Journal of Universal Computer Science, vol. 12, no. 1 (2006), 115-125
submitted: 6/5/05, accepted: 13/12/05, appeared: 28/1/06 © J.UCS

A Web-Based Decision Support System for Chronic Diseases

Chi-Chang Chang
(Yuan Ze University, Taiwan, ROC
threec@mail.yust.edu.tw)

Chuen-Sheng Cheng
(Yuan Ze University, Taiwan, ROC
ieccheng@saturn.yzu.edu.tw)

Yeu-Shiang Huang
(National Cheng Kung University, Taiwan, ROC
yshuang@mail.ncku.edu.tw)

Abstract: Individuals vary in survival chances due to differences in genetics, environmental exposures, and gene-environment interactions. These chances, as well as the contribution of each factor to mortality, change as individuals get older. In general, human physiological systems are constructed by collecting more than one part to perform either single or multiple functions. In addition, the successive times between failures are not necessarily identically distributed. More generally, they can become smaller (an indication of deterioration). However, if any critical deterioration is detected, then the decision of when to take the intervention, given the costs of diagnosis and therapeutics, is of fundamental importance. At the time of the decision, the degree of future physiological system deterioration, which is likely to be uncertain, is of primary interest for the decision maker. This paper develops a possible Web-based decision support system by considering the sensitivity analysis as well as the optimal prior and posterior decisions for aging chronic diseases. The proposed design of Bayesian decision support systems facilitates the effective use of the computing capability of computers and provides a systematic way to integrate the expert's opinions and the sampling information which will furnish decision makers with valuable support for quality decision making.

Keywords: Chronic Diseases, nonhomogeneous Poisson process (NHPP), Bayesian Decision Theory, Decision Support System

Categories: F.4.2, G.1.10, G.3, H.1.0, H.2.1, H.4.2

1 Introduction

Demographic shifts in the population will lead to a further increase in the proportion of elderly and consequently of people with chronic diseases. For example, almost 75 percent of the elderly (age 65 and over) have at least one chronic and about 50 percent have at least two chronic diseases [Calkins et al., (99)]. In addition, aging is a strong socially appealing issue with many implications for human as well as providers of healthcare. In general, human physiological systems are constructed by collecting more than one part to perform either single or multiple functions. However, the

successive times between failures are not necessarily identically distributed. More generally, they can become smaller (an indication of deterioration). If any critical deterioration is detected, then the decision of when to take the intervention, given the costs of treatments and failures, is of fundamental importance. At the time of the decision, the degree of future physiological deterioration, which is likely to be uncertain, is of primary interest for the decision maker (e.g., determining the prevalence of disease, doing a population survey, or measuring the level of a toxin). Naturally, gathering additional data will not always be economical. It is of special interest to determine analytically or numerically the conditions under which it will be worthwhile to collect additional information. Therefore, we propose a Bayesian decision process to provide a significantly improved methodology for dealing with the decision problems of human physiological systems which can determine the conditions for taking the different actions, and thereby help the decision maker maximize expected profit (or minimize expected loss). This paper reviews several plausible approaches and develops a Web-based decision support system (DSS) for carrying out the complex computations involved in the decision process to determine the conditions for taking different actions, and thereby provides guidelines for decision-making and furnishes decision makers with valuable support for making reliable and robust decisions.

2 Bayesian Decision Model for Aging Physiological System

In many studies in the reliability, engineering, public health, medical, actuarial, and economic settings, the event of primary interest is recurrent, so that for a given unit the event could be observed more than once during the study [Zenia et al., 99]. For example, the breakdown of electromechanical systems (e.g., motor vehicles, computers, nuclear power plants) are recurrent events in the reliability and engineering setting. Example of recurrent events in public health and medical settings include outbreak of diseases (e.g., encephalitis), repeated hospitalization of end-stage renal disease patients, recurrence of tumors, and angina pectoris in patients with chronic coronary disease [Byar, (80)], [Gail et al., 80], [Klein et al., 89], [Lawless, 87], [Thall and Lachin, 88], and [Wei et al., 89]. Recurrent events pervade many studies in variety of fields, and hence it is of paramount importance to have appropriate models and methods of statistical analyses for such data [Chang and Cheng, 06].

2.1 The Failure Intensity Functions

In many longitudinal studies of chronic diseases, the diseases are modeled by either renewal processes or nonhomogeneous Poisson processes. A renewal process is based on the assumption that the system will be “as good as new” after repair (e.g., Hip replacement) [Cox, 62], [Lawless, 82], [Nelson, 82], [Aalen and Husebye, 91], [Cook et al., 99], and [Abouammoh and Qamber, 03]. A nonhomogeneous Poisson process is based on the assumption that the system will be “as good as old” after repair. For example, [Cook et al., 96] developed a robust test for kidney transplant based on recurrent event responses. [Wang et al., 00] used a nonhomogeneous Poisson process to analyze the times between failures to determine the optimum first metastases.

[Dewanji and Moolgavkar, 00] present a position process formulation for chronic respiratory disease in relation to air pollution indices. [Aggarwal et al., 05] also investigated the mean time between failures for dorsal cochlear nucleus neurons. In practice, the former assumption may be reasonable for only one part, while the latter assumption seems more plausible for complex human physiological systems consisting of many organs, each of which has its own failure mode.

In this study, we assume an aging physiological system behaves according to a nonhomogeneous Poisson process and that the physiological system failure process is time-dependent. The intensity function of the failure process is usually assumed to be of the form $\lambda(x)=\lambda_0 h(\beta;x)$, where λ_0 is the scale factor, β is the aging rate, x is the elapsed time, and $h(.)$ can be any function that reflects the aging process. Several different failure intensity functions have been proposed. The three commonly used models are linear, exponential, and power law failure models.

- The linear model is of the form $\lambda(x) = \lambda_0(1+\beta x)$. This model assumes that the failure rate increases (or decreases) linearly over time, with an initial failure rate of λ_0 and a linear aging rate β. Where β, measured in units of 1/time, may be negative, corresponding to survivability growth; however, it must be large enough that $\lambda(x)$ does not become negative in the time period of interest.
- The exponential model is of the form $\lambda(x) = \lambda_0 \, exp(\beta x)$. This model assumes that the failure rate increases (or decreases) exponentially over time, with an initial failure rate of λ_0 and an exponential deterioration rate β. If β is negative, the survivability growth is indicated, while β is positive, then the physiological system is deteriorating over time.
- The power-law model is of the form $\lambda(x) = \lambda_0 \beta x^{\beta-1}$, where β is effectively unit less and must be positive to ensure that $\lambda(x)$ does not become negative. For $\beta<1$ the failure intensity is decreasing (corresponding to survivability growth), and for $\beta>1$ the failure intensity is increasing (corresponding to deterioration).

One of the other models studied by [Cox and Lewis, 66], called the log-linear model, $\lambda(x) = \exp(\alpha + \beta x)$, is essentially the same as the exponential model if we let $\lambda_0 = \exp(\alpha)$. Some other models have also been proposed to model deterioration, but in more complicated ways. Unlike the three models given above, these models often have more than two parameters.

2.2 Bayesian Decision Model

Suppose that human being has a lifetime (i.e., time horizon) T and the decision has to be made at time t. The crucial two-action decision is whether at time t, the failure rate will be too high (in which case some risk reduction action needs to be taken), or whether it will still be within an acceptable range (in which case we can according to the status quo). Another option is to gather additional information before the final decision is made. The basic elements of the Bayesian decision process are as follows [Huang and Chang, 04]:

- Parameter space Θ:{(λ_0,β)| λ_0>0}, where λ_0 is the initial factor and β is the aging rate. Both parameters are uncertain and can be estimated through experts' opinions.
- Action space A:{a_1,a_2}, where a_1 is the status quo, and a_2 is the risk reduction action. (We eventually expand this to consider a third possible action, the collection of additional information).
- Loss function L: a real function defined on $\Theta \times A$. If we decide to keep continuing the status quo, then the loss we face is $L(\theta,a_1)$; if we decide to take the risk reduction action, then the loss we face is $L(\theta,a_2)$.
- Sample space S: the additional information available to be collected. For example, the successive failure times till the *nth* failure can be denoted as the likelihood function of the form

$$f_{X_1,X_2,\cdots,X_n}(x_1,x_2,\cdots,x_n)=[\prod_{i=1}^{n}\lambda(x_i)]\exp(-\Lambda(x_n)) \quad (1)$$

  where $\Lambda(x)=\int_0^x\lambda(u)du$ is the mean number of failures by time x in the NHPP. The cost of collecting this additional information should also be reflected in the decision process. The following terminology will be used throughout this paper:
- C_A: the cost of a failure if it occurs.
- C_R: the cost of the proposed risk reduction action.
- C_I: the cost of collecting additional information.
- ρ: the reduction in failure rate that would result from the proposed risk reduction action ($0<\rho<1$).
- M: the expected number of failures during the time period $[t,T]$ under the status quo.

The decision variable we are dealing with is then the expected number of failures during the time period $[t,T]$, i.e.,

$$M\equiv M(T,t,\lambda_0,\beta)=\int_t^T\lambda(s)ds \quad (2)$$

Note that the expected number of failures M is itself a random variable, since it is a function of the two uncertain parameters λ_0 and β, and this is the case where Bayesian analysis can be effectively performed. Suppose that the risk reduction action will reduce the failure intensity by a fraction ρ, where $0 < \rho < 1$, then the expected number of failures in $[t\,;\,T]$, if the risk reduction action is taken is given by

$$\int_t^T\lambda(s)(1-\rho)ds = (1-\rho)\ M. \quad (3)$$

On the basis of the assumptions given above, we therefore have a two-action problem with a linear loss function, where the loss for taking action a_1 (i.e., continuing with the status quo) is C_AM and the loss for taking action a_2 (i.e., undertaking the risk reduction action) is $C_A(1-\rho)M+C_R$. The expected loss for the status quo is simply $C_A\boldsymbol{E}\{M\}$, and the expected loss for the risk reduction action is $C_A(1-\rho)\boldsymbol{E}\{M\}+C_R$. Since the prior and posterior density functions for M are functions of λ_0 and β, some prior and posterior mean values of M can be derived by the bivariate transformation technique. For example, the prior mean value of M for the power law failure model, with the assumptions that λ_0 and β are independent and their prior distributions are Gamma(α ; γ) and Uniform (a,b), respectively. However, closed forms for the prior and posterior means of M are not always

available which are the typical cases for the Bayesian analysis. Nevertheless, Bayesian prior and posterior analyses can still be performed by computing the prior and posterior mean values of M using the numerical integration technique and comparing them with the cut of value $M_C=C_R/(C_A\rho)$ (i.e., the cutoff value of $E\{M\}$ for taking the risk reduction action). If the relevant mean is smaller than M_C, then we should keep the status quo; if not, then we should perform the risk reduction action.

2.3 The CGD Case

We have used the chronic granulomatous disease (CGD) case study to illustrate the use of the models developed in the previous sections. Real failure data from a trial of immunotherapy for the CGD are studied [International Chronic Granulomatous Disease Cooperative Study Group, 91]. CGD is an inherited disease caused by defects in superoxide-generating nicotinamideadenine dinucleotide phosphate (NADPH) oxidase of phagocytes. Impairment of oxygen-dependent intracellular killing mechanisms results in severe bacterial or fungal infections with catalase-producing Staphylococcus aureus, Burkholderia cepacia, or Aspergillus Species. Antimicrobial prophylaxis is efficient in reducing the incidence of severe bacterial infections [International Chronic Granulomatous Disease Cooperative Study Group, 91]. The birthdate of the studied subject was 1-May-1973, and the observation period was from 24-August-1988, to 1-September-1989. The failure dates for the subject during the observation period were: {26-Sep-88, 26-Oct-88, 25-Nov-88, 25-Dec-88, 24-Jan-89, 23-Feb-89, 25-Mar-89, 24-Apr-89, 5-May-89, 24-May-89, 23-Jun-89, 23-Jul-89, 15-Aug-89, 22-Aug-89}.

Table 1 summarizes the results of the results for *EVPI* (Expected Value of Perfect Information). Prior and posterior analyses can be performed by comparing the prior and posterior mean values of λ_0 with the cutoff value τ_C. If the relevant mean is smaller than τ_C, then we should keep the status quo; if not, then we should perform the risk reduction action (e.g., gene therapy). As can be seen from that table 2, the effect of incorporating the observed data is to switch the optimal decision from the status quo to the risk reduction action for linear failure models. This is because the increasing cost of the risk reduction action is justified for the large failure rates assumed in the prior distributions. However, in the exponential and power law failure model, the observed data generally support the adoption of the risk reduction action.

	Linear Model	Power Law Model	Exponential Model
Prior $E\{M\}$	6.1428	11.0327	26.4416
Range of $E\{M\}$ for Collecting Information	5.327~10.152	2.538~27.968	3.622~55.860
EVPI	19046.56	32231.42	28224.74
Prior $E\{\lambda_0\}$	0.1	0.1	0.1
Posterior $E'\{\lambda_0\}$	0.1694	0.1422	0.1793
Prior $E\{\beta\}$	0.5	1.57	0.17
Posterior $E'\{\beta\}$	0.7849	1.8488	0.1809
Cutoff Value of $E\{M\}$ for Risk Reduction	7.2	7.2	7.2
Prior Decision	Status Quo	Risk Reduction	Risk Reduction
Posterior $E'\{M\}$	9.9132	10.2495	23.4724
Posterior Decision	Risk Reduction	Risk Reduction	Risk Reduction

Table 1: Summary Table for the Case of CGD

3 The Decision Support System

Decision support systems provide practitioners with patient-specific assessments or recommendations to help them make decisions about treatment. According to [Watson and Sprague, (93)], a general notion about a DSS is that it is an interactive computerized system consisting of three major components: a dialog subsystem, a database subsystem, and a model base subsystem. With the knowledge and other capabilities embodied in these components, a DSS is intended to help a decision maker interactively solve managerial decision problems. However, the three component architecture is capable of managing data; fitting data into models; and providing methods to reach decisions. By manipulating models and data, the decision maker is able to examine various scenarios and their consequences. Therefore, these three components, as a whole, contribute to the quality of decisions that are taken by a decision maker.

Nevertheless, a system which provides passive decision support barely achieves its design objectives as the user's experiences, knowledge, and expertise change [Chuang and Yadav, 98]. The Bayesian decision process mentioned in the previous section is capable of not only dealing with the uncertainties but also taking into account prior knowledge. In such a case, the decision is based only on the uncertainties quantified by decision makers. Also, the DSS has to notify the decision maker whether collecting additional information is desirable or not. Therefore, the DSS has the ability to allow decision makers change each uncertain entity (e.g. aging rate), and therefore a range of such uncertain entities within which the optimal decision remains unchanged should be derive. It is of important interest for decision makers to learn what the resulting decision will be once they change some parameters.

According to the discussion above, the inputs to the DSS would be the uncertainties mentioned previously and the failure data (if available), and the outputs would be the optimal prior decision, the optimal posterior decision, and the results of the sensitivity analysis and what-if analysis [Erto, 82]. In order to perform the required complicated numerical integration for the decision process, the specifications of hardware and software should be closely considered. The hardware should have the ability to correctly and quickly respond to decision makers before they get impatient and the software should be easily and reliably programmed and maintained. Figure 1 and 2 show the follow chart and the framework of the DSS, respectively. The DSS has input, output, and process three major parts. The detailed information descriptions of each part are as follows:

- ***Prior information part***: the prior information has eleven elements and there are: (1) *Lifetime*: the expected lifetime of the physiological system. (2) *Initial date*: the birthdate. (3) *Decision time*: the actual time for decision makers to make the decision of whether maintaining the status quo or undertaking a risk reduction action. (4) *Cost of failure*: the cost or loss once the failure actually occurs. (5) *Cost of risk reduction action*: the cost for undertaking the risk reduction action. (6) *Risk reduction factor*: the fraction of the original function of the physiological system that the risk reduction action can retrieve. (7) *Cost of collecting information*: the cost of collecting the failure data. (8) *E{Scale Factor}*: the expected value of the initial factor. (9) *SD{Scale Factor}*: the standard deviation of the initial factor. (10) *E{Deterioration Rate}*: the expected value of the deterioration rate. (11) *SD{Deterioration Rate}*: the standard deviation of the deterioration rate.
- ***Sampling information part***: The sampling information is for inputting the observed failure data.
- ***Decision part***: The decision part provides the optimal decisions that are suggested by the DSS. There are five output elements that can be valuable to decision makers for making the final decision. There are: (1) *Expected value of sampling information (EVSI)*: the EVSI can be treated as an indicator for determining whether to collect the failure data. In particular, if the EVSI were greater than the cost of collecting information applied in the prior information area, then collecting the failure data would be desirable; otherwise, collecting the failure data is not desirable. (2) *Prior E{# of Failure}*: the expected number of failures under the status quo which is estimated by using the prior information only. This value shows the performance of the system if no risk reduction action is considered. (3) *Prior decision*: the suggested decision is based only on the prior information. It could be either maintaining the status quo or undertaking the risk reduction action. If collecting the failure data is evaluated as not desirable, then the prior decision suggested by the DSS should be considered as the optimal decision. (4) *Posterior E{# of Failure}*: the expected number of failures under status quo which is estimated by using both the prior information and the failure data. This value shows the performance of the system if no risk reduction action is considered when the prior knowledge of the system and the failure data are both applied to evaluate the system. (5) *Posterior decision*: the suggested decision is based on both the prior information and the failure data. It could be either maintaining the status quo or

undertaking the risk reduction action. Once the failure data is applied, the posterior decision suggested by the DSS should be considered as the optimal decision.

Once the decision area shows the decisions suggested by the DSS, the decision maker can perform further analysis to ensure the suggested optimal decisions are reliable. Sensitivity analysis can show the degree of importance for each prior parameter and study how they affect the optimal decisions. The DSS can provide one-way sensitivity analysis by using each element in the prior information as the changing factor. The results would be the ranges of the prior parameters that are of special interest in which the optimal decisions remain unchanged. The DSS also provides what-if analysis by changing the values of the prior parameters in the prior information and see if the optimal decisions are still unchanged or not.

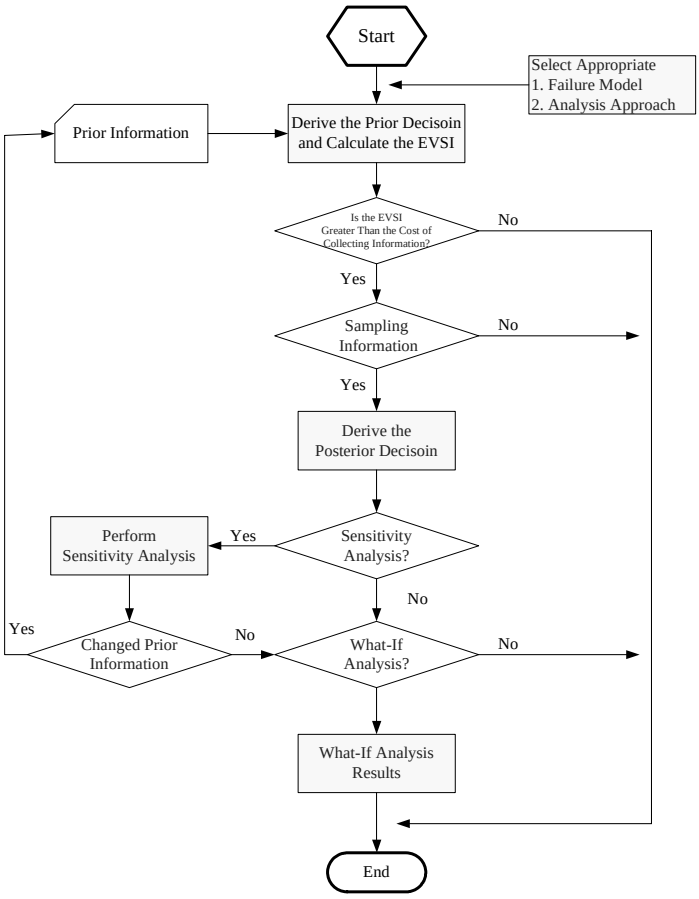

Figure 1: The follow chart of the DSS

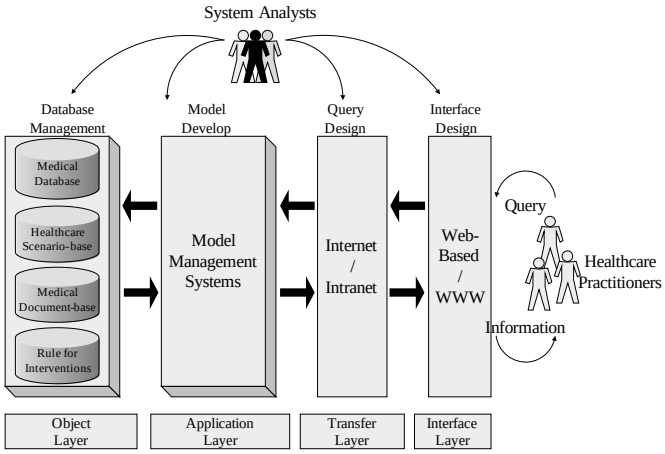

Figure 2: The framework of the DSS

4 Conclusions

Aging refers to the regular changes that occur in mature genetically representative organisms living under representative environmental conditions as they advance in the chronological age. Recurrent events or failures occur frequently in studies in which the failures are not necessarily fatal. Information is valuable because it reduces the expected costs of uncertainty surrounding a decision making. The expected costs of uncertainty are determined by the probability that a treatment decision based on existing information (prior knowledge) will be wrong and by the consequences if the wrong decision is made. This paper develops a possible Web-based decision support system by considering the sensitivity analysis as well as the optimal prior and posterior decisions for aging chronic diseases. It can deal with uncertain prior knowledge about the physiological system by considering the optimal prior decision, the sensitivity analysis, and possibly, the optimal posterior decision (if actual failure data were available), and provide decision makers the effective support for quality decision making.

References

[Aalen and Husebye, 91] Aalen, O. O. and Husebye, E.: "Statistical analysis of repeated events forming renewal processes"; Statistics in Medicine, 10 (1991), 1227–1240.

[Abouammoh and Qamber, 03] Abouammoh M. A. R. and Qamber, I. S.: "New better than renewal-used classes of life distributions"; IEEE Transactions on Reliability, 52, 2 (2003), 150-153.

[Aggarwal et al., 05] Aggarwal, P. S., Lowen, S. B., Colburn, H. S., Dolphin, W. F.: "Intrinsic oscillations in spike trains indicate non-renewal statistics due to convergence of inputs in dorsal cochlear nucleus neurons"; Hearing Research, 200 (2005), 10-28.

[Byar, (80)] Byar, D.: "Multiple tumour recurrence data for patients with bladder cancer" in data: a collection of problems from many fields for the student and research worker; Springer-Verlag / New York (1980).

[Calkins et al., (99)] Calkins E, Boult C, and Wagner, E.: "New ways to care for older people Building systems based on evidence"; Springer / New York (1999).

[Chang and Cheng, 06] Chang, C. C., and Cheng, C. S.: "A Bayesian decision analysis with Fuzzy interpretability for aging chronic disease"; International Journal of Technology Management, In press (2006).

[Chuang and Yadav, 98] Chuang, T. T., and Yadav, S. B.: "The development of an adaptive decision support system"; Decision Support System, 24 (1998), 73-87.

[Cook et al., 96] Cook, R. J., Lawess, J. F. and Nadeau, C.: "Robust test for treatment comparisons based on recurrent event responses"; Biometrics, 52 (1996), 557-571.

[Cook et al., 99] Cook, R. J., Edmund T. M., Jayanti M. and David V.: "Two-state mixed renewal processes for chronic disease"; Statistics in Medicine, 18 (1999), 175–188.

[Cox and Lewis, (66)] Cox, D. R. and Lewis, P. A. W.: "The Statistical Analysis of Series of Events"; Chapman and Hall / London (1966).

[Cox, 62] Cox, D. R.: "Renewal theory "; Methuen and Co / London (1962).

[Dewanji and Moolgavkar, 00] Dewanji, A. and Moolgavkar, S. H.: "A poisson process approach for recurrent event data with environmental covariates"; Environmetrics, 11 (2000), 665-673.

[Erto, 82] Erto, P.: "New practical Bayes estimators for the 2-parameter Weibull distribution"; IEEE Transactions on Reliability, 31, 2, (1982), 194–197.

[Gail et al., 80] Gail, M., Santner, T., and Brown, C.: "An analysis of comparative carcinogenesis experiments based on multiple times to tumour"; Biometrics, 36 (1980), 255-2662.

[Huang and Chang, 04] Huang, Y. S. and Chang, C. C.: "A study of defuzzification with experts' knowledge for deteriorating repairable systems"; European Journal of Operational Research, 157, 3 (2004), 658-670.

[International Chronic Granulomatous Disease Cooperative Study Group, 91] International Chronic Granulomatous Disease Cooperative Study Group: "A controlled trial of interferon gamma to prevent infection in chronic granulomatous disease"; New England Journal of Medicine, 324 (1991), 509-516.

[Klein et al., 89] Klein, J., Keiding, N., and Kamby, C.: "Semiparametric Marshall-Olkin models applied to the occurrence of metastases at multiple sites after breast cancer"; Biometrics, 45 (1989), 1073-1086.

[Lawless, 82] Lawless, J. F.: "Statistical Models and Methods for Lifetime Data"; John Wiley / New York (1989).

[Lawless, 87] Lawless, j.: "Regression methods for Poisson process data"; Journal of the American statistical association, 82, 399 (1987), 808-815.

[Nelson, 82] Nelson, W.: "Applied Life Data Analysis"; John Wiley / New York (1982).

[Thall and Lachin, 88] Thall, P., and Lachin, J.: "Analysis of recurrent events: nonparametric methods for random-interval count data"; Journal of the American statistical association, 83 (1988), 339-347.

[Wang et al., 00] Wang, C. C., Chen, M. L., Hsu, K. H., Lee, S. P., Chen, T. C., Chang, Y. S., Tsang, N. M., and Hong, J. H.: "Second malignant tumors in patients with nasopharyngeal carcinoma and their association with Epstein-Barr virus"; International Journal of Cancer, 87 (2000), 228-231.

[Watson and Spragure, (93)] Watson, H. J., and Spragure, R. H.: "The components of an architecture for DSS"; Prentice-Hall / NJ (1993).

[Wei et al., 89] Wei, L., Lin, D., and Weissfeld, L.: "Regression analysis of multivariate incomplete failure time data by modelling marginal distributions"; Journal of the American statistical association, 84 (1989), 1065-1073.

[Zenia et al., 99] Zenia, M., Agustin, N., and Pena, A.: "Order statistic properties, random generation, and goodness-of-fit testing for a minimal repair model"; Journal of the American statistical association, 94, 445 (1999), 266-272.