natural logarithm of age attained (flage70sq, flage70qsp, mlage70sq & mlage70qsp) on the baseline risk.

When testing the linear no-threshold model without effect modifiers (H_1) against the baseline model(H_0) the significance of the dose effect is at question. Using the $\mathcal{X}^2$ distribution the difference in likelihood is significant (on confidence level α=0.05) - using the $\mathcal{X}^2$ distribution changes this test decision. Additionally the simulated distribution does not differ significantly (p=0.59) from the respective $\mathcal{X}^2$ distribution - it follows that the decision which distribution should be used is disputable. As stated in section 7.2 the preferred approach - provided enough simulation runs - is to use the simulated 95% quantile.
Testing LNT with effect modifiers against the baseline models leaves no doubt about the significance of the dose effect. Also - as in the respective test for female breast cancer - the simulated likelihood ratios are not random values from the corresponding $\mathcal{X}^2$ distribution.

	LNT	LNT-noEM
lr	27.78	4.13
$\mathcal{X}^2$ 95%	7.81	3.84
Simulated 95%	10.83	4.29
KS-test p=	<0.001	0.59

Table 9.2: BCC - Likelihood-Ratio Tests vs. Baseline

[15]In figures 9.4, 9.6 and 9.7 the x-axis is labelled to highlight the dose intervals used for stratification.

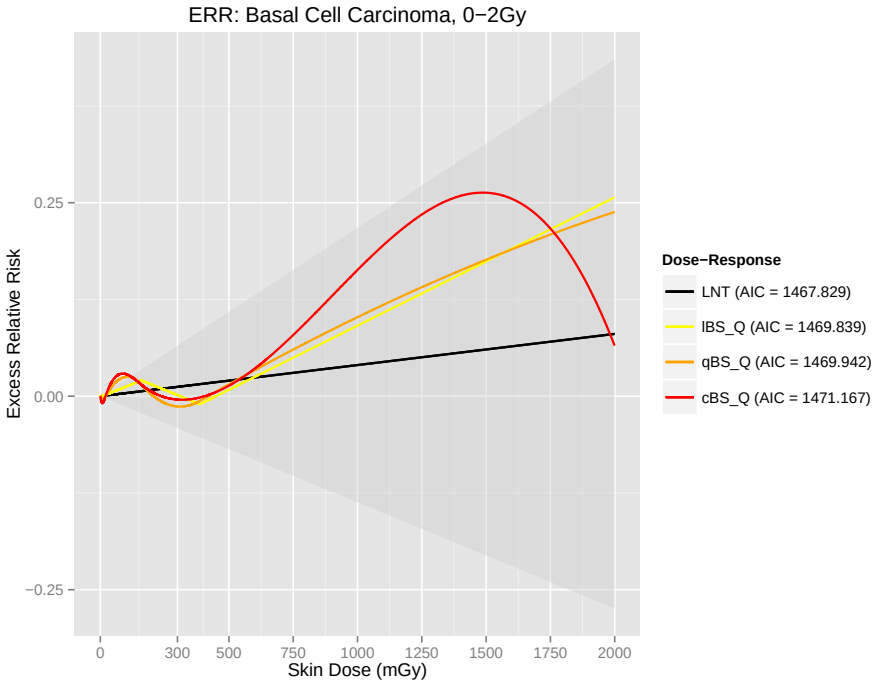

Figure 9.6: BCC - Splines - Knots at Quartiles

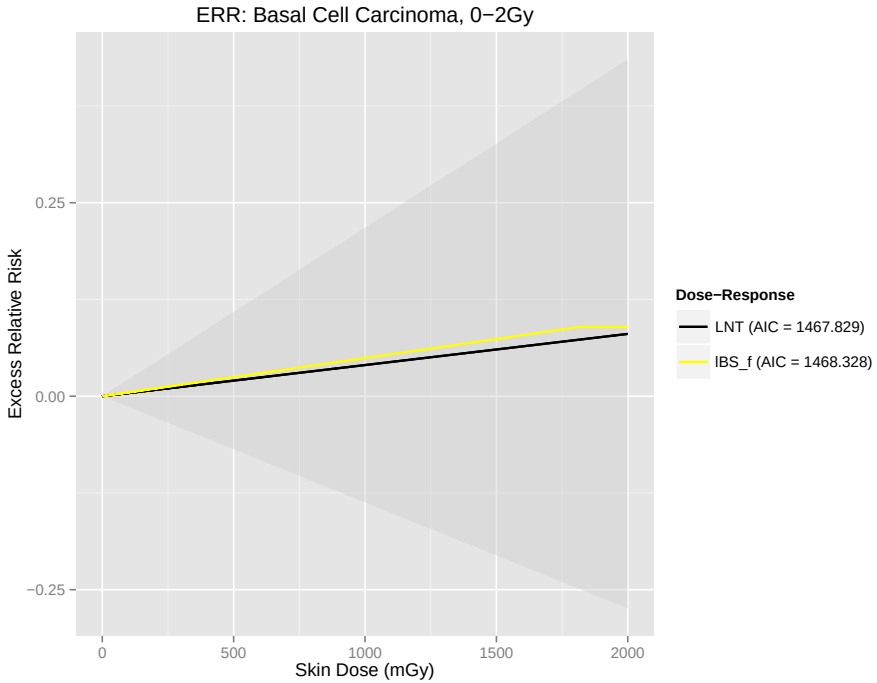

Figure 9.7: BCC - Splines - Free Knot

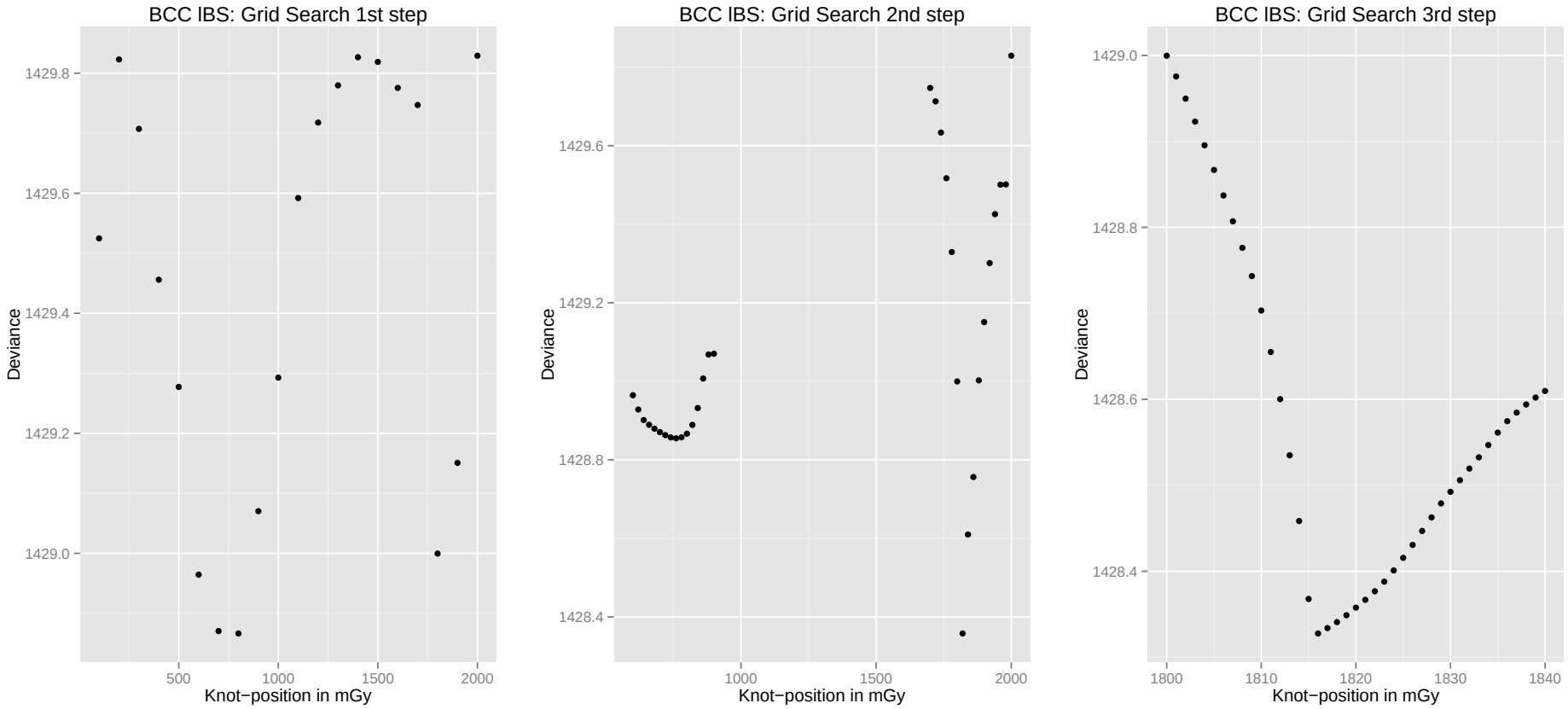

Figure 9.8: BCC - Grid Search B-Spline degree 1 (1 knot)

9.2 Testing for Non-Linearity

	Deviance	#Par.	AIC
Baseline	1457.613	16	1489.613
LNT noEM	1453.481	17	1487.481
LNT	1429.829	19	1467.829
QNT	1432.488	19	1470.488
LQNT	1429.644	20	1469.644
LT	1429.796	20	1469.796
lBS_M	1429.743	20	1469.743
qBS_M	1427.166	21	1469.166
cBS_M	1425.982	22	1469.982
lBS_Q	1425.839	22	1469.839
qBS_Q	1423.942	23	1469.942
cBS_Q	1423.167	24	1471.167
lBS_f	1428.328	21	1470.328

Table 9.3: BCC - Model Comparison

As stated above the preferred model by AIC is LNT. Testing for non-linearity (see table 9.4) shows that the small differences in the deviance are not significant - there is no significant non-linearity. Comparing the simulated distribution of lr with the $\mathcal{X}^2$ distribution again shows that the assumption of the latter is inappropriate.

	QNT	LQNT	LT	lBS_M	qBS_M	cBS_M	lBS_Q	qBS_Q	cBS_Q	lBS_f
lr	-2.66	0.19	0.03	0.09	2.66	3.85	3.99	5.89	6.66	1.50
$\mathcal{X}^2$ 95%	—	3.84	3.84	3.84	5.99	7.81	7.81	9.49	11.07	5.99
Simulated 95%	2.77	4.51	3.80	4.51	7.35	9.41	9.49	12.14	14.48	8.35
KS-test p=	—	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001

Table 9.4: BCC - Likelihood-Ratio Tests vs. LNT

10 Non-melanoma Skin Cancer: Squamous Cell Carcinoma Type

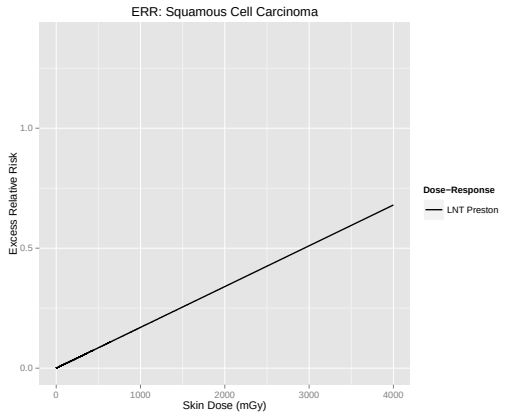

Figure 10.1: SCC - in Preston et al.(2007)

Dose	PYR	Cases
-0.005	1,598,934	71
-0.1	641,495	26
-0.2	149,943	12
-0.5	160,286	7
-1	113,556	6
-2	63,808	6
2+	36,704	3
Total	2,764,726	131

Table 10.1: PYR & Cases by Dose

Squamous cell carcinoma type non-melanoma skin cancer (SCC) was diagnosed 131 times. By restricting the data to up to 2 Gy ionizing radiation dose on the skin 36,704 person years at risk are lost and 128 cases remain. When compared to basal cell carcinoma only a small number of cases were diagnosed in strata groups with high radiation doses.
Trying to replicate the excess relative risk estimate (0.17 per Gy) in Preston et al. (2007) - which is not significantly better than the baseline model - yields that the preferred model is probably LNT without any effect modifiers with the same baseline risk model as in the analysis for all non-melanoma skin cancers and for basal cell carcinoma (see section 9).

10.1 Dose-Effect Estimation

After restricting the data the preferred non-spline excess relative risk model (see figure 10.2) by AIC is quadratic no-threshold (excess relative risk is 0.24 at 1 Gy). It has a slightly higher likelihood than LNT while using the same number of parameters. The pattern of QNT - that low doses are having a lower and high doses a higher excess relative risks than LNT - is consistent with every other model estimated in this analysis. Adding a linear term results in negative excess risk estimates for lower doses. LQNT, however, yields only a small increase in likelihood compared to QNT. The preferred threshold in LT is at 1268 mGy (grid search can be retraced in figure 10.4. After this very high threshold, estimated excess relative risks increase by 3.67 per Gy. The likelihood is higher for LT than LQNT but still not high enough to be preferred by AIC when compared to LNT and QNT.

As depicted in figure 10.3 the linear spline with a knot at the median (160 mGy) of the dose distribution has a dose-response curve plot similar to LNT and has almost no increase in likelihood at all. The respective degree 2 and 3 splines show a rather similar pattern to the linear threshold model: Very low - partially even negative - excess relative risks at doses below 1 Gy and a very steep increase for higher doses. Both spline models don't yield a big enough increase in likelihood to be preferred by AIC.
The spline models with additional knots at the quartiles (18 mGy and 338 mGy) of the dose

distribution show similar patterns (see figure 10.5). This steep increase in the dose-response can - for the same reasons as in section 9 - not be a consequence of poor data support at high doses.
The maximum likelihood estimate of the knot position in the degree 1 spline with a free knot is at 1257 mGy. Excess relative risks decrease before and increase strongly after this knot. The increase in likelihood in this model compared to LNT is, however, too low to be preferred by the AIC.
Surprisingly, when comparing all models with the AIC the 3rd degree spline with knots at the quartiles of the dose distribution is the preferred model even though it needs 5 parameters more than LNT and QNT. Only this spline is significantly better than the baseline model and all simulated distributions differ significantly from the respective $\mathcal{X}^2$ distributions.

	LNT	QNT	cBS_Q
lr	0.85	1.43	13.58
$\mathcal{X}^2$ 95%	3.84	3.84	12.59
Simulated 95%	4.29	4.63	12.79
KS-test p=	<0.001	0.033	<0.001

Table 10.2: SCC - Likelihood-Ratio Tests vs. Baseline

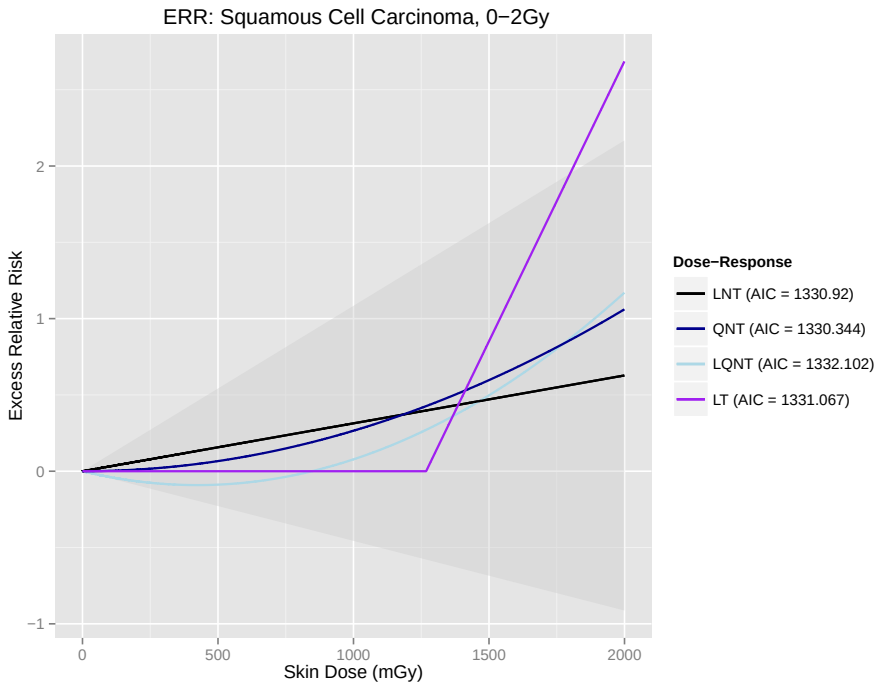

Figure 10.2: SCC - LNT, QNT, LQNT, LT

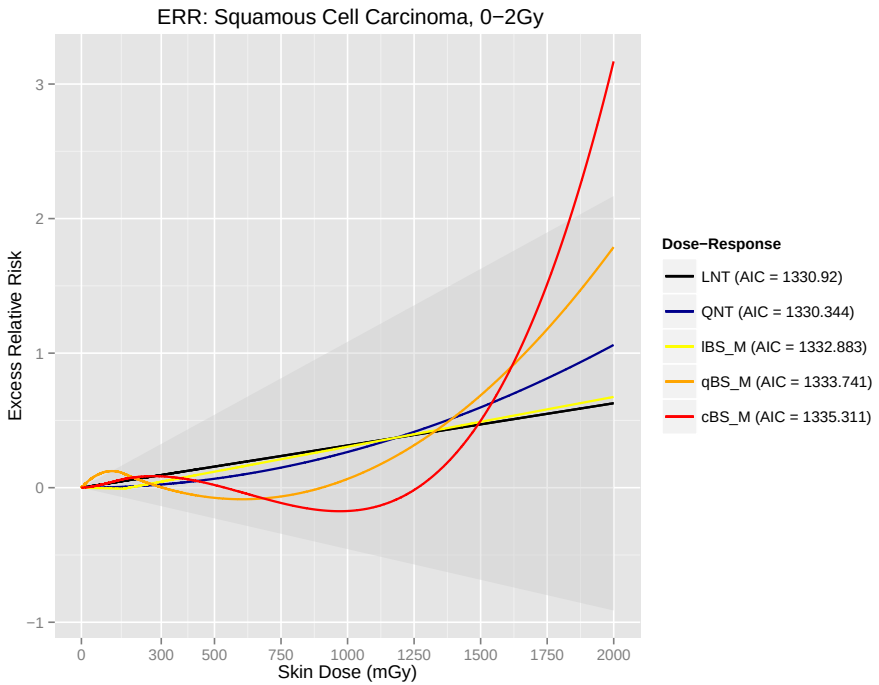

Figure 10.3: SCC - Splines - Knot at Median

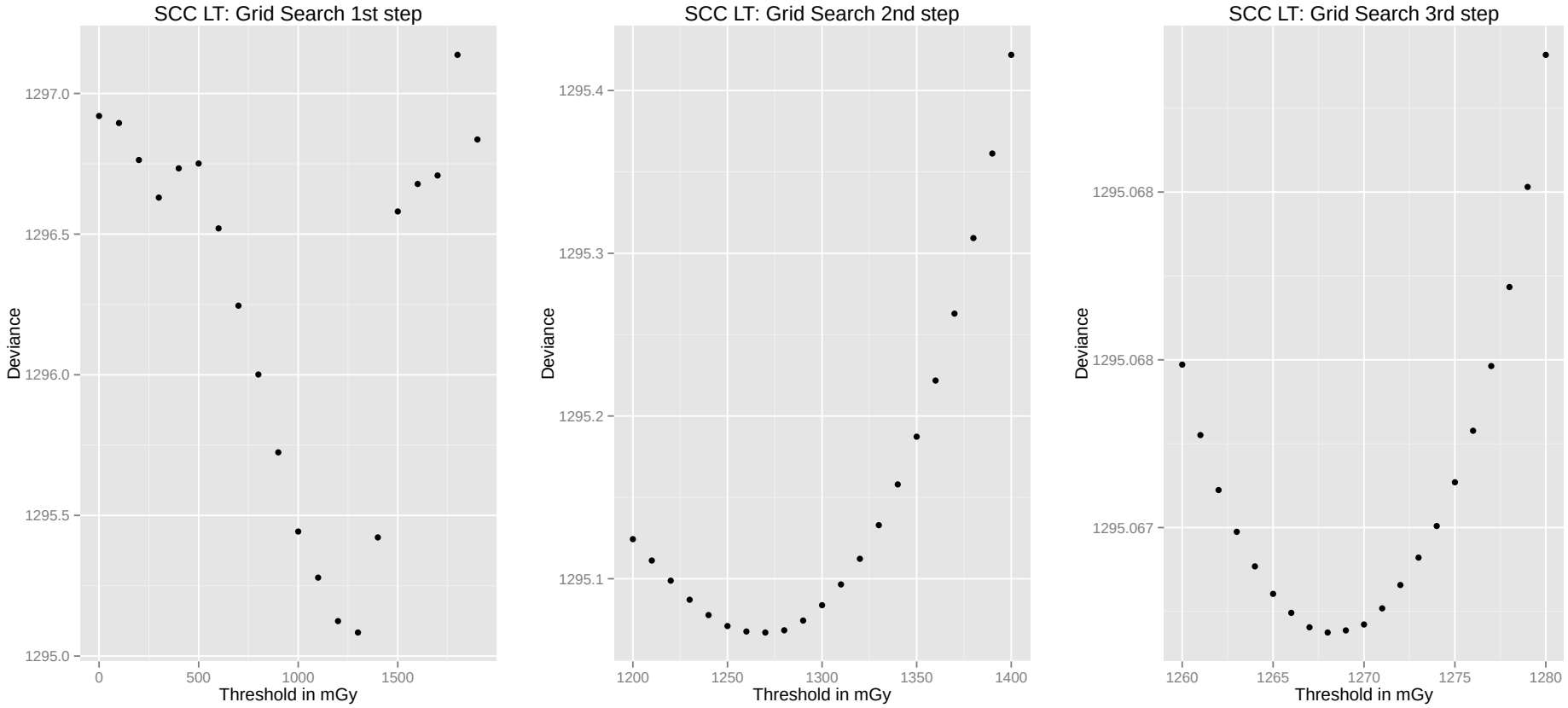

Figure 10.4: SCC - Grid Search Linear Threshold

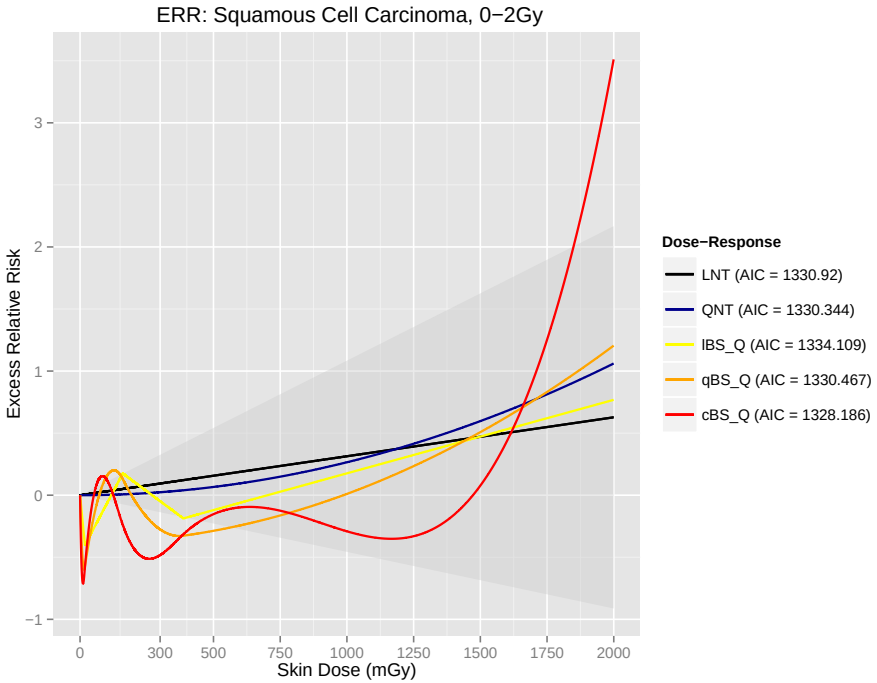

Figure 10.5: SCC - Splines - Knots at Quartiles

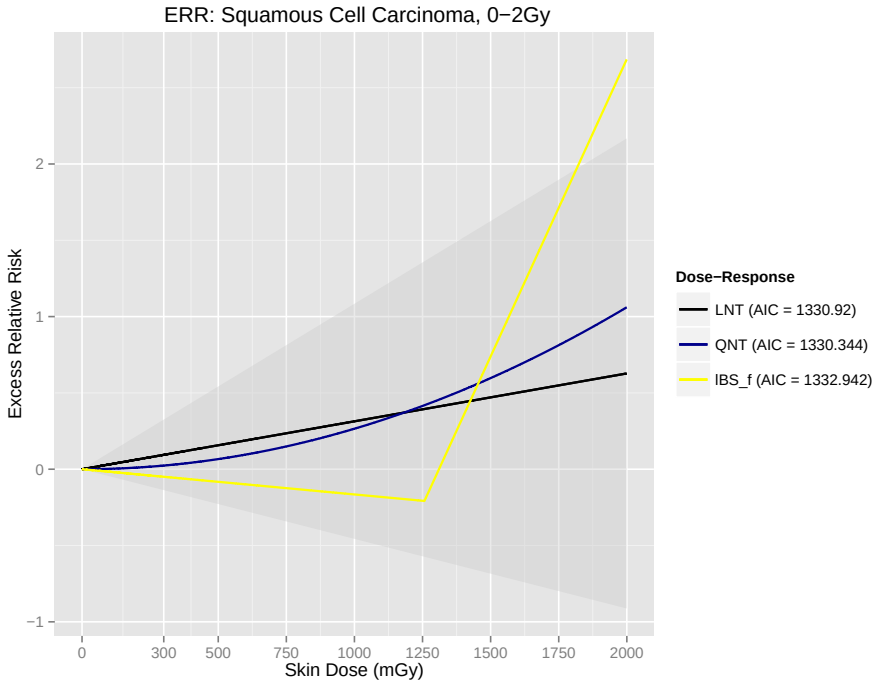

Figure 10.6: SCC - Splines - Free Knot

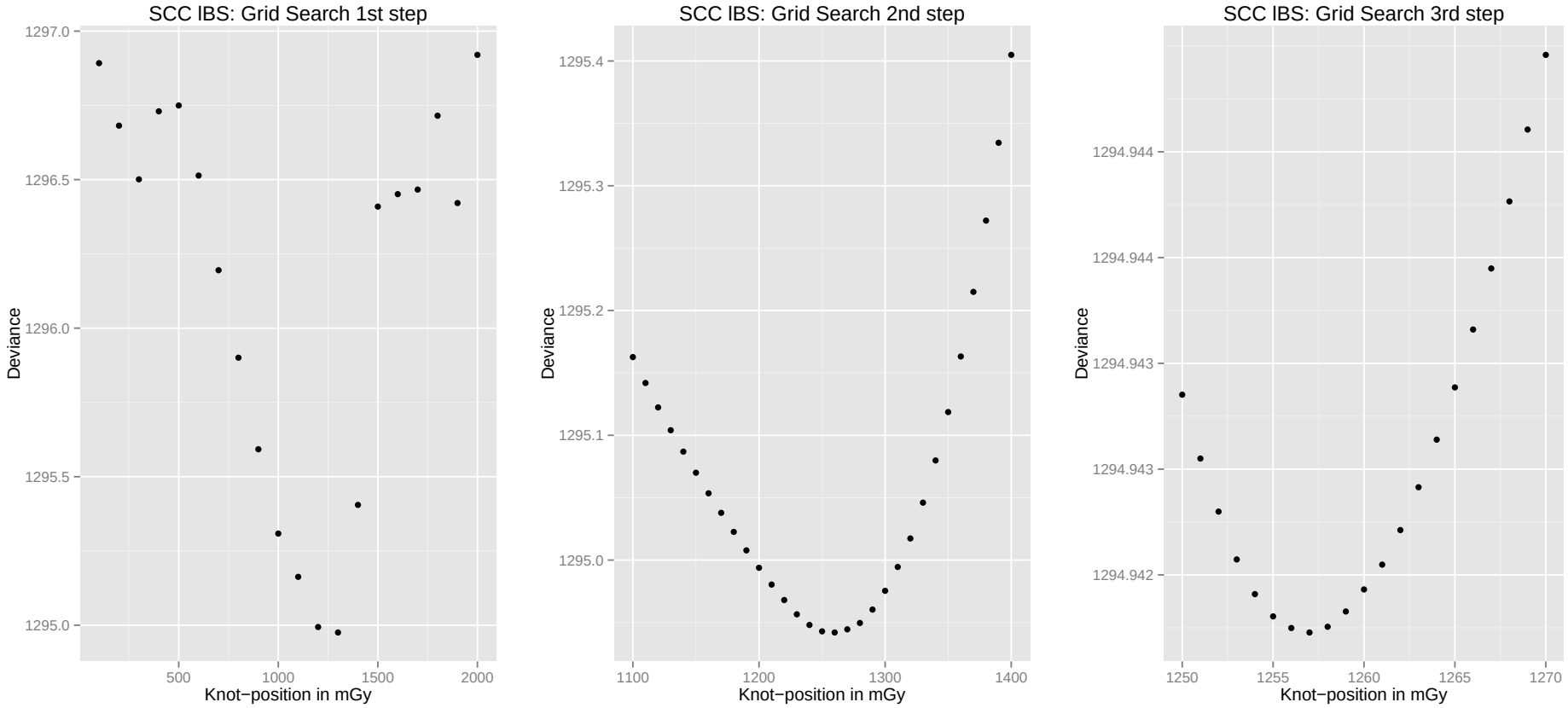

Figure 10.7: SCC - Grid Search B-Spline degree 1 (1 knot)

10.2 Testing for Non-Linearity

Figure 10.8: SCC - 95% CI: LNT cBS_Q

	Deviance	#Par	AIC
Baseline	1297.769	16	1329.769
LNT	1296.920	17	1330.920
QNT	1296.344	17	1330.344
LQNT	1296.102	18	1332.102
LT	1295.067	18	1331.067
lBS_M	1296.883	18	1332.883
qBS_M	1295.741	19	1333.741
cBS_M	1295.311	20	1335.311
lBS_Q	1294.109	20	1334.109
qBS_Q	1288.467	21	1330.467
cBS_Q	1284.186	22	1328.186
lBS_f	1294.942	19	1332.942

Table 10.3: SCC - Model Comparison

The preferred model by AIC is the degree 3 spline with knots at the quartiles of the dose distribution.

In table 10.4 one can see that the difference in log-likelihood of the degree 3 spline with one knot at the parameter (cBS_Q) and LNT is high enough to be significant on a 5% level. Simulating the distribution of lr under H_0 shows that the $\mathcal{X}^2$ assumption is comparable to the results in most cases. The Kolmogorov-Smirnov test shows for most nested models that the simulated lrs could be random values from the respective $\mathcal{X}^2$ distribution. Differences in the distributions are only significant for both degree 3 splines and all models with knots as parameters. Due to the conservativeness of the Kolmogorov-Smirnov test and the large differences in the corresponding 95% quantiles the asymptotic properties still are at question. Nevertheless, this could also be a sign that in some cases - maybe depending on model complexity - the asymptotic distribution is approached faster (in terms of required event counts) than in others.

	QNT	LQNT	LT	lBS_M	qBS_M	cBS_M	lBS_Q	qBS_Q	cBS_Q	lBS_f
lr	0.58	0.82	1.58	0.04	1.18	1.61	2.81	8.45	12.73	1.98
$\mathcal{X}^2$ 95%	—	3.84	3.84	3.84	5.99	7.81	7.81	9.49	11.07	5.99
Simulated 95%	1.45	4.38	4.41	3.93	6.30	8.71	8.10	10.01	12.02	7.61
KS-test p=	—	0.27	<0.001	0.923	0.286	0.026	0.240	0.241	0.016	<0.001

Table 10.4: SCC - Likelihood-Ratio Tests vs. LNT

11 Conclusion

For the most part restricting the dose to 2 Gy does not result in very different dose-response models. Dose-related excess relative acute myeloid leukemia risks are according to AIC best described by a quadratic no-threshold dose-response and female breast cancer risks by linear no-threshold dose-response. The preferred model for radiation related increases in chronic myeloid leukemia incidence is - after excluding models that overfit the data - also still linear no-threshold.
Linear no-threshold is the preferred dose-response for excess relative basal cell carcinoma type non-melanoma skin cancer risk. In Preston et al. (2013) it was, however, a degree 1 spline with a knot at the median - this difference in preferred models is to be expected because the spline was necessary to fit large increases in rates for strata that are defined by high levels of radiation exposure. These strata were excluded for the analysis in this thesis, hence the spline is not necessary any more. Restricting the dataset also led to a much lower dose response for cohort members with attained age 70 after being 30 years old at exposure. While in Preston et al. (2013) linear no-threshold (with no significant dose effect), the preferred non-spline dose-response model for squamous cell carcinoma type non-melanoma skin cancer risks after the restriction is quadratic no-threshold. A spline of degree 3 with knots at the quartiles estimating primarily negative excess relative risks for doses below 1500 mGy and a steep increase in risks after this threshold is, however, preferred by AIC.
For all of these cancer types a significant dose effect is observable as the preferred model is always significantly better than the baseline model.

In general the addition of flexible spline models for excess relative risk estimation can contribute a lot to detecting non-linearity in the dose-response. Splines with flexible knots and splines with a larger number of knots are, however, prone to overfit the data. This is the case for most spline estimates of chronic myeloid leukemia.

A problem for detecting non-linearity is that in Preston et al. (2007) and Hsu et al. (2013) the parameters of the baseline model as well as the effect modifiers get chosen with linear no-threshold as the dose-response model. Both the chosen baseline and effect modification model are used for all other dose-response models as well. Considering that the relevance of those parameters can differ for different dose-response models (see the effect modifying power of age attained in section 9) the goodness-of-fit comparison and significance testing is always biased towards linear no-threshold. The parameters are chosen to maximize the goodness-of-fit of linear no-threshold and all other models have to make do with these parameters.

Putting that aside, the results suggest that proper testing for the significance of parameters (and generating sensible confidence intervals for parameter estimates) is more complex - and definitely numerically more expensive - than practised in other papers that analyse the radiation related cancer risks of the atomic bomb survivor cohort. Stating that likelihood ratio test statistics (when testing nested models against each other) generally follow a $\mathcal{X}^2$ distribution under H_0 is not true for most of the cancer types analysed in this thesis. Whether or not the

asymptotic properties hold seems to depend on the number of events. When simulating the distribution the 95% quantiles tend to be higher than the quantile of the $\mathcal{X}^2$ distribution. It follows that dose responses - if the $\mathcal{X}^2$ distribution is assumed - are too often considered significantly non-linear.
In addition, using simulations to approximate the true distribution of lr under H_0 also enables testing the baseline model against models with effect modifiers and linear no-threshold against dose-response models with free knots. In both instances the restriction of certain parameters disable additional parameters - what results in problems for the likelihood-ratio test. In this case the $\mathcal{X}^2$ assumption is wrong even if it (for example in the case of female breast cancer risks) seems appropriate for all other tests.
When likelihood-ratio testing linear no-threshold against splines with free knots, the 95% quartile of the simulated distribution for the test-statistic under the null-hypothesis is consistently higher than the value suggested by the respective $\mathcal{X}^2$ distribution. It seems that making the knot-position in a spline available for estimation results in a much better fit to the data than to include a different non-knot parameter.
Simulation even can utilize the difference in deviance of non-nested models to establish a reasonable criterion for stating that one of them is unlikely to be the true data generating process. This is a useful tool when testing for non-linearity because the quadratic no-threshold model - in which linear no-threshold is not nested in - is very often the preferred model by AIC.

Considering these results, there are no signs of non-linearity of the dose-response related excess relative risks for chronic myeloid leukemia, female breast cancer and basal cell carcinoma type non-melanoma skin cancer. The dose-response for excess relative acute myeloid leukemia and squamous cell carcinoma type non-melanoma skin cancer risks is, however, significantly non-linear.

References

Armitage, P. and Colton, T. (eds) (1998). *Encyclopedia of biostatistics*, 2 edn, Wiley, Chichester.

Broyden, C. G. (1970). The convergence of a class of double-rank minimization algorithms 1. general considerations, *IMA Journal of Applied Mathematics* **6**(1): 76–90.

Deuflhard, P. (2011). *Newton methods for nonlinear problems: Affine invariance and adaptive algorithms*, Vol. 35 of *Springer Series in Computational Mathematics*, Springer, Berlin and New York.

Fahrmeir, L., Kneib, T., Lang, S. and Marx, B. (2013). *Regression: Models, methods and applications*, Springer-Verlag, Berlin Heidelberg.

Fletcher, R. (1970). A new approach to variable metric algorithms, *The computer journal* **13**(3): 317–322.

Folley, J. H., Borges, W. and Yamawaki, T. (1952). Incidence of leukemia in survivors of the atomic bomb in Hiroshima and Nagasaki, *The American journal of medicine* **13**(3): 311–321.

Fox, P. A., Hall, A. D. and Schryer, N. L. (1978). The PORT mathematical subroutine library, *ACM Transactions on Mathematical Software (TOMS)* **4**(2): 104–126.

Gilbert, P. and Varadhan, R. (2012). *numDeriv: Accurate Numerical Derivatives*. R package version 2012.9-1.
URL: *http://CRAN.R-project.org/package=numDeriv*
(Last called upon: June 25, 2015)

Goldfarb, D. (1970). A family of variable-metric methods derived by variational means, *Mathematics of computation* **24**(109): 23–26.

Hsu, W.-L., Preston, D. L., Soda, M., Sugiyama, H., Funamoto, S., Kodama, K., Kimura, A., Kamada, N., Dohy, H., Tomonaga, M., Iwanaga, M., Miyazaki, Y., Cullings, H. M., Suyama, A., Ozasa, K., Shore, R. E. and Mabuchi, K. (2013). The Incidence of Leukemia, Lymphoma and Multiple Myeloma among Atomic Bomb Survivors: 1950-2001, *Radiation Research* **179**: 361–382.

Karagas, M. R., McDonald, J. A., Greendberg, E. R., Stukel, T. A., Weiss, J. E., Baron, J. A. and Stevens, M. M. (24). Risk of basal cell and squamous cell skin cancers after ionizing radiation therapy, *Journal of the National Cancer Institute* **88**(1996): 1848–1853.

Krestinina, L., Preston, D. L., Davis, F. G., Epifanova, S., Ostroumova, E., Ron, E. and Akleyev, A. (2010). Leukemia incidence among people exposed to chronic radiation from the contaminated Techa River, 1953-2005, *Radiation and Environmental Biophysics* **49**(2): 195–201.

Matanoski, G. M., Seltser, R., Sartwell, P. E., Diamond, E. L. and Elliott, E. A. (1975). The current mortality rates of radiologists and other physician specialists: specific causes of death, *American Journal of Epidemiology* **101**(3): 199–210.

Muirhead, C. R., O'Hagan, J. A., Haylock, R., Phillipson, M. A., Willcock, T., Berridge, G. and Zhang, W. (2009). Mortality and cancer incidence following occupational radiation exposure: third analysis of the National Registry for Radiation Workers, *British Journal of Cancer* **100**: 206–212.

Nelder, J. A. and Mead, R. (1965). A simplex method for function minimization, *The computer journal* **7**(4): 308–313.

Novomestky, F. (2012). *matrixcalc: Collection of functions for matrix calculations*. R package version 1.0-3.
URL: *http://CRAN.R-project.org/package=matrixcalc*
(Last called upon: June 25, 2015)

Pierce, D. A. and Preston, D. L. (2000). Radiation-Related Cancer Risks at Low Doses among Atomic Bomb Survivors, *Radiation Research* **154**: 178–186.

Pierce, D. A., Stram, D. O. and Vaeth, M. (1990). Allowing for Random Errors in Radiation Dose Estimates for the Atomic Bomb Survivor Data, *Radiation Research* **123**(3): 275–284.

Pierce, D. A., Vaeth, M. and Cologne, J. B. (2008). Allowance for Random Dose Estimation Errors in Atomic Bomb Survivor Studies: A Revision, *Radiation Research* **170**(1): 118–126.

Preston, D. L., Lubin, J. H., Pierce, D. A. and McConney (1993). Epicure user's guide, *Seattle, WA: Hirosoft International Corporation* .

Preston, D. L., Mattson, A., Holmberg, E., Shore, R. E., Hildreth, N. G. and Boice Jr, J. D. (2002). Radiation effects on breast cancer risk: a pooled analysis of eight cohorts, *Radiation Research* **158**(2): 220–235.

Preston, D. L., Ron, E., Tukuoka, S., Funamoto, S., Nishi, N., Soda, M., Mabuchi, K. and Kodama, K. (2007). Solid Cancer Incidence in Atomic Bomb Survivors: 1958-1998, *Radiation Research* **168**: 1–64.

Preston, D. L., Shimizu, Y., Pierce, D. A., Suyama, A. and Mabuchi, K. (2003). Studies of mortality of atomic bomb survivors. Report 13: Solid cancer and noncancer disease mortality: 1950-1997, *Radiation Research* **160**(4): 381–407.

R Core Team (2014). *R: A Language and Environment for Statistical Computing (version 3.1.1)*, R Foundation for Statistical Computing, Vienna, Austria.
URL: *http://www.R-project.org/*
(Last called upon: June 25, 2015)

Rericha, V., Kulich, M., Řeřicha, R., Shore, D. L. and Sandier, D. P. (2006). Incidence of leukemia, lymphoma, and multiple myeloma in Czech uranium miners: a case-cohort study, *Environmental health perspectives* **114**(6): 818–822.

Robert, C. P. and Casella, G. (2004). *Monte Carlo statistical methods*, 2nd ed edn, Springer, New York.

Romanenko, A. Y., Finch, S. C., Hatch, M., Lubin, J. H., Bebeshko, V. G., Bazyka, D. A., Gudzenko, N., Dyagil, I. S. and Reiss, R. F. (2008). The Ukrainian-American Study of Leukemia and Related Disorders among Chornobyl Cleanup Workers from Ukraine: III. Radiation Risks, *Radiation Research* **170**(6): 711–720.

Ronckers, C. M., Erdmann, C. A. and Land, C. E. (2005). Radiation and breast cancer: a review of current evidence, *Breast Cancer Res* **7**(1): 21–32.

Schöllnberger, H., Kaiser, J.-C., Jacob, P. and Walsh, L. (2012). Dose-responses from multi-model inference for the non-cancer disease mortality of atomic bomb survivors, *Radiation and Environmental Biophysics* **51**(2): 165–178.

Sevcova, M., Sevc, J. and Thomas, J. (1978). Alpha irradiation of the skin and the possibility of late effects, *Health Physics* **35**(6): 803–806.

Shanno, D. F. (1970). Conditioning of quasi-Newton methods for function minimization, *Mathematics of computation* **24**(111): 647–656.

Smith, V. K. (1973). *Monte Carlo methods: their role for econometrics*, Lexington Books, Lexington, Mass.

Thomas, D. C. (2009). *Statistical Methods in Environmental Epidemiology*, Oxford University Press, Oxford and New York.

Turner, H. and Firth, D. (2012). *Generalized nonlinear models in R: An overview of the gnm package*. R package version 1.0-7.
URL: *http://CRAN.R-project.org/package=gnm*
(Last called upon: June 25, 2015)

Tutz, G. (2012). *Regression for categorical data*, Cambridge University Press, Cambridge and New York.

Wickham, H. (2009). *ggplot2: elegant graphics for data analysis*, Springer New York.
URL: *http://had.co.nz/ggplot2/book*
(Last called upon: June 25, 2015)

Yamada, M., Wong, F. L., Fujiwara, S., Akahoshi, M. and Suzuki, G. (2004). Noncancer disease incidence in atomic bomb survivors, 1958-1998, *Radiation Research* **161**(6): 622–632.

Young, R. W. and Kerr, George, D. (eds) (2005). *Reassessment of the Atomic Bomb Radiation Dosimetry for Hiroshima and Nagasaki–Dosimetry System 2002 (DS02)*, Radiation Effects Research Foundation, Hiroshima.

Appendix

Acute Myeloid Leukemia

Age Attained / B. M. Dose	-45		-60		-75		+75		Total	
	PYR	#AML	PYR	#AML	PYR	#AML	PYR	#AML	PYR	#AML
-0.005	772,003	10	587,457	16	487,008	29	192,627	22	2,039,095	77
-0.1	373,613	3	273,883	7	220,572	16	89,821	10	957,889	36
-0.2	74,734	2	58,055	2	48,649	2	20,497	3	201,935	9
-0.5	76,591	3	59,863	2	49,880	6	20,414	1	206,749	12
-1	43,925	3	34,686	2	29,171	6	10,073	0	117,855	11
-2	25,166	5	19,140	4	15,084	4	4,733	5	64,123	18
2+	11,452	2	7,653	5	5,333	1	1,322	5	25,761	13
Total	1,377,484	28	1,040,738	38	855,697	64	339,486	46	3,613,406	176

Table A.1: AML - Person-years and Cases by Bone Marrow Dose and Age Attained

Age at Exp. / B. M. Dose	-15		-30		-45		+45		Total	
	PYR	#AML	PYR	#AML	PYR	#AML	PYR	#AML	PYR	#AML
-0.005	728,595	12	613,842	19	455,409	38	241,249	8	2,039,095	77
-0.1	368,908	7	254,659	13	211,314	13	123,008	3	957,889	36
-0.2	74,159	3	52,152	5	47,638	0	27,986	1	201,935	9
-0.5	70,175	2	57,575	4	50,824	4	28,174	2	206,749	12
-1	35,518	2	41,682	3	26,326	3	14,328	3	117,855	11
-2	19,776	6	22,888	3	14,913	5	6,546	4	64,123	18
2+	9,668	4	9,106	3	4,400	4	2,587	2	25,761	13
Total	1,306,799	36	1,051,903	50	810,824	67	443,879	23	3,613,406	176

Table A.2: AML - Person-years and Cases by Bone Marrow Dose and Age at Exposure

	Estimate	Std. Error	z-value	Pr(> |z|)	95%CI	
Intercept	0.5794	0.16868	3.4348	5.93e-04***	0.2488	0.9100
female	-0.9229	0.17574	-5.2516	1.51e-07***	-1.2674	-0.5785
bcohort	0.1799	0.08528	2.1093	3.49e-02*	0.0127	0.3470
bcsq	-0.1878	0.05080	-3.6978	2.18e-04***	-0.2874	-0.0883
lage70	4.3326	0.64832	6.6829	2.34e-11***	3.0620	5.6033
lage70sq	1.5578	0.46620	3.3416	8.33e-04***	0.6441	2.4716
f_lage70	-1.2484	0.79690	-1.5666	1.17e-01	-2.8104	0.3135
f_lage70sq	-0.2532	0.53913	-0.4696	6.39e-01	-1.3098	0.8035
nic_naga	-0.1386	0.39387	-0.3518	7.25e-01	-0.9105	0.6334
nic_hiro	-0.2707	0.24369	-1.1108	2.67e-01	-0.7483	0.2069
dosegy	1.2328	0.51532	2.3923	1.67e-02*	0.2228	2.2428
EM:e30	0.1858	0.20230	0.9183	3.58e-01	-0.2107	0.5823
EM:e30sq	0.2295	0.09163	2.5043	1.23e-02*	0.0499	0.4090
EM:lage70	-1.2639	0.80884	-1.5627	1.18e-01	-2.8492	0.3214
Devianz:	1695.003					
AIC:	1723.003					

°p<0.1, *p<0.05, **p<0.01, ***p<0.001

Table A.3: AML - LNT (with EM), 0-2Gy

	Estimate	Std. Error	z-value	Pr(> |z|)	95%CI	
Intercept	0.6077	0.16244	3.7410	1.83e-04***	0.2893	0.9260
female	-0.9128	0.17587	-5.1900	2.10e-07***	-1.2574	-0.5681
bcohort	0.1725	0.08073	2.1372	3.26e-02*	0.0143	0.3308
bcsq	-0.1760	0.04806	-3.6622	2.50e-04***	-0.2702	-0.0818
lage70	4.3043	0.64483	6.6750	2.47e-11***	3.0404	5.5681
lage70sq	1.5644	0.45778	3.3417	6.32e-04***	0.6672	2.4617
f_lage70	-1.2389	0.79647	-1.5555	1.20e-01	-2.8000	0.3221
f_lage70sq	-0.2386	0.53755	-0.4439	6.57e-01	-1.2922	0.8150
nic_naga	-0.1934	0.39154	-0.4940	6.23e-01	-0.96081	0.5740
nic_hiro	-0.3260	0.24017	-1.3574	1.75e-01	-0.7967	0.1447
dosegysq^{16}	1.0689	0.43285	2.2048	2.75e-02*	0.1060	1.8027
EM:e30	0.2074	0.20193	1.0272	3.04e-01	-0.1883	0.6032
EM:e30sq	0.2447	0.09331	2.6222	8.74e-03*	0.0618	0.4276
EM:lage70	-1.3216	0.75829	-1.7428	8.14e-02°	-2.8078	0.1647
Devianz:	1686.190					
AIC:	1714.190					

°p<0.1, *p<0.05, **p<0.01, ***p<0.001

Table A.4: AML - QNT (with EM), 0-2Gy

Chronic Myeloid Leukemia

Age Attained / B. M. Dose	-45		-60		-75		+75		Total	
	PYR	#CML	PYR	#CML	PYR	#CML	PYR	#CML	PYR	#CML
-0.005	772,003	1	587,457	7	487,008	7	192,627	7	2,039,095	22
-0.1	373,613	4	273,883	4	220,572	5	89,821	4	957,889	17
-0.2	74,734	0	58,055	2	48,649	0	20,497	0	201,935	2
-0.5	76,591	3	59,863	4	49,880	3	20,414	1	206,749	11
-1	43,925	2	34,686	2	29,171	2	10,073	0	117,855	6
-2	25,166	5	19,140	0	15,084	3	4,733	1	64,123	9
2+	11,452	7	7,653	0	5,333	1	1,322	0	25,761	8
Total	1,377,484	22	1,040,738	19	855,697	21	339,486	13	3,613,406	75

Table B.1: CML - Person-years and Cases by Bone Marrow Dose and Age Attained

Age at Exp. / B. M. Dose	-15		-30		-45		+45		Total	
	PYR	#CML	PYR	#CML	PYR	#CML	PYR	#CML	PYR	#CML
-0.005	728,595	8	613,842	1	455,409	12	241,249	1	2,039,095	22
-0.1	368,908	3	254,659	2	211,314	6	123,008	6	957,889	17
-0.2	74,159	1	52,152	1	47,638	0	27,986	0	201,935	2
-0.5	70,175	1	57,575	1	50,824	4	28,174	5	206,749	11
-1	35,518	3	41,682	0	26,326	2	14,328	1	117,855	6
-2	19,776	0	22,888	3	14,913	4	6,546	2	64,123	9
2+	9,668	4	9,106	2	4,400	2	2,587	0	25,761	8
Total	1,306,799	20	1,051,903	10	810,824	30	443,879	15	3,613,406	75

Table B.2: CML - Person-years and Cases by Bone Marrow Dose and Age at Exposure

	Estimate	Std. Error	z-value	Pr(> |z|)	95%CI	
Intercept	-1.4968	0.29462	-5.0804	3.77e-04***	-2.0742	-0.9193
female	-0.1121	0.30317	-0.3697	7.12e-01	-0.7063	0.4821
lage70	1.2340	0.69489	1.7758	7.58e-02°	-0.1280	2.5959
f_lage70	1.9271	0.79915	2.4114	1.59e-02*	0.3608	3.4934
nic_naga	-0.8469	1.01935	-0.8308	4.06e-01	-2.8448	1.1510
nic_hiro	-0.1891	0.42495	-0.4450	6.56e-01	-1.0220	0.6438
dosegy	5.3842	2.54661	2.1143	3.45e-02*	0.3930	10.3755
EM:naga	-1.9520	1.16296	-1.6785	9.33e-02°	-4.2313	0.3274
EM:ltsx25	-1.6508	0.38726	-4.2627	2.02e-05***	-2.4098	-0.8918
EM:lage55	-0.6868	0.87064	-0.7888	4.30e-01	-2.3932	1.0197
Devianz:	804.117					
AIC:	824.117					

°p<0.1, *p<0.05, **p<0.01, ***p<0.001

Table B.3: CML - LNT (with EM), 0-2Gy

Female Breast Cancer

Age Attained / Breast Dose	-45		-60		-75		+75		Total	
	PYR	#FBC	PYR	#FBC	PYR	#FBC	PYR	#FBC	PYR	#FBC
-0.005	294,291	63	307,535	205	274,221	188	118,798	90	994,844	546
-0.1	129,047	34	130,241	87	114,930	67	51,911	29	426,129	217
-0.2	27,234	6	28,576	20	25,975	23	11,934	9	93,719	58
-0.5	32,217	16	33,809	41	30,772	28	14,399	19	111,197	104
-1	11,982	10	15,568	17	16,894	16	6,441	8	50,885	51
-2	11,391	11	10,990	27	15,084	3	2,829	7	34,115	60
2+	4,678	8	4,413	22	3,548	6	925	1	13,565	37
Total	510,840	148	531,133	419	475,244	343	207,238	163	1,724,454	1073

Table C.1: FBC - Person-years and Cases by Breast Dose and Age Attained

Age at Exp. / Breast Dose	-15		-30		-45		+45		Total	
	PYR	#FBC	PYR	#FBC	PYR	#FBC	PYR	#FBC	PYR	#FBC
-0.005	294,825	150	358,474	214	246,394	137	95,152	45	994,844	546
-0.1	137,809	66	135,720	84	108,370	54	44,231	13	426,129	217
-0.2	26,952	17	32,305	23	24,362	12	10,101	6	93,719	58
-0.5	34,142	26	36,561	50	28,460	23	12,034	5	111,197	104
-1	9,526	11	18,626	23	17,550	14	5,183	3	117,855	6
-2	10,019	16	14,658	29	7,195	11	2,243	4	34,115	60
2+	4,654	15	5,168	15	2,678	6	1,064	1	13,565	37
Total	517,926	301	601,512	438	435,007	257	170,009	77	1,724,454	1073

Table C.2: FBC - Person-years and Cases by Breast Dose and Age at Exposure

	Estimate	Std. Error	z-value	Pr(> |z|)	95%CI	
Intercept	2.0883	0.06832	30.5682	3.24e-205***	1.9544	2.2223
naga	-0.1647	0.07976	-2.0655	3.88e-02*	-0.3211	-0.0084
lage50preqsp	-10.7834	2.99245	-3.6035	3.14e-04***	-16.6484	-4.9183
lage70	2.6306	0.64826	4.0579	4.95e-05***	1.3560	3.9011
lage70sq	2.2941	1.56842	1.4627	1.44e-01	-0.7800	5.3681
lage70qsp	-2.0418	5.69423	-0.3586	7.20e-01	-13.2023	9.1187
e30	-0.3888	0.03638	-10.6860	1.18e-26***	-0.4601	-0.3175
e30sq	-0.0131	0.01494	-0.8743	3.82e-01	-0.0424	0.0162
nic_naga	0.0613	0.16255	0.3769	7.06e-01	-0.2573	0.3799
nic_hiro	-0.0472	0.08913	-0.5298	5.96e-01	-0.2219	0.1275
brea02w10	0.9652	0.28328	3.4072	6.56e-04***	0.4100	1.5204
EM:e30	-0.0133	0.14710	-0.0908	9.28e-01	-0.3017	0.2750
EM:lage70	-1.9686	0.86539	-2.2748	2.29e-02*	-3.6647	-0.2725
Devianz:	4782.969					
AIC:	4808.969					

°p<0.1, *p<0.05, **p<0.01, ***p<0.001

Table C.3: FBC - LNT (with EM), 0-2Gy

Non-melanoma Skin Cancer: Basal Cell Carcinoma Type

Age Attained / Skin Dose	-45		-60		-75		+75		Total	
	PYR	#BCC	PYR	#BCC	PYR	#BCC	PYR	#BCC	PYR	#BCC
-0.005	522,396	0	492,561	4	415,584	33	168,394	40	1,598,934	77
-0.1	224,727	2	193,683	2	156,062	12	67,022	16	641,494	32
-0.2	46,966	0	46,472	2	40,429	5	16,075	4	149,943	11
-0.5	48,372	0	48,005	1	44,272	1	19,637	6	160,287	8
-1	36,904	1	34,943	1	29,981	3	11,728	3	113,556	8
-2	19,305	1	20,174	4	17,975	4	6,353	1	63,808	10
2+	12,678	0	11,928	9	9,437	8	2,662	3	36,704	20
Total	911,349	4	847,766	23	713,740	66	291,871	73	2,764,726	166

Table D.1: BCC - Person-years and Cases by Skin Dose and Age Attained

Age at Exp. / Skin Dose	-15		-30		-45		+45		Total	
	PYR	#BCC	PYR	#BCC	PYR	#BCC	PYR	#BCC	PYR	#BCC
-0.005	578,864	4	495,928	17	364,368	37	159,773	19	1,598,934	77
-0.1	263,687	3	171,954	8	140,025	13	65,828	8	641,494	32
-0.2	53,812	1	43,939	3	35,581	5	16,610	2	149,943	11
-0.5	52,438	0	46,183	1	42,261	4	19,405	3	160,287	8
-1	39,640	2	36,633	1	25,142	2	12,141	3	113,556	8
-2	17,382	4	24,099	5	16,758	1	5,568	0	63,808	10
2+	11,957	6	13,824	11	7,602	1	3,320	2	36,704	20
Total	1,017,781	20	832,560	46	631,738	63	282,647	37	2,764,726	166

Table D.2: BCC - Person-years and Cases by Skin Dose and Age at Exposure

	Estimate	Std. Error	z-value	Pr(> |z|)	95%CI	
female	0.1026	0.18098	0.5669	0.5707	-0.2521	0.4573
male	0.5336	0.24211	2.2037	0.0275*	0.0590	1.0081
naga	0.2262	0.19708	1.1478	0.2510	-0.1601	0.6125
inahs	0.2983	0.19342	1.5422	0.1230	-0.0808	0.6774
nic_naga	-0.3494	0.44305	-0.7886	0.4304	-1.2177	0.5190
nic_hiro	-0.2820	0.24745	-1.1397	0.2544	-0.7670	0.2030
f_lage70	7.7313	2.24037	3.4509	0.0006***	3.3403	12.1224
m_lage70	6.7957	2.68240	2.5334	0.0113*	1.5383	12.0531
f_lage70sq	-10.5193	10.14987	-1.0364	0.3000	-30.4127	9.3741
m_lage70sq	-10.4942	10.16441	-1.0324	0.3019	-30.4160	9.4277
f_lage70sp	-1.4453	18.53053	-0.0780	0.9378	-37.7645	34.8738
m_lage70sp	0.0603	22.55760	0.0027	0.9979	-44.1517	44.2724
f_e30	-0.0555	0.13921	-0.3986	0.6902	-0.3283	0.2174
m_e30	-0.1279	0.15722	-0.8136	0.4158	-0.4361	0.1802
f_e30sq	-0.0514	0.05758	-0.8921	0.3723	-0.1642	0.0615
m_e30sq	-0.1317	0.07821	-1.6834	0.0923°	-0.2849	0.0216
skia02w10	0.0402	0.09051	0.4446	0.6566	-0.1372	0.2176
EM:e30	-1.2748	0.52654	-2.4212	0.0155*	-2.3068	-0.2429
EM:lage70	-13.2719	7.96046	-1.6672	0.0955°	-28.8741	2.3303
Devianz:	1429.829					
AIC:	1467.829					

°p<0.1, *p<0.05, **p<0.01, ***p<0.001

Table D.3: BCC - LNT (with EM), 0-2Gy

Non-melanoma Skin Cancer: Squamous Cell Carcinoma Type

Skin Dose \ Age Attained	-45		-60		-75		+75		Total	
	PYR	#SCC	PYR	#SCC	PYR	#SCC	PYR	#SCC	PYR	#SCC
-0.005	522,396	5	492,561	8	415,584	13	168,394	45	1,598,934	71
-0.1	224,727	2	193,683	4	156,062	8	67,022	12	641,494	26
-0.2	46,966	2	46,472	0	40,429	5	16,075	5	149,943	12
-0.5	48,372	0	48,005	0	44,272	2	19,637	5	160,287	7
-1	36,904	0	34,943	1	29,981	1	11,728	4	113,556	6
-2	19,305	1	20,174	0	17,975	1	6,353	4	63,808	6
2+	12,678	1	11,928	1	9,437	1	2,662	0	36,704	3
Total	911,349	11	847,766	14	713,740	31	291,871	75	2,764,726	131

Table E.1: SCC: Person-years and Cases by Skin Dose and Age Attained

Skin Dose \ Age at Exp.	-15		-30		-45		+45		Total	
	PYR	#SCC	PYR	#SCC	PYR	#SCC	PYR	#SCC	PYR	#SCC
-0.005	578,864	1	495,928	15	364,368	35	159,773	20	1,598,934	71
-0.1	263,687	3	171,954	4	140,025	12	65,828	7	641,494	26
-0.2	53,812	1	43,939	3	35,581	3	16,610	5	149,943	12
-0.5	52,438	1	46,183	0	42,261	5	19,405	1	160,287	7
-1	39,640	0	36,633	0	25,142	4	12,141	2	113,556	6
-2	17,382	1	24,099	1	16,758	3	5,568	1	63,808	6
2+	11,957	1	13,824	1	7,602	1	3,320	0	36,704	3
Total	1,017,781	8	832,560	24	631,738	63	282,647	36	2,764,726	131

Table E.2: SCC - Person-years and Cases by Skin Dose and Age at Exposure

	Estimate	Std. Error	z-value	Pr(> |z|)	95%CI	
female	-0.4765	0.20754	-2.2962	2.17e-02*	-0.8833	-0.0698
male	-0.0717	0.25499	-0.2814	7.78e-01	-0.5715	0.4280
naga	-0.2213	0.23258	-0.9517	3.41e-01	-0.6772	0.2345
inahs	0.2555	0.22270	1.1475	2.51e-01	-0.1809	0.6920
nic_naga	-0.9991	0.73841	-1.3531	1.76e-01	-2.4464	0.4481
nic_hiro	-0.5112	0.26575	-1.9236	5.44e-02°	-1.0321	0.0097
f_lage70	6.8123	1.42374	4.7848	1.71e-06***	4.0218	9.6028
m_lage70	2.4845	1.47585	1.6835	9.23e-02°	-0.4081	5.3771
f_lage70sq	5.9415	1.12231	5.2940	1.20e-07***	3.7418	8.1412
m_lage70sq	1.0481	1.46692	0.7145	4.75e-01	-1.8270	3.9233
f_lage70sp	0.3331	7.77885	0.0428	9.66e-01	-14.9132	15.5793
m_lage70sp	27.8083	8.30393	3.3488	8.12e-04***	11.5329	44.0837
f_e30	0.5394	0.22072	2.4437	1.45e-02*	0.1068	0.9720
m_e30	0.0678	0.14999	0.4518	6.51e-01	-0.2262	0.3617
f_e30sq	-0.2295	0.08125	-2.8244	4.74e-03**	-0.3887	-0.0702
m_e30sq	-0.1251	0.06436	-1.9436	5.19e-02°	-0.2512	0.0011
skia02w10	0.3138	0.39293	0.7986	4.25e-01	-0.4563	1.0839
Devianz:	1296.920					
AIC:	1330.920					

°p<0.1, *p<0.05, **p<0.01, ***p<0.001

Table E.3: SCC - LNT, 0-2Gy

	Estimate	Std. Error	z-value	Pr(> |z|)	95%CI	
female	-0.1521	0.23986	-0.6343	5.26e-01	-0.6223	0.3180
male	0.2087	0.27861	0.7491	4.54e-01	-0.3374	0.7548
naga	-0.1994	0.23754	-0.8393	4.01e-01	-0.6649	0.2662
inahs	0.1429	0.22848	0.6255	5.32e-01	-0.3049	0.5907
nic_naga	-1.2939	0.74203	-1.7437	8.12e-02°	-2.7483	0.1604
nic_hiro	-0.7993	0.28702	-2.7849	5.35e-03**	-1.3619	-0.2368
f_lage70	6.8090	1.42307	4.7848	1.71e-06***	4.0199	9.5982
m_lage70	2.4650	1.47842	1.6673	9.54e-02°	-0.4326	5.3626
f_lage70sq	5.9439	1.12195	5.2978	1.17e-07***	3.7449	8.1429
m_lage70sq	1.0422	1.47276	0.7076	4.79e-01	-1.8444	3.9287
f_lage70sp	0.2321	7.77297	0.0299	9.76e-01	-15.0024	15.4670
m_lage70sp	28.2346	8.36703	3.3745	7.39e-04***	11.8355	44.6337
f_e30	0.5441	0.22100	2.4619	1.38e-02*	0.1109	0.9772
m_e30	0.0658	0.15021	0.4377	6.62e-01	-0.2287	0.3602
f_e30sq	-0.2295	0.08121	-2.8262	4.71e-03**	-0.3887	-0.0704
m_e30sq	-0.1230	0.06420	-1.9155	5.54e-02°	-0.2488	0.0029
α_1	-0.9211	0.17823	-5.1680	2.37e-07***	-1.2704	-0.5718
α_2	0.6950	0.50047	1.3888	1.65e-01	-0.2859	1.6759
α_3	-0.8569	0.45875	-1.8680	6.18e-02°	-1.7561	0.0422
α_4	1.2469	1.32480	0.9412	3.47e-01	-1.3497	3.8434
α_5	-2.7521	2.37408	-1.1592	2.46e-01	-7.4052	1.9010
α_6	3.5160	2.97117	1.1834	2.37e-01	-2.3074	9.3394
Devianz:	1284.186					
AIC:	1328.186					

°p<0.1, *p<0.05, **p<0.01, ***p<0.001

Table E.4: SCC - Spline, degree 3, knots at quartiles, 0-2Gy

Electronic Appendix

See enclosed CD.

Acknowledgements

This report makes use of data obtained from the Radiation Effects Research Foundation (RERF) in Hiroshima and Nagasaki, Japan. RERF is a private, non-profit foundation funded by the Japanese Ministry of Health, Labour and Welfare (MHLW) and the U.S. Department of Energy (DOE), the latter through the National Academy of Sciences. The data include information obtained from the Hiroshima City, Hiroshima Prefecture, Nagasaki City, and Nagasaki Prefecture Tumor Registries and the Hiroshima and Nagasaki Tissue Registries. The conclusions in this report are those of the authors and do not necessarily reflect the scientific judgment of RERF or its funding agencies.