

CALCULATING THE EFFECT OF LENGTH  
BIASED SAMPLING ON SCREEN-DETECTED  
CASES IN RANDOMIZED CONTROLLED  
SCREENING TRIALS

by

Jean McKenzie Ethredge

B.S.Ed., University of Georgia, 1972

B.S., Metropolitan State College, 1992

A thesis submitted to the  
University of Colorado at Denver  
in partial fulfillment  
of the requirements for the degree of  
Master of Science  
Applied Mathematics  
December 2000

This thesis for the Master of Science

degree by

Jean McKenzie Ethredge

has been approved

by

---

Karen Kafadar

---

William Briggs

---

Craig Johns

---

Date

Ethredge, Jean McKenzie (M.S., Applied Mathematics)

Calculating the Effect of Length Biased Sampling on Screen-detected Cases  
in Randomized Controlled Screening Trials

Thesis directed by Professor Karen Kafadar

## ABSTRACT

Screening tests are used frequently for the early detection of diseases such as cancer. The benefit gained by these early detection tests can be assessed by comparing the mortality rates or survival rates between a study arm and a control arm in a randomized controlled screening trial. The comparison of the two arms is affected by two important biases: lead time bias and length bias. Lead time is the amount of time by which a preclinical diagnosis is advanced over a clinical diagnosis as a result of the screening test. If survival is measured from the time of diagnosis, then the comparison between study and control arms is biased by the lead time. Trial results affected by a lead time bias can indicate an increase in survival time of a study arm over a control arm even when there is no benefit at all from early screening. This bias can be eliminated by the use of randomized controlled screening trials where survival is measured since the entry into the study. Length-biased sampling, caused by periodic screening, leads to an over-representation in the study arm of cases having slower growing disease with better prognosis than in the general

population. This phenomenon can also lead to higher survival rates in the study arm versus the control arm, even in the absence of any screening benefit; but, unlike lead time bias, can not be eliminated by the study design as in a randomized screening trial.

In this thesis, I calculate the mean and variance of the increase in survival times that arise because of length biased sampling, when the sojourn times are gamma distributed. I further show that by ignoring the length bias, screening will appear more beneficial than it actually is. This bias must be considered to avoid over-optimistic conclusions about the benefit of screening programs.

This abstract accurately represents the content of the candidate's thesis. I recommend its publication.

Signed \_\_\_\_\_  
Karen Kafadar

## DEDICATION

To Jack and to our four daughters: Kristen, Alicia, Elizabeth and Jacquelyn.

## ACKNOWLEDGMENTS

It has been an honor and privilege to have done the calculations and research to complete this thesis under the direction of Karen Kafadar, my graduate adviser. I would like to thank her for her encouragement, patience, guidance and generosity.

I would also like to thank William Briggs for his meticulous reading of my thesis and his suggestions for improvement. Further thanks go to Craig Johns for serving on my committee and suggesting that I work with Karen in the first place.

Throughout the years of my mathematical studies, I have received valuable advice and assistance from Thomas Kelley, an undergraduate professor and from my friend Dave Brown. I would like to extend my appreciation to them both. Additional thanks go to Cary Miller and Randy Chase for teaching me enough about UNIX to make this thesis possible. To these and others, I extend my thanks.

Finally, I thank Jack who helped maintain our family and home in so many ways so “Mom” could pursue math.

I thank you all.

## CONTENTS

Figures . . . . .	ix
Tables . . . . .	x
1. Introduction . . . . .	1
1.1 Natural History of Cancer . . . . .	2
1.2 The Screening Process . . . . .	3
1.3 Lead Time Bias . . . . .	3
1.4 Length Bias . . . . .	4
1.5 The Accuracy of Screening Tests . . . . .	7
2. The Assessment of Screening Programs . . . . .	9
2.1 The HIP Breast Cancer Random Control Study . . . . .	10
2.2 Detection of Breast Cancers of the HIP Study . . . . .	12
2.3 Mortality in the HIP Trial . . . . .	12
2.4 Conclusion of the HIP Trial . . . . .	15
2.5 The CNBSS-2 Breast Cancer Random Control Study . . . . .	16
2.6 Detection of Breast Cancers by CNBSS . . . . .	17
2.7 Mortality in the CNBSS-2 Trial . . . . .	19
2.8 Conclusion of CNBSS-2 Trial . . . . .	20
3. The Exponential Distribution of Sojourn Times . . . . .	22
4. The Gamma Distribution of Sojourn Times . . . . .	38

4.1	The Gamma Distribution for Sojourn Times when $r = 2$ . . . . .	40
4.2	The Gamma Distribution for Sojourn Times for $r = 4$ . . . . .	49
5.	Conclusion . . . . .	52
	Notation Index . . . . .	53
	Appendix	
	A. . . . .	56
	References . . . . .	63

## FIGURES

1.1	Natural History of Cancer Progression . . . . .	2
1.2	Lead Time Bias . . . . .	4
1.3	Length-Biased Sampling or Length Bias . . . . .	6

## TABLES

2.1	HIP Compliance of the Study Arm by Number of Screening Exams [9, p. 19] . . . . .	12
2.2	HIP: Cumulative Detection of Breast Cancers by Year and Arm [9, p. 50] . . . . .	13
2.3	HIP: Cumulative Cases of Mortality Due to Breast Cancer .	14
2.4	HIP: Survival Rates [9, p. 71] . . . . .	15
2.5	CNBSS: Detection of Invasive Cancers [5, p. 1492] . . . . .	18
2.6	CNBSS: Mortality Cases by Cause to the End of 1993 by Study Arm [5, p. 1494] . . . . .	19
2.7	CCBSS: Cumulative Deaths due to Breast Cancer . . . . .	20
3.1	Exponential Distribution: Means for $j = 1, 2, 3, 4$ . . . . .	29
3.2	$E(Y^*)$ : The Mean Preclinical Duration of Cases Eligible for Detection by Screening . . . . .	30
3.3	Exponential Distribution Ratio Comparison: $E(Y^*)/E(Y)$ .	31
3.4	Relative Increase in the Mean: $(E(Y^*)/\mu - 1) \cdot 100\%$ . . . . .	31
3.5	$Var(Y^*)$ : Variances of Sojourn Times of Cases Eligible for Detection by Screening . . . . .	34
3.6	$SD(Y^*)$ : Standard Deviations of Sojourn Times of Cases Eligible for Detection by Screening . . . . .	34
3.7	Exponential Distribution Ratio $(SD(Y^*)/\mu)$ . . . . .	35

3.8	Relative Error in the Standard Deviations . . . . .	36
3.9	Exponential Distribution Summary of $Y^*$ when $r = 1$ . . . . .	37
4.1	Gamma Distribution $r = 2$ Means for $j = 1, 2, 3, 4$ . . . . .	42
4.2	Gamma $r = 2, E(Y^*)$ : The Mean Preclinical Duration of Cases Eligible for Detection by Screening with $\beta = 0.95$ . . . . .	42
4.3	Gamma $r = 2, E(Y^*)$ : The Mean Preclinical Duration of Cases Eligible for Detection by Screening with $\beta = 0.95$ . . . . .	43
4.4	Values of $E(Y_j^*)/\mu$ when $f_{Y^*}(y)$ is a Gamma Distribution $r =$ $2$ . . . . .	43
4.5	Gamma Distribution Ratio when $r = 2 E(Y^*)/\mu$ . . . . .	44
4.6	Gamma Distribution Relative Increase in the Means . . . . .	44
4.7	Gamma Formulas for the Expected Value of $Y_j^{2*}$ when $r = 2$ . . . . .	45
4.8	Gamma Formulas for the Expected Value Squared when $r =$ $2$ . . . . .	46
4.9	Gamma Distribution Variances for $r = 2, j = 1, 2, 3, 4$ . . . . .	47
4.10	Gamma Distribution $r = 2, \beta = 0.95 Var(Y^*)$ . . . . .	47
4.11	Gamma Distribution $r = 2$ , Standard Deviations of $Y^*$ . . . . .	48
4.12	Gamma Distribution $r = 2$ Ratios of Standard Deviations . . . . .	48
4.13	Gamma Distribution $r = 2$ ; Relative Error ( %) in Standard Deviations of $Y^*$ . . . . .	49
4.14	Summary of $Y^*$ when $r = 2$ : The Means, Relative Increase, Variances, Standard Deviations, and Relative Errors . . . . .	50
4.15	Summary of $Y^*$ when $r = 4$ : The Means, Relative Increase, Variances, Standard Deviations, and Relative Errors . . . . .	51

A.1	Exponential Distribution Ratios $E(Y_j^*)/\mu$ . . . . .	57
A.2	Exponential Distribution Variances: $Var(Y_j^*)$ . . . . .	57
A.3	Exponential Distribution Standard Deviations for $Y_j^*$ . . . . .	58
A.4	Exponential Distribution (Standard Deviations $(Y_j^*)$ )/ $\mu$ . . . . .	58
A.5	Gamma Distribution Means $r = 4$ for $j = 1, 2, 3, 4$ . . . . .	59
A.6	Gamma Distribution Ratio: $\frac{E(Y_j^*)}{\mu}$ when $r = 4$ . . . . .	60
A.7	Gamma Distribution $r = 4$ $\frac{E(Y_j^*)}{\mu}$ . . . . .	60
A.8	$r = 4$ Variances for $E(Y_j^*)$ for $j = 1, 2, 3, 4$ . . . . .	61
A.9	$r = 4$ Ratio of Sampled Standard Deviations/ $\frac{\mu}{2}$ . . . . .	62
A.10	Gamma Distribution $r = 4$ Ratio Standard Deviations of $Y^*/\sigma$ . . . . .	62

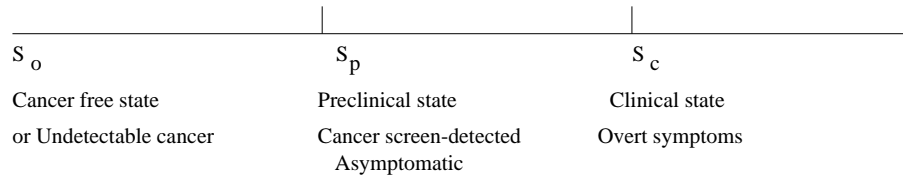
## **1. Introduction**

The expression “a stitch in time, saves nine” parallels a common belief that therapy is more effective when administered at an earlier stage of a developing cancer than at a later one. That early detection of cancer leads to treatment at an earlier stage of disease development and therefore to a decline in cancer mortality is a hypothesis, not a fact. Screening tests are developed for the early detection and diagnosis of disease. The value and effectiveness of screening programs that utilize the screening tests must be shown to either reduce the mortality rate or increase the survival time for the disease targeted by the screening. The benefits gained by early screen detection of a disease can be assessed through the use of randomized screening trials. The actual increase in survival time of participants offered screening over the survival time among participants not offered screening is one useful measure of screening effectiveness. Another measure is the reduction in mortality; but, the focus of this thesis is the former.

## 1.1 Natural History of Cancer

The idealized disease process is defined by using a three-state progressive disease model whereby an individual in a screened population is assumed to be in one of three states: the *disease free* state,  $S_o$ , the *preclinical* state,  $S_p$ , or the *clinical* state  $S_c$ [8]. Individuals in the disease free state,  $S_o$  are either free of the disease or have disease characteristics which are undetectable by a screening test. Participants in the preclinical state,  $S_p$ , do not have any apparent clinical symptoms, but have asymptomatic cancer that is detectable by a screening test. These individuals are unaware of their illness. The transition from the disease free into the preclinical state is assumed to take place at the first point in time at which a disease is detectable by the screening test. In the clinical stage,  $S_c$ , the disease is characterized by overt signs or symptoms leading to diagnosis. This state follows the preclinical state and is marked by the point of clinical diagnosis. The duration of the preclinical state is also called the *sojourn time*. The transition from  $S_o \rightarrow S_p \rightarrow S_c$  is the basic structure of the cancer screening models [4, p. 601].

**Figure 1.1.** Natural History of Cancer Progression



## 1.2 The Screening Process

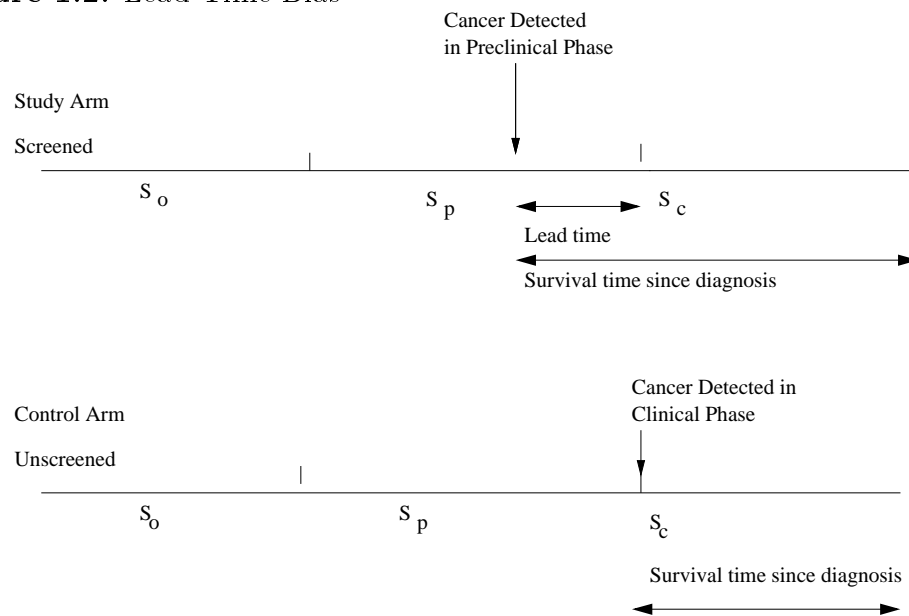
The screening process used for the detection of a disease involves periodic screenings of participants performed at regular intervals of time. The initial screen is designated as  $j = 0$  and marks the entry into a study. At the initial screening, all participants in a study are disease free. The time interval between screenings is generally assumed to be a constant. (In this thesis, the constant is denoted by  $\delta$ .) The term *screen-detected* refers to those individuals whose disease is detected in the preclinical stage by the screening process. Preclinical individuals collectively have a distribution of sojourn times. The screening process therefore samples from that distribution of the natural disease process; that is, the screen-detected individuals form a sampling of the preclinical durations. This sampling gives rise to several biases; two of which are: *lead time bias* which is a consequence of the sampling and *length bias* directly caused by sampling.

## 1.3 Lead Time Bias

Lead time is the time span by which a disease is diagnosed earlier as a result of screening than it would have been in the absence of screening. It is defined as the length of time by which the diagnosis is advanced over clinical detection by virtue of the screening procedure. The participant has the

opportunity to begin earlier treatment during the lead time interval because the diagnosis was made prior to the clinical phase of the disease. Survival time is automatically lengthened for cases detected by screening, even if there is no increased therapeutic benefit, when survival is measured from time of diagnosis. This phenomenon is known as a *lead time bias*.

**Figure 1.2.** Lead Time Bias



#### 1.4 Length Bias

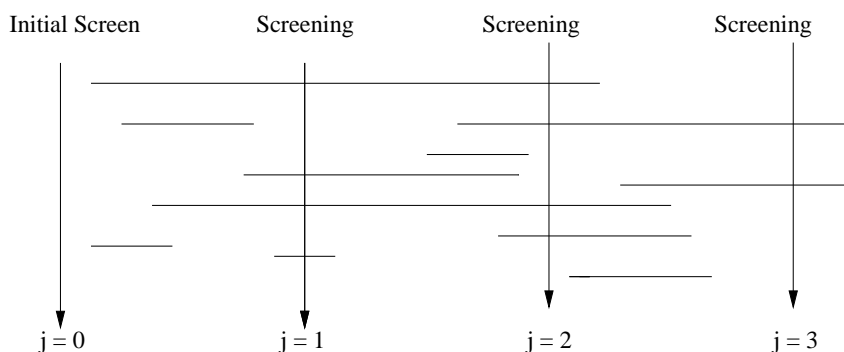
Length bias or length-biased sampling is another important, but subtle screening bias. This bias is a major factor in determining which preclinical disease cases will be detected early; that is, which cases will become part of the sample representing the distribution of preclinical durations. The cases of

cancer for all participants of a study in the preclinical phase of the natural disease process are not all equally likely to be detected by periodic screening. The sojourn time probability density function (pdf) for those screen-detected individuals is different from the sojourn time pdf for the population. The screening process does not detect people at random, but rather favors those with longer sojourn times; that is, the probability of being in the sample is a function of the sojourn time itself, and hence the sampled sojourn times do not represent a random sample from the individuals in the study population [2, p. 604].

The length of the preclinical duration, or sojourn time, varies from person to person due to the different growth rates of cancer. It is commonly believed that cases with long preclinical phases indicate a slowly advancing cancer, while cases with short preclinical phases indicate a rapidly spreading disease. With periodic screening, those cases having longer preclinical durations are more likely to be detected; that is, they have a higher probability of being detected than do cases of shorter preclinical durations. As a result, the longer duration preclinical cases are over-represented among the screen-detected cases. It is not unreasonable to assume that the clinical course of the disease is positively correlated with the preclinical course. Therefore the diagnostic screen will automatically select those individuals who are more likely to have longer survival times regardless of whether or not screening offers a

survival benefit due to early detection and treatment [2, p. 604]. If cancer advances slowly in the preclinical disease and then progresses to a slow-growing clinical disease, the cancer will tend to have characteristics of a good prognosis. Such cases would have more favorable outcomes even in the absence of screening; thus, studies may show increased survival rates or a decrease in the mortality rate because of this bias. Periodic sampling due to screening will affect the survival distribution by making the survival times appear longer than they would be in the general population, even in the absence of a screening benefit. In Figure 1.3 the horizontal lines represent sojourn times. Detection of cancer in the preclinical phase corresponds to a horizontal line intersecting a vertical screening line.

**Figure 1.3.** Length-Biased Sampling or Length Bias



Periodic screening occurs at regularly scheduled intervals. In cancer studies, the shorter sojourn time, those with poorer prognosis, may be under-represented.

## 1.5 The Accuracy of Screening Tests

Cancer screening is the testing of apparently healthy volunteers from the general population for the purpose of separating them into groups with high and low probabilities of having a given disorder [8, p. 225].

A screening program involves the designation and recruitment of participants, the performance of the screening test at certain ages or frequencies, and the provision for follow-up of suspicious and positive screening results, especially for the diagnosis and treatment of those testing positive for the disease.

A good screening test should possess properties of high sensitivity and high specificity. *Sensitivity* is the proportion of individuals designated positive by the screening test among all individuals who have the disease. *Specificity* is the proportion designated negative by the test among all those who do not have the disease. ( Low specificity leads to a large number of false positive cases and unnecessary treatment.) A screening test should have not only high specificity to avoid needless treatment of negative cases, but also high sensitivity to deliver the appropriate care to positive cases. The *predictive value* of a positive test is the proportion of individuals with a positive test who have the disease. The probability of correctly identifying those who have the disease should be high; in this thesis, test sensitivity will be denoted by  $\beta$ .

Once an effective screening test is developed it becomes part of a

screening program involving therapy and follow-up which then must be evaluated in terms of disease outcome.

## 2. The Assessment of Screening Programs

Breast cancer is the most frequent type of cancer diagnosed among women in the United States. Despite changes between 1940 and 1980 in economic and social conditions, nutritional status, health care, and fertility patterns, the mortality rate due to breast cancer had remained relatively constant. Approximately one woman in ten develops clinically detectable breast cancer in her lifetime [9, p. 8]. According to the National Center for Health Statistics, the death rate due to breast cancer is 23 deaths per 100,000 females [9, p. 2]. Mammography, clinical breast examinations, and breast self-examinations are three tests currently used in screening programs to detect breast cancer.

An objective method for evaluating the effectiveness of a screening program is a randomized control trial or RCT. There have been several long-term studies measuring the reduction in breast cancer mortality. Two of the most prominent studies, one conducted by Health Insurance Plan of New York (HIP) during the 1960's and another conducted in the 1990s by the National Cancer Institute of Canada, have led to some very interesting findings.

The HIP study compared the use of mammography plus clinical breast examination versus usual medical care and found a 30% reduction in breast cancer mortality over a ten-year period [9, p. 2]. The Canadian study compared mammography, clinical breast examination, and breast self-examination versus clinical and breast self-examination, but found no significant difference in the mortality rate of the two arms when mammography was added to the screening program [5].

## **2.1 The HIP Breast Cancer**

### **Random Control Study**

The HIP project was initiated by the National Cancer Institute in the United States to determine whether periodic breast cancer screening with mammography and clinical examination of the breast would lead to lower mortality. It was the first long-term randomized control trial of its kind with a follow-up ending eighteen years from the date of entry into the trial.

The participants for this project were randomly selected from 80,300 women, ages 40 - 64, covered by the HIP group insurance plan. The selected participants entered the project between December 1963 and June 1966. Every woman from the insurance plan who was eligible to be in the study arm received an initial mammogram and clinical breast examination followed by

annual reexaminations for three years. Of the 30,131 women comprising this arm, approximately one third (10,800 women) refused to participate; their results were nonetheless maintained as part of the study arm (intention-to-treat analysis).

The counterpart to this arm, the control arm, included 30,565 women who received an initial clinical examination of the breasts and then followed their own usual practices in obtaining medical care.

Sample sizes for the trial needed to be large enough to detect a 20% or greater reduction in breast cancer mortality at an alpha level of 0.05 [9]. The two arms were highly comparable; however, within the study arm, the characteristics of the “refusers” differed from the women in compliance. Because the arms were comparable, shifting this group of refusers to the control arm would have biased the trial. Although they did not participate, data were still maintained on them through insurance and other records.

Of the 30,131 women in the mammography and clinical examination group, 20,200 (66.8%) appeared for the initial screening [9, p. 19]. Half of all study arm participants had three or four screenings; see Table 2.1. Even in the face of such low compliance, the results of the HIP trial nonetheless showed that screening was beneficial.

**Table 2.1.** HIP Compliance of the Study Arm by Number of Screening Exams [9, p. 19]

Number of Exams	Number of Participants	Total (%) Study Arm	(%) Study Arm Participants
One or more	20,128	66.8	100.0
Two or more	17,476	58.0	86.8
Three or more	15,096	50.1	75.1
Four or more	11,932	39.6	59.3

## 2.2 Detection of Breast Cancers of the HIP Study

The follow-up of all those detected with breast cancer from both arms of the trial lasted eighteen years. By the fifth, sixth, and tenth year from entry into the trial, breast cancer cases in the study and control arms had “equalized” [9, p.60] (Table 2.2). It is at these points of equalization that the mortality and survival of the two arms can be compared.

## 2.3 Mortality in the HIP Trial

When the year of diagnosis was between one and five years after entry, the number of breast cancer deaths within the first five years in the study arm and control arms had accumulated to 39 and 63, respectively, implying a 38% reduction in mortality due to breast cancer. (When designing the sample size for their study, the Canadian researchers [5] used this approximately 40% figure as a target for achieving a desired power level of 0.80.) Within ten years from entry, the reduction in mortality was 28.6% or approximately 30%, a frequently

**Table 2.2.** HIP: Cumulative Detection of Breast Cancers by Year and Arm [9, p. 50]

Year from Entry	Study Arm Cases	Control Arm Cases	Cumulative Study Arm Cases	Cumulative Control Arm Cases
1	79	58	79	58
2	59	66	138	124
3	49	41	187	165
4	62	54	249	219
5	55	76	304	295
6	63	69	367	364
7	59	75	426	439
8	71	51	497	490
9	61	75	558	565
10	59	52	617	617
11	80	63	697	680
12	70	60	767	740
13	59	59	826	799
14	63	75	889	874
15	57	53	946	927

quoted percentage. By year 18, the reduction in mortality had dropped to 22.6 % [9, p. 65], (Table 2.3).

**Table 2.3.** HIP: Cumulative Cases of Mortality Due to Breast Cancer  
**Cumulative Deaths by Trial Arms**  
**Mortality Ratios, Percent Reduction in Mortality**  
**with 95% Confidence Intervals [9, p. 63]**

Years from Entry	Cumulative Study Arm Deaths	Cumulative Control Arm Deaths	$\frac{\text{Cum. Study}}{\text{Cum. Control}}$	Reduction in Mortality	95% Confidence Intervals
1 – 5	39	63	0.619	38.1	(8.9 – 59.5)
1 – 10	95	133	0.714	28.6	(7.4 – 45.5)
1 – 18	126	163	0.773	22.6	(2.7 – 39.0)

Most studies usually focus on mortality rates among patients with breast cancer rather than survival rates because of the biases due to lead time and length bias. The HIP project provided the first opportunity to develop and apply models for estimating lead time that utilized a randomized control group. The estimates were subject to large sampling errors and differed depending on the model used [9, p. 36]. The lead time estimates for this trial varied by age group. The estimated lead time for women between 40 and 49 was 5.2 months; for 50 – 59, 21.9 months and for women 60 – 64, there was no clear evidence of any lead time [9, p.36]. Survival rates calculated in the HIP project used an adjustment of one year for lead time [9, p.71]. Table 2.3 shows that survival among the study arm cases exceeded that in the control arm with or without

the lead time adjustment. No estimate of length-biased sampling was available to adjust the survival rate. This length bias is the focus of this thesis.

**Table 2.4.** HIP: Survival Rates [9, p. 71]

**Survival Rate Among Confirmed Breast Cancer Cases  
Diagnosed During the First Five Years after Entry By Arms**

Trial Arm	Number of Cases	Five Years from Entry	Ten Years from Entry
Study Arm	304	74	54.9
Control Arm	295	59.7	46.4
Study Arm Adj*	304	71.6	53.8

\*The lead time adjustment was used to calculate the survival rate.

## 2.4 Conclusion of the HIP Trial

In conclusion, by ten years from entry into the trial, there were about 30% fewer breast cancer deaths in the study arm than in the control arm; however, over the long term of eighteen years the reduction in the mortality rate was approximately 23 %. The HIP project provided strong evidence that periodic screening was beneficial, but the reduction in mortality as a result of mammography alone was not demonstrated.

## 2.5 The CNBSS-2 Breast Cancer Random Control Study

The Canadian National Breast Screening Study-2 (CNBSS-2) was designed to compare the incremental effect on the mortality rate of adding mammography to the screening program. So the study arm was offered annual screening using mammography, physical breast examination plus breast self-examination; the control arm was offered annual physical breast examination and breast self-examination only. The sample size of 40,000 participants was determined on the basis of an estimated power of 0.80 to detect a 40% reduction in breast cancer mortality, similar to the reduction observed in the HIP trial at five years from entry [5, p. 1491].

Participants for the study were recruited by general publicity, by personal invitation, group mailings, and through physicians. From January 1980 through March 1985, 39,405 women, 50 – 59 years of age, were randomly and individually assigned to one of the trial's arms. The study arm, denoted by M-Plus, was comprised of 19,711 women who received annual mammograms and clinical breast examinations. For the control arm (denoted BE), 19,694 women received annual clinical breast examination without mammography. Both groups were taught breast self-examination and were encouraged to use this technique regularly.

Five annual examinations were offered to the first 62% of the women

entering the trial and four examinations were offered to the remainder. Compliance in the study arm (M-plus) varied from 100% at the first screening to 86.7% by the fifth; 1.8% to 3.2% refused mammography at various screenings and 6.1% or 1196 participants had interval mammograms.

The compliance of the control arm, BE, after the first screen varied between 89.1% and 85.4% by the fifth screening; 16.9% (3300) of the participants of this group had one or more interval mammographs; 8% of the 3300 had mammograms between screens four and five. These statistics reflect a much better compliance than was observed in the HIP study.

## **2.6 Detection of Breast Cancers by CNBSS**

The last screens were conducted in 1988 (all clinics closed), but the annual follow-ups for all women known to have breast cancer continued until June 30, 1996. The mean follow-up from entry into the trial was 13 years. Cases of breast cancer detection were classified as either in-situ or invasive cancers. In the M-Plus arm, there were 71 in-situ cases in comparison to 16 in the BE arm. Invasive cancers were then classified into one of three categories: screen-detected, interval (those cases occurring within twelve months of a negative screening), and incident cancers (those cases occurring twelve or more months after the previous CNBSS screening examination).

**Table 2.5.** CNBSS: Detection of Invasive Cancers [5, p. 1492]

	<b>M-Plus</b>			<b>BE</b>		
	Screen Detected	Interval	Incident	Screen Detected	Interval	Incident
Year 1	118	114	0	64	16	0
Years 2 -5	149	36	32	84	72	47
Years 6 - 9	0		175			217
Total	267	50	207	148	88	264

The invasive cancers totaled 524 and 500 for the M-Plus and BE participants respectively (see Table 2.5). By December 1993, the totals rose to 622 and 610 in the respective groups. The number of invasive cancers approximately equalized by the end of the thirteenth year of the study [5, p. 1493]. The detection rates were higher in the study arm than in the control arm throughout screening, resulting in lead time for the M-Plus participants during which they received earlier diagnosis and treatment. Despite this earlier treatment, mortality rates between the two arms were almost identical (see section 2.2.2). The average lead time for M-Plus participants has been estimated to be 3.6 years (95.% CI: 2.7 - 5.5) in contrast to only 1.5 years for the BE. (CI: 2.7 - 5.5). The lead time gained by the study arm over the control arm was 2.1 years on average [5, p. 1492]. As a result of mammography, more diagnostic procedures were recommended and performed on the M-Plus arm than on the participants of the BE group. Interestingly, the cancers detected by mammography alone

were less likely to be lymph node positive than those detected by physical examination. Small tumors were less likely to be lymph node positive than large tumors. The M-Plus arm had a “higher biopsy rate for benign lesions and an excess of mastectomies due to uncertainty over the appropriate treatment” [5, p. 1496].

## 2.7 Mortality in the CNBSS-2 Trial

The total number of deaths from various causes were similar in the study and control groups (734 and 690 respectively) illustrating the comparability of the two arms (see Table 2.6) [5, p. 1494].

**Table 2.6.** CNBSS: Mortality Cases by Cause to the End of 1993 by Study Arm [5, p. 1494]

Causes	M-Plus	BE
Breast cancer	88	90
Other cancers	376	313
Non-cancer	270	287
Total	734	690

Among those diagnosed with breast cancer during the first five years after entry into the study, the mortality ratio,  $\frac{M-Plus}{BE}$ , is 1.09. Similar ratios resulted as deaths accumulated [5, p. 1495] ( see Table 2.7).

**Table 2.7.** CCBSS: Cumulative Deaths due to Breast Cancer  
 Number of deaths from breast cancer to June 30, 1996 by study arm and year  
 of detection

Year	M-Plus Deaths	BE only Death	Rate M-Plus/BE	95% Confidence Interval
Year 5	74	68	1.09	(0.78 - 1.51)
Year 6	84	76	1.10	(0.81 - 1.51)
Year 7	93	83	1.12	(0.83 - 1.50)
Year 8	99	89	1.11	(0.48 - 1.48)
Year 9	104	97	1.07	(0.81 - 1.41)
Beyond Year 9	107	105	1.02	(0.78 - 1.33)

## 2.8 Conclusion of CNBSS-2 Trial

The CNBSS-2 is the only trial that has evaluated the effect of mammography over and above physical breast examination and breast self-examination in women aged 50 - 59 at entry into the trial. The Canadian trial concluded that screening women in this age category with yearly mammography in addition to physical examination detected more lymph node negative and small breast cancers than screening with physical breast examination alone but had no impact on mortality from breast cancer [5, 1496]. This is the first study to show *lead time without benefit* for breast cancer screening [5, 1497].

In the HIP trial, 87.5 % of women diagnosed with impalpable cancers were still alive at ten years. This is in contrast to the Canadian trial showing 89.9% of the cancers detected in the BE arm experienced survival at ten years [5, 1497]. According to the Canadian trial, the survival in the HIP trial is

“almost certainly due to lead time and length bias....at least 70% of the benefit time may have come from the physical examinations” of the breasts rather than the mammography [5, 1497].

### 3. The Exponential Distribution of Sojourn Times

The beginning of a preclinical phase can not be determined unless screenings are conducted continuously rather than periodically. So the question is how can the mean length of the preclinical phase or the mean sojourn time be estimated given that the preclinical stage is not totally observable in the disease process? The clinical phase can be observed because symptoms are overt. (Many believe that the clinical and preclinical phases are positively correlated, but the correlation is unknown.) The survival time is also observable if the measurement is taken from the time of entry into the trial, given there are no detectable cases in the healthy population originally. This survival measurement would not be affected by the lead time bias because survival would no longer be measured from time of diagnosis.

Consider the length-biased density of sojourn times when only one screening is conducted. Let  $Y^*$  be the random variable that denotes the pre-clinical duration of a screen detected case and  $Y$  be the random variable that denotes the preclinical duration for a case in the general population.

The following notation will be used:

- $n$  = the number of observations.
- $Y_1, Y_2, \dots, Y_n$  are the sojourn times of the  $n$  cases.
- $f_Y(\cdot)$  = the probability density function (pdf) of preclinical durations.
- $f_{Y^*}(y)$  = the probability density function of the *sampled* preclinical durations.
- $\mu_Y$  = mean of  $(Y_1, Y_2, \dots, Y_n)$  i.e.,  $\mu = \int_0^\infty y f_Y(y) dy$
- $n_y$  = the number of preclinical durations,  $Y_i$ , with lengths  $y$

Then, the density of the sampled preclinical durations,  $f_{Y^*}(y)$ , can be derived from the pdf of *all* preclinical durations as follows: [4, p. 4]

$$\begin{aligned}
 f_{Y^*}(y) &= \lim_{n \rightarrow \infty} \left( \text{proportion of } \sum_{i=1}^n Y_i \text{ due to intervals of length } y \right) \\
 &= \lim_{n \rightarrow \infty} \frac{y n_y}{\sum_{i=1}^n Y_i} \\
 &= \lim_{n \rightarrow \infty} \frac{(n_y y) / n}{\sum_{i=1}^n Y_i / n} \\
 &= y \frac{f_Y(y)}{\mu_Y}.
 \end{aligned}$$

Thus, the density of  $Y^*$  is related to the density of  $Y$  [4, p. 4]. In this thesis, I compare the mean of the sampled preclinical durations with the mean of the population preclinical durations. If there were no effect from length-biased sampling, the mean from the sampled preclinical durations would equal

the population mean of preclinical durations. Thus, this ratio gives an estimate of the increase due to length-biased sampling and can be expressed in terms of the coefficient of variation as follows:

$$\frac{E(Y^*)}{E(Y)} = (1 + CV_Y^2) \text{ where the coefficient of variation } CV_Y = \frac{\sigma}{\mu}.$$

Now, consider screening occurring at regular or periodic intervals. The following notation is used to derive the cumulative distribution for preclinical durations at the first screening.

- $n$  = the number of observed preclinical durations.
- $\delta$  denotes the interval of time between screenings.
- $X$  represents the beginning of a preclinical phase. Assume that  $X$  is uniformly distributed over the interval  $[0, \delta]$  because the disease exhibits no particular preference for any time between the initial and subsequential screenings. [4].
- $Y$  denotes the continuous random variable of the preclinical duration of a case in the general population.
- $y_1, y_2, \dots, y_n$  denote the observed preclinical durations of a population.
- $f_Y(y)$  is the probability density function of all preclinical durations.
- $Y_j^*$  = is the continuous random variable of a screen-detected preclinical duration observed at the  $j^{th}$  screen,  $j = 1, 2, 3, 4$ .
- $F_{Y_j}(y)$  is the cumulative distribution function (cdf) for the preclinical

durations of cases detected at the  $j^{th}$  screening,  $j = 1, 2, \dots$

Cases that should be detected by the first screening will be those that meet the conditions:  $0 < X \leq \delta$  and  $X + Y > \delta$ . (If  $X + Y < \delta$ , the case is an interval case which has progressed to a clinical case at the time of screening.) Thus, the cdf of the sojourn times for these cases is derived as follows using a conditional probability that includes the conditions above.

$$\begin{aligned}
F_{Y_1}(y) &= P\{Y \leq y | X + Y > \delta \text{ and } 0 < X \leq \delta\} \\
&= \frac{\int_{\max(0, \delta-y)}^{\delta} P\{Y \leq y \text{ and } X + Y \geq \delta \text{ and } 0 < X \leq \delta \text{ and } X = x\} f_x(x) dx}{\int_0^{\delta} P\{X + Y > \delta \text{ and } 0 < x \leq \delta \text{ and } X = x\} f_x(x) dx} \\
&= \frac{\int_{\max(0, \delta-y)}^{\delta} P\{\delta - x \leq Y \leq y\} \frac{1}{\delta} dx}{\int_0^{\delta} P\{Y > \delta - x\} \frac{1}{\delta} dx} \\
&= \frac{yF_Y(y) - \int_0^y F(u) du}{\delta - \int_0^{\delta} F(u) du} \text{ where } y \leq \delta \\
&= \frac{\delta \cdot F_Y(y) - \int_0^{\delta} F(u) du}{\delta - \int_0^{\delta} F(u) du} \text{ where } y > \delta.
\end{aligned}$$

Thus, the pdf of  $Y_1^*$  and the mean,  $E(Y_1^*)$ , for the screenable sojourn times at the first screening can be calculated as follows: [4]

$$\begin{aligned}
f_{Y_1}(y) &= \frac{\min(y, \delta) \cdot f_Y(y)}{\delta - \int_0^{\delta} F(y) dy} \\
E(Y_1) &= \int y f_Y(y) dy \\
E(Y_1) &= \frac{1}{\delta - \int_0^{\delta} F(y) dy} \left[ \int_0^{\delta} y^2 f_Y(y) dy + \delta \int_{\delta}^{\infty} y f_Y(y) dy \right]
\end{aligned}$$

While the preclinical durations are not directly observable, studies

have suggested that the preclinical durations from the HIP study have a probability distribution that is approximately exponentially distributed [2]. This distribution is used frequently as a model for the distribution of times between the occurrence of successive events.

The exponential pdf, is the special case of the general gamma pdf with  $r = 1$ . Let the preclinical durations of a population,  $Y$ , have an *exponential distribution*.

$$f(y; \lambda) = \begin{cases} \lambda e^{-\lambda y} & \text{for } y \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

where  $\lambda > 0$ .

For some fixed value  $y$ , the probability that the observed value of  $Y$  will be at most  $y$  can be determined by employing the cumulative distribution function.

$$F(y; \lambda) = P\{Y \leq y\} = \int_0^y f(u; \lambda) du = \int_0^y \lambda e^{-\lambda u} du \text{ for } y \geq 0.$$

It follows that

$$F(y; \lambda) = \begin{cases} 1 - e^{-\lambda y} & \text{for } y \geq 0 \\ 0 & y < 0 \end{cases}$$

Using the exponential distribution for  $Y$  (unsampled) in the case of periodic screenings, the probability density function of the sojourn times detected at the  $j^{\text{th}}$  screening,  $f_{Y_j}(y; \lambda)$ , is derived in two intervals as follows:

$$f_{Y_j^*}(y; \lambda) = \begin{cases} \frac{[y - (j-1)\delta]\lambda e^{-\lambda y}}{\delta - [\int_0^{j\delta} (1 - e^{-\lambda y}) dy - \int_0^{(j-1)\delta} (1 - e^{-\lambda y}) dy]} & \text{for } (j-1)\delta \leq y \leq j\delta \\ \frac{\delta \lambda e^{-\lambda y}}{\delta - [\int_0^{j\delta} (1 - e^{-\lambda y}) dy - \int_0^{(j-1)\delta} (1 - e^{-\lambda y}) dy]} & \text{for } y > j\delta \end{cases}$$

$E(Y_j^*)$  represents the mean preclinical duration among cases that started before  $\delta$  and extended to at least  $j\delta$ , the time of the  $j^{\text{th}}$  screen. It is given by:

$$E(Y_j^*) = \int_{(j-1)\delta}^{j\delta} \left( y \frac{[y - (j-1)\delta]\lambda e^{-\lambda y}}{\delta - [\int_0^{j\delta} (1 - e^{-\lambda y}) dy - \int_0^{(j-1)\delta} (1 - e^{-\lambda y}) dy]} \right) dy + \int_{j\delta}^{\infty} \left( y \frac{\delta \lambda e^{-\lambda y}}{\delta - [\int_0^{j\delta} (1 - e^{-\lambda y}) dy - \int_0^{(j-1)\delta} (1 - e^{-\lambda y}) dy]} \right) dy.$$

using the notation  $D_{j\delta}$  for the denominator, we have,

$$D_{j\delta} = \delta - [\int_0^{j\delta} (1 - e^{-\lambda y}) dy - \int_0^{(j-1)\delta} (1 - e^{-\lambda y}) dy] = \frac{e^{\delta\lambda} - 1}{\lambda e^{\delta j\lambda}}.$$

Now,  $E(Y_j)$  can be simplified as:

$$\begin{aligned} E(Y_j) &= \int_{(j-1)\delta}^{j\delta} \left( y \frac{[y - (j-1)\delta]\lambda e^{-\lambda y}}{D_{j\delta}} \right) dy \\ &+ \int_{j\delta}^{\infty} \left( y \frac{\delta \lambda e^{-\lambda y}}{D_{j\delta}} \right) dy. \\ &= \frac{-2 - j\delta\lambda + e^{\delta\lambda}(2 - \delta\lambda + j\delta\lambda)}{\lambda(e^{\delta\lambda} - 1)} \end{aligned}$$

Letting the parameter  $\lambda = \frac{1}{\mu}$ , it follows that  $\lim_{\delta \rightarrow 0} E(Y_j) = \mu$ . That is, if screening occurred continuously instead of periodically ( $\delta \rightarrow 0$ ), then in the limit  $E(Y_j^*) \rightarrow E(Y) = \frac{1}{\lambda} = \mu$  and  $Var(Y_j) = \frac{1}{\lambda^2} = \mu^2$ .

Using *Mathematica* and different values for  $\mu$ , the values of  $E(Y_j)$  were calculated: (see Table 3.1).

Suppose the time interval between screenings is  $\delta = 2$  years. From Table 3.1, the mean length of the preclinical phase through the third screening, ( $j = 3$ ) given  $\mu = 0.5$ , is 4.96 years. Furthermore, with that table,  $E(Y^*)$ , the mean preclinical duration of all cases that are eligible for detection by screening given that four screenings have occurred can be calculated as follows:

$$\begin{aligned} E(Y^*) &= \sum_{j=0}^{\infty} E(Y_j^*) \cdot P\{\text{missed on } (j-1) \text{ previous screens,} \\ &\quad \text{detected on the } j^{\text{th}} \text{ screen}\} \\ &= \beta E(Y_1^*) + (1-\beta)\beta E(Y_2^*) + (1-\beta)^2\beta E(Y_3^*) + (1-\beta)^3\beta E(Y_4^*) \\ &\quad \text{where } \beta \text{ is the test sensitivity [4, p. 7].} \end{aligned}$$

The results of applying this equation to the matrices  $E(Y_j^*)$  for  $j = 1, 2, 3, 4$  and setting  $\beta$  equal to 0.95 follow in Table 3.2.

When the true mean preclinical duration is 1 year long, the sampled duration is 1.47 years long, almost a half of a year longer. Thus, even in the absence of any screening benefit, the survival time since diagnosis would be almost one half of a year longer. One might erroneously conclude an increase benefit from

**Table 3.1.** Exponential Distribution: Means for  $j = 1, 2, 3, 4$

$E(Y_j^*)$ : The mean length of the sampled sojourn times at the  $j^{\text{th}}$  screening.

		Values of $\delta$				
$E(Y_j^*)$	$\mu$	1	2	3	4	5
$E(Y_1^*)$	0.5	0.84	0.96	0.99	1.00	1.00
	1.0	1.42	1.69	1.84	1.93	1.97
	2.0	2.46	2.84	3.14	3.37	3.55
	5.0	5.48	5.93	6.35	6.73	7.09
	10.0	10.49	10.97	11.42	11.87	12.29
$E(Y_2^*)$	0.5	1.84	2.96	3.99	5.00	6.00
	1.0	2.42	3.69	4.84	5.93	6.97
	2.0	3.46	4.84	6.14	7.37	8.55
	5.0	6.48	7.93	9.35	10.74	12.09
	10.0	11.49	12.97	14.43	15.87	17.29
$E(Y_3^*)$	0.5	2.84	4.96	6.99	9.00	11.00
	1.0	3.42	5.69	7.84	9.93	11.97
	2.0	4.46	6.84	9.14	11.37	13.55
	5.0	7.48	9.93	12.35	14.74	17.09
	10.0	12.49	14.97	17.43	19.87	22.29
$E(Y_4^*)$	0.5	3.84	6.96	9.99	13.00	16.00
	1.0	4.42	7.69	10.84	13.93	16.97
	2.0	5.46	8.84	12.14	15.37	18.55
	5.0	8.48	11.93	15.35	18.74	22.09
	10.0	13.49	16.97	20.43	23.87	27.29

**Table 3.2.**  $E(Y^*)$ : The Mean Preclinical Duration of Cases Eligible for Detection by Screening

		Values of $\delta$				
$\mu$		1	2	3	4	5
0.5		0.90	1.07	1.15	1.210	1.26
1.0		1.47	1.79	2.00	2.14	2.23
2.0		2.51	2.94	3.30	3.58	3.82
5.0		5.54	6.04	6.51	6.95	7.35
10.0		10.54	11.07	11.58	12.08	12.56

screening. The table of ratios of  $E(Y_j^*)$  divided by  $\mu$  for  $j = 1, 2, 3, 4$  is given in the appendix (see table A.11). Using the four matrices within that table, the ratios:

$$E(Y^*)/\mu = \beta E(Y_1^*)/\mu + (1-\beta)\beta E(Y_2^*)/\mu + (1-\beta)^2\beta E(Y_3^*)/\mu + (1-\beta)^3\beta E(Y_4^*)/\mu$$

are given in Table 3.3.

Using the ratios in Table 3.3, the relative increase in the estimation of the sampled mean can be calculated by subtracting one from each value and multiplying by 100% to obtain percentages. The results are shown in Table 3.4.

**Table 3.3.** Exponential Distribution Ratio Comparison:  $E(Y^*)/E(Y)$ 

$\mu$	Values of $\delta$				
	1	2	3	4	5
0.5	1.79	2.14	2.30	2.42	2.53
1.0	1.47	1.79	2.00	2.14	2.23
2.0	1.26	1.47	1.65	1.79	1.91
5.0	1.11	1.21	1.30	1.39	1.47
10.0	1.05	1.11	1.16	1.21	1.26

**Table 3.4.** Relative Increase in the Mean:  $(E(Y^*)/\mu - 1) \cdot 100\%$ 

$\mu$	Values of $\delta$				
	1	2	3	4	5
0.5	79	113	130	142	153
1.0	47	79	100	114	123
2.0	26	47	65	79	91
5.0	11	21	30	39	47
10.0	5	11	16	21	26

From Table 3.4, when screening occurs every two years and the true mean sojourn time of the population is five years, the average sojourn time among the screen-detected cases is 21% larger than the overall mean. As the screening interval increases, the estimations of the preclinical durations get even larger. (Intuitively, diseases with shorter durations would need more frequent

screenings.)

The formula for the variance or the mean squared distance of the sojourn times from the mean is

$$\begin{aligned}
\sigma_Y^2 &= \text{Var}(Y) \\
&= \int_0^\infty (y - \mu_Y)^2 f_Y(y) dy \text{ for } y \geq 0. \\
&= \int_0^\infty (y - E(Y))^2 f_Y(y) dy \text{ for } y \geq 0.
\end{aligned}$$

The variance for the exponential distribution at the  $j^{\text{th}}$  screening becomes:

$$\begin{aligned}
\text{Var}(Y_j^*) &= \int y^2 f_{Y_j}(y) dy - \left( \int y f_{Y_j}(y) dy \right)^2 \\
&= E(Y_j^{2*}) - [E(Y_j^*)]^2 \\
&= \int_{(j-1)\delta}^{j\delta} (y^2 (y - (j-1)\delta) \lambda e^{-\lambda y}) / D_{j\delta} + \int_{j\delta}^\infty (y^2 \delta \lambda e^{-\lambda y}) / D_{j\delta} \\
&\quad - \left( \int_{(j-1)\delta}^{j\delta} (y (y - (j-1)\delta) \lambda e^{-\lambda y}) / D_{j\delta} + \int_{j\delta}^\infty (y \delta \lambda e^{-\lambda y}) / D_{j\delta} \right)^2 \\
&= \frac{2 - e^{\delta\lambda} (4 - 2e^{\delta\lambda} + \delta^2 \lambda^2)}{\lambda^2 (e^{\delta\lambda} - 1)^2} \\
&\quad \text{where } D_{j\delta} = \delta - \left[ \int_0^{j\delta} (1 - e^{-\lambda y}) dy - \int_0^{(j-1)\delta} (1 - e^{-\lambda y}) dy \right] = \\
&\quad \frac{e^{\delta\lambda} - 1}{\lambda e^{\delta j \lambda}} \\
&= \frac{2 - e^{\delta\lambda} (4 - 2e^{\delta\lambda} + \delta^2 \lambda^2)}{\lambda^2 (e^{\delta\lambda} - 1)^2} \\
&= \frac{-\delta^2 e^{\delta/\mu} + \mu^2 (2 - 4e^{\delta/\mu} + 2e^{2\delta/\mu})}{(e^{\delta/\mu} - 1)^2} \text{ where } \lambda = 1/\mu.
\end{aligned}$$

Note that the variance of sojourn times is independent of  $j$  so  $Y_1^* = Y_2^* = Y_3^* = Y_4^*$  and is therefore equal to  $Y^*$ . In this thesis, the calculation for  $Y^*$  is made anyway using equation 3.3 verifying this fact (Table 3.5).  $Var(Y^*)$  is calculated in the same manner as  $E(Y^*)$ , again letting  $\beta$  be the test sensitivity, we have [4, p. 7]

$$Var(Y^*) = \sum_{j=0}^{\infty} Var(Y_j)P\{\text{missed on } (j-1) \text{ previous screens,} \quad (3.1)$$

$$\text{detected on } j^{\text{th}} \text{ screen}\} \quad (3.2)$$

$$= \beta Var(Y_1^*) + (1-\beta)\beta Var(Y_2^*) + (1-\beta)^2 \beta Var(Y_3^*) \quad (3.3)$$

$$+ (1-\beta)^3 \beta Var(Y_4^*)$$

The variances of the sampled preclinical durations,  $Y^*$ , are shown in Table 3.5; they are indeed the same as the variances for each  $Y_j^*$  for the various values of  $\mu$ . From the variances, the standard deviations are calculated by taking the square root. ( $SD(Y^*) = [Var(Y^*)]^{\frac{1}{2}}$ ) (Table 3.6)

**Table 3.5.**  $Var(Y^*)$ : Variances of Sojourn Times of Cases Eligible for Detection by Screening

		Values of $\delta$				
$\mu$	1	2	3	4	5	
0.5	0.32	0.42	0.48	0.49	0.50	
1.0	1.08	1.28	1.50	1.70	1.83	
2.0	4.08	4.32	4.67	5.10	5.56	
5.0	25.08	25.33	25.74	26.29	26.98	
10.0	100.08	100.33	100.75	101.32	102.06	

**Table 3.6.**  $SD(Y^*)$ : Standard Deviations of Sojourn Times of Cases Eligible for Detection by Screening

		Values of $\delta$				
$\mu$	1	2	3	4	5	
0.5	0.56	0.65	0.69	0.70	0.71	
1.0	1.04	1.13	1.23	1.30	1.35	
2.0	2.02	2.08	2.16	2.26	2.36	
5.0	5.01	5.03	5.07	5.13	5.19	
10.0	10.00	10.01	10.04	10.07	10.10	

**Table 3.7.** Exponential Distribution Ratio  $(SD(Y^*))/\mu$ 

$\mu$	Values of $\delta$				
	1	2	3	4	5
0.5	1.13	1.30	1.38	1.41	1.41
1.0	1.04	1.13	1.23	1.30	1.35
2.0	1.01	1.04	1.08	1.13	1.18
5.0	1.00	1.01	1.01	1.03	1.04
10.0	1.00	1.00	1.00	1.01	1.01

Subtracting one from the values in Table 3.7 and multiplying by 100 will produce the relative error, in percentages, between the standard deviation of the sampling distribution and the population distribution of sojourn times. These values are shown in Table 3.8. Note that the relative error in the standard deviations is smaller than that in the means but nonetheless shows an increase. As with the means (Table 3.4), the relative error increases as  $\delta$  increases for a given value of  $\mu$  and decreases as  $\mu$  increases for a given value of  $\delta$ . That is, the relative error, in both the mean and standard deviation increases as the screening interval increases and as the mean preclinical durations decrease.

**Table 3.8.** Relative Error in the Standard Deviations

$\mu$	Values of $\delta$				
	1	2	3	4	5
0.5	13	30	38	41	41
1.0	4	13	23	30	35
2.0	1	4	8	13	18
5.0	0	1	1	3	4
10.0	0	0	0	1	1

Table 3.9 is a summary of the findings for the exponential distribution. As the interval between screenings gets larger, the estimation of the sample means also gets larger. The relative increase also gets larger. For example: When  $\delta = 1$  and  $\mu = 0.5$ , the mean of  $Y^*$  is estimated to be 0.90 or 79% larger than the overall mean of  $Y$ . The relative error in standard deviations gets larger for each value of  $\mu$  as  $\delta$  increases but decreases as  $\mu$  gets larger.

**Table 3.9.** Exponential Distribution Summary of  $Y^*$  when  $r = 1$   
The Means, Relative Increase, Variances, Standard Deviations, and Relative Errors  
**The Preclinical Durations of Cases Eligible for Detection by Screening**

		Values of $\delta$				
$E(Y^*)$	$\mu$	1	2	3	4	5
Sampled Means	0.5	0.90	1.07	1.15	1.21	1.26
	1.0	1.47	1.79	2.00	2.14	2.23
	2.0	2.51	2.94	3.30	3.58	3.82
	5.0	5.54	6.04	6.51	6.95	7.35
	10.0	10.54	11.07	11.58	12.08	12.56
Relative Increase	$\mu$	1	2	3	4	5
$(\frac{E(Y^*)}{\mu} - 1)100$	0.5	79	113	130	142	153
	1.0	47	79	100	114	123
	2.0	26	47	65	79	91
	5.0	11	21	30	39	47
	10.0	5	11	16	21	26
$Var(Y^*)$	$\mu$	1	2	3	4	5
Variances of sampled sojourn times $E(Y^{*2}) - (E(Y^*))^2$	0.5	0.32	0.42	0.48	0.49	0.50
	1.0	1.08	1.28	1.50	1.70	1.83
	2.0	4.08	4.32	4.67	5.10	5.56
	5.0	25.08	25.33	25.74	26.29	26.98
	10.0	100.08	100.33	100.75	101.32	102.06
$sd(Y^*)$	$\mu$	1	2	3	4	5
Standard Deviations of sampled sojourn times	0.5	0.56	0.65	0.69	0.70	0.71
	1.0	1.04	1.13	1.23	1.30	1.35
	2.0	2.02	2.08	2.16	2.26	2.36
	5.0	5.01	5.03	5.07	5.13	5.19
	10.0	10.00	10.01	10.04	10.07	10.10
Relative Error	$\mu$	1	2	3	4	5
$(sd(Y^*)/\mu - 1)100$	0.5	13	30	38	41	41
	1.0	4	13	23	30	35
	2.0	1	4	8	13	18
	5.0	0	1	1	3	4
	10.0	0	0	0	1	1

#### 4. The Gamma Distribution of Sojourn Times

The gamma distribution is used to model waiting times or lifetimes such as human longevity after age twenty or so [7, p. 189]. It is appropriate to use this distribution as a model for preclinical durations. The exponential pdf is a special case of the general gamma probability density function for  $r = 1$ ; therefore, this chapter will consider the cases  $r = 2$  and  $r = 4$ . Once again, let the preclinical durations of a population be represented by the continuous random variable  $Y$ . This random variable is said to have a *gamma distribution* if the pdf of  $Y$  is

$$f(y; r, \lambda) = \begin{cases} \frac{\lambda^r y^{r-1} e^{-\lambda y}}{\Gamma(r)} & \text{for } y \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

where  $r > 0$  and  $\lambda > 0$ . The gamma function is defined by  $\Gamma(r) = \int_0^\infty y^{r-1} e^{-y} dy$  but for any positive integer  $r$ ,  $\Gamma(r) = (r - 1)!$ ; thus, for  $r = 2$ ,  $\Gamma(2) = 1$  and  $\Gamma(4) = 6$ . The frequencies of preclinical durations that are  $y$  units of time long are given by the gamma distribution.

The probability that the observed value of  $Y$  will be at most  $y$  for some fixed value  $y$ , can be determined by use of the cumulative distribution function

(cdf) which is obtained by integrating the gamma distribution as shown:

$$F(y; r, \lambda) = P(Y \leq y) = \int_0^y f(u; r, \lambda) du = \int_0^y \frac{\lambda^r u^{r-1} e^{-\lambda u}}{\Gamma(r)} du$$

for  $y \geq 0$ .

$$F(y; r, \lambda) = \begin{cases} 1 - \frac{1 + \lambda y}{e^{\lambda y}} & \text{for } y \geq 0 \text{ and } r = 2 \\ 0 & y < 0 \end{cases}$$

When the distribution of unsampled preclinical durations ( $Y$ ) is gamma distributed and screenings are periodic, the pdf of the sojourn times detected at the  $j^{\text{th}}$  screening,  $f_{Y_j}(y; r, \lambda)$ , is calculated over two intervals,

$(j-1)\delta \leq y \leq j\delta$  and  $j\delta < y < \infty$ :

$$f_{Y_j}(y; r, \lambda) = \begin{cases} \frac{[y - (j-1)\delta] \lambda^r y^{r-1} e^{-\lambda y} / \Gamma(r)}{\delta - [\int_0^{j\delta} (-1 + e^{\lambda y} - \lambda y) / e^{\lambda y} dy - \int_0^{(j-1)\delta} (-1 + e^{\lambda y} - \lambda y) / e^{\lambda y} dy]} & \text{for } (j-1)\delta \leq y \leq j\delta \\ \frac{\delta \lambda^r y^{r-1} e^{-\lambda y} / \Gamma(r)}{\delta - [\int_0^{j\delta} (-1 + e^{\lambda y}) - \lambda y) / dy - \int_0^{(j-1)\delta} (-1 + e^{\lambda y} - \lambda y) / e^{\lambda y} dy]} & \text{for } y > j\delta. \end{cases}$$

The denominator of the above density function is denoted by  $D_{j\delta}$

where  $D_{j\delta} = \frac{e^{\delta\lambda}(2 - \delta\lambda + \delta j\lambda) - \delta\lambda j - 2}{\lambda e^{\delta j\lambda}}$ . Using this notation the density

function becomes

$$f_{Y_j}(y; r, \lambda) = \begin{cases} \frac{[y - (j-1)\delta] \lambda^r y^{r-1} e^{-\lambda y} / \Gamma(r)}{D_{j\delta}} & \text{for } (j-1)\delta \leq y \leq j\delta \\ \frac{\delta \lambda^r y^{r-1} e^{-\lambda y} / \Gamma(r)}{D_{j\delta}} & \text{for } y > j\delta \end{cases}$$

#### 4.1 The Gamma Distribution for Sojourn Times when $r = 2$

The main purpose of this section is to generate the mean sojourn time of screen-detected cases and compare them to the overall sojourn times, as well as calculate the percentage increase between the two. In this section, all calculations use the gamma distribution with  $r = 2$ . In the next section, rather than repeat the exact same steps, only the summary results are given for  $r = 4$ . Tables from which the summary is derived for  $r = 4$  are located in the appendix.

The first goal is to calculate the sampled mean of sojourn times,  $E(Y^*)$ . To do so, each  $E(Y_j)$  for  $j = 1, 2, 3$ , and 4 are needed, where  $E(Y_j)$  represents the mean preclinical duration among cases that started before  $\delta$  but no later than  $j\delta$ . The expected value or mean of the distribution of sojourn times,  $\int y f_{Y_j}(y) dy$ , by the  $j^{\text{th}}$  screening, is given by

$$E(Y_j) = \int_{(j-1)\delta}^{j\delta} \left( y \frac{[y - (j-1)\delta] \lambda^r y^{r-1} e^{-\lambda y} / \Gamma(r)}{D_{j\delta}} \right) dy + \int_{j\delta}^{\infty} \left( y \frac{\delta \lambda^r y^{r-1} e^{-\lambda y} / \Gamma(r)}{D_{j\delta}} \right) dy.$$

which becomes for any  $j > 0$ ,

$$E(Y_j) = \frac{6 + \delta\lambda(-4 + 4j + \delta\lambda - 2\delta\lambda j + j^2\delta^2\lambda^2)}{\lambda(2 - \delta\lambda + \delta j\lambda)} + \frac{-2\delta + \delta^2\lambda(2 - 4j + \delta j\lambda - \delta j^2\lambda)}{(2 - \delta\lambda + \delta j\lambda)(-2 - \delta j\lambda + e^{\delta\lambda}(2 - \delta\lambda + \delta j\lambda))}$$

Screenings performed continuously are found by letting  $\delta \rightarrow 0$ . We find that  $\lim_{\delta \rightarrow 0} E(Y_j) = \mu$  when  $\lambda = \frac{2}{\mu}$ . The table of means (Table 4.1) was generated using *Mathematica* with different values for  $\mu$  and setting  $\lambda = \frac{2}{\mu}$  in the equation for  $E(Y_j)$  above.

Repeating the process that was used for the exponential distribution, the expected preclinical durations for the cases eligible for detection by screening was calculated using Table 4.1 to produce Table 4.3.

$$\begin{aligned}
 E(Y^*) &= \sum_{j=0}^{\infty} E(Y_j^*) \cdot P\{\text{missed on } (j-1) \text{ previous screens,} \\
 &\quad \text{detected on the } j^{\text{th}} \text{ screen}\} \\
 &= \beta E(Y_1^*) + (1-\beta)\beta E(Y_2^*) + (1-\beta)^2\beta E(Y_3^*) + (1-\beta)^3\beta E(Y_4^*) \\
 &\quad \text{where } \beta \text{ is the test sensitivity [4, p. 7] (See Table 4.3) .}
 \end{aligned}$$

From Table 4.3, the ratios of  $E(Y^*)/\mu$ , are calculated producing Table 4.5. These ratios being greater than one indicate that the means of the sampled durations are larger than the overall means, thus illustrating that the sample is length-biased.

The relative increase in the means for the gamma density is obtained by subtracting one from each value above and multiplying by 100 to obtain percentages. The percentage of increase in the sojourn times for  $r = 2$  over the actual expected preclinical durations are found in Table 4.6.

The variance of preclinical durations at the  $j^{\text{th}}$  screening follows using the

**Table 4.1.** Gamma Distribution  $r = 2$  Means for  $j = 1, 2, 3, 4$

		Values of $\delta$				
$E(Y_j^*)$	$\mu$	1	2	3	4	5
$E(Y_1^*)$	0.5	0.70	0.75	0.75	0.75	0.75
	1.0	1.22	1.40	1.47	1.49	1.50
	2.0	2.18	2.44	2.66	2.81	2.90
	5.0	5.10	5.32	5.58	5.85	6.11
	10.0	10.06	10.20	10.40	10.64	10.90
$E(Y_2^*)$	0.5	1.55	2.55	3.54	4.53	5.52
	1.0	2.02	3.11	4.11	5.10	6.08
	2.0	2.86	4.03	5.15	6.21	7.23
	5.0	5.56	6.59	7.74	8.92	10.08
	10.0	10.35	11.13	12.10	13.19	14.32
$E(Y_3^*)$	0.5	2.52	4.53	6.52	8.51	10.51
	1.0	2.96	5.05	7.06	9.05	11.05
	2.0	3.73	5.91	8.03	10.10	12.12
	5.0	6.24	8.26	10.42	12.61	14.78
	10.0	10.84	12.48	14.43	16.52	18.67
$E(Y_4^*)$	0.5	3.51	6.52	9.51	12.51	15.51
	1.0	3.93	7.02	10.04	13.04	16.03
	2.0	4.66	7.85	10.98	14.05	17.07
	5.0	7.03	10.07	13.25	16.45	19.63
	10.0	11.44	14.06	17.03	20.14	23.31

**Table 4.2.** Gamma  $r = 2, E(Y^*)$ : The Mean Preclinical Duration of Cases Eligible for Detection by Screening with  $\beta = 0.95$

		Values of $\delta$				
$\mu$		1	2	3	4	5
0.5		0.75	0.84	0.90	0.95	1.00
1.0		1.26	1.49	1.61	1.68	1.74
2.0		2.22	2.53	2.79	2.99	3.13
5.0		5.13	5.39	5.70	6.01	6.32
10.0		10.07	10.25	10.49	10.78	11.08

**Table 4.3.** Gamma  $r = 2, E(Y^*)$ : The Mean Preclinical Duration of Cases Eligible for Detection by Screening with  $\beta = 0.95$

Values of $\delta$					
$\mu$	1	2	3	4	5
0.5	0.75	0.84	0.90	0.95	1.00
1.0	1.26	1.49	1.61	1.68	1.74
2.0	2.22	2.53	2.79	2.99	3.13
5.0	5.13	5.39	5.70	6.01	6.32
10.0	10.07	10.25	10.49	10.78	11.08

**Table 4.4.** Values of  $E(Y_j^*)/\mu$  when  $f_{Y^*}(y)$  is a Gamma Distribution  $r = 2$ .

Values of $\delta$						
Screen	$\mu$	1	2	3	4	5
$j = 1$	0.5	1.40	1.49	1.50	1.50	1.50
	1.0	1.22	1.40	1.47	1.49	1.50
	2.0	1.09	1.22	1.33	1.40	1.45
	5.0	1.02	1.06	1.17	1.17	1.22
	10.0	1.01	1.02	1.04	1.06	1.09
$j = 2$	0.5	3.11	5.10	7.07	9.06	11.05
	1.0	2.02	3.11	4.11	5.10	6.08
	2.0	1.43	2.02	2.58	3.11	3.61
	5.0	1.11	1.32	1.55	1.78	2.02
	10.0	1.04	1.11	1.21	1.32	1.43
$j = 3$	0.5	5.05	9.05	13.04	17.03	21.02
	1.0	2.96	5.05	7.06	9.05	11.15
	2.0	1.87	2.96	4.02	5.05	6.06
	5.0	1.25	1.65	2.08	2.52	2.96
	10.0	1.08	1.25	1.44	1.65	1.87
$j = 4$	0.5	7.02	13.04	19.03	25.02	31.02
	1.0	3.93	7.02	10.04	10.04	16.03
	2.0	2.33	3.93	5.49	7.02	8.54
	5.0	1.41	2.01	2.65	3.29	3.93
	10.0	1.14	1.41	1.70	2.01	2.33

**Table 4.5.** Gamma Distribution Ratio when  $r = 2 E(Y^*)/\mu$

$\mu$	Values of $\delta$				
	1	2	3	4	5
0.5	1.49	1.68	1.79	1.90	2.00
1.0	1.26	1.49	1.61	1.68	1.74
2.0	1.11	1.26	1.40	1.49	1.56
5.0	1.03	1.08	1.14	1.20	1.26
10.0	1.01	1.03	1.05	1.08	1.11

**Table 4.6.** Gamma Distribution Relative Increase in the Means

$\mu$	Values of $\delta$				
	1	2	3	4	5
0.5	49	68	79	90	100
1.0	26	49	61	68	74
2.0	11	26	40	49	56
5.0	3	8	14	20	26
10.0	1	3	5	8	11

**Table 4.7.** Gamma Formulas for the Expected Value of  $Y_j^{2*}$  when  $r = 2$

$E(Y_j^{2*})$	Equations
$E(Y_1^{2*})$	$\frac{24 - 24e^{\delta\lambda} + 18\delta\lambda + 6\delta^2\lambda^2 + \delta^3\lambda^3}{\lambda^2(2 - 2e^{\delta\lambda} + \delta\lambda)}$
$E(Y_2^{2*})$	$\frac{-24 + e^{\delta\lambda}(24 + 18\delta\lambda + 6\delta^2\lambda^2 + \delta^3\lambda^3) - 4\delta\lambda(9 + 6\delta\lambda + 2\delta^2\lambda^2)}{\lambda^2(-2 - 2\delta\lambda + e^{\delta\lambda}(2 + \delta\lambda))}$
$E(Y_3^{2*})$	$\frac{-24 - 27\delta\lambda[2 + \delta\lambda(2 + \delta\lambda)] + 4e^{\delta\lambda}[6 + \delta\lambda(9 + \delta\lambda + 2\delta^2\lambda^2)]}{\lambda^2[-2 - 3\delta\lambda + 2e^{\delta\lambda}(1 + \delta\lambda)]}$
$E(Y_4^{2*})$	$\frac{8(-3 - 9\delta\lambda - 12 - \delta^2\lambda^2 - 18\delta^3\lambda^3) + 3e^{\delta\lambda}[8 + 9\delta\lambda(2 + 2\delta\lambda + \delta^2\lambda^2)]}{\lambda^2[-2 - 4\delta\lambda + e^{\delta\lambda}(2 + 3\delta\lambda)]}$

gamma distribution:

$$\begin{aligned}
 Var(Y_j) &= \int y^2 f_{Y_j}(y) dy - \left( \int y f_{Y_j}(y) dy \right)^2 \\
 &= E(Y^2) - [E(Y)]^2 \\
 &= \int_{(j-1)\delta}^{j\delta} \frac{y^2(y - (j-1)\delta)\lambda^r y^{r-1} e^{-\lambda y} / \Gamma(r)}{D_{j\delta}} + \int_{j\delta}^{\infty} \frac{y^2 \delta \lambda^r y^{r-1} e^{-\lambda y} / \Gamma(r)}{D_{j\delta}} \\
 &\quad - \left( \int_{(j-1)\delta}^{j\delta} \frac{y(y - (j-1)\delta)\lambda^r y^{r-1} e^{-\lambda y} / \Gamma(r)}{D_{j\delta}} + \int_{j\delta}^{\infty} \frac{y \delta \lambda^r y^{r-1} e^{-\lambda y} / \Gamma(r)}{D_{j\delta}} \right)^2 \\
 &\quad \text{where } D_{j\delta} = \frac{e^{\delta\lambda}(2 - \delta\lambda + \delta j\lambda) - \delta\lambda j - 2}{\lambda e^{\delta j\lambda}}.
 \end{aligned}$$

Unlike the exponential, the variances of sojourn times using this pdf for  $r = 2$ , are not independent of  $j$ . The formulas in Table 4.7 and Table 4.8 can be used to calculate  $E(Y_j^{2*}) - [E(Y_j^*)]^2$ , the variances for  $Y_1^*, Y_2^*, Y_3^*$  and  $Y_4^*$ .

The  $\lim_{\delta \rightarrow 0} Var(Y_j^*) = \frac{\mu^2}{2}$  when  $\lambda = \frac{2}{\mu}$ . To calculate the variances, let  $\lambda = \frac{2}{\mu}$  in the formula tables 4.6 and 4.7.

This is the same value that was used when the mean sojourn times

**Table 4.8.** Gamma Formulas for the Expected Value Squared when  $r = 2$

$(E(Y_j^*))^2$	Equations
$(E(Y_1^*))^2$	$\left(\frac{6 - 6e^{\delta\lambda} + 4\delta\lambda + \delta^2\lambda^2}{\lambda(2 - 2e^{\delta\lambda} + \delta\lambda)}\right)^2$
$(E(Y_2^*))^2$	$\left(\frac{-2[3 + 2\delta\lambda(2 + \delta\lambda)] + e^{\delta\lambda}[6 + \delta\lambda(4 + \delta\lambda)]}{\lambda[-2(1 + \delta\lambda) + e^{\delta\lambda}(2 + \delta\lambda)]}\right)^2$
$(E(Y_3^*))^2$	$\left(\frac{-6 - 12\delta\lambda - 9\delta^2\lambda^2 + 2e^{\delta\lambda}(3 + 4\delta\lambda + 2\delta^2\lambda^2)}{\lambda[-2 - 3\delta\lambda + 2e^{\delta\lambda}(1 + \delta\lambda)]}\right)^2$
$(E(Y_4^*))^2$	$\left(\frac{-6 - 16\delta\lambda - 16\delta^2\lambda^2 + 3e^{\delta\lambda}(2 + 4\delta\lambda + 3\delta^2\lambda^2)}{\lambda(-2 - 4\delta\lambda + e^{\delta\lambda}(2 + 3\delta\lambda))}\right)^2$

were calculated. The variances using this substitution are found in Table 4.6.

The same equation for  $E(Y^*)$  can now be applied to  $Var(Y^*)$ ; that is, [4, p. 7]

$$\begin{aligned}
 Var(Y^*) &= \sum_{j=0}^{\infty} Var(Y_j^*) \cdot P\{\text{missed on } (j-1) \text{ previous screens,} \\
 &\quad \text{detected on the } j^{th} \text{ screen}\} \\
 &= \beta Var(Y_1^*) + (1-\beta)\beta Var(Y_2^*) + (1-\beta)^2\beta Var(Y_3^*) \\
 &\quad + (1-\beta)^3\beta Var(Y_4^*)
 \end{aligned}$$

where  $\beta$  is the test sensitivity [4, p. 7].

To obtain a comparison, that is a ratio between the sampled standard deviations,  $Y^*$ , and the overall standard deviations of  $Y$ , each value in Table 4.9 is divided by  $\frac{\mu}{\sqrt{2}}$  to obtain Table 4.12.

The relative differences in the sampled standard deviations and the standard deviations of preclinical duration are very small since each ratio is close to one. This completes the analysis for the gamma distribution for  $r = 2$ .

**Table 4.9.** Gamma Distribution Variances for  $r = 2$ ,  $j = 1, 2, 3, 4$

		Values of $\delta$				
$Var(Y_j^*)$	$\mu$	1	2	3	4	5
$Var(Y_1^*)$	0.5	0.15	0.18	0.19	0.19	0.19
	1.0	0.49	0.59	0.68	0.73	0.74
	2.0	1.92	1.95	2.12	2.34	2.56
	5.0	12.31	12.06	11.93	11.97	12.17
	10.0	49.75	49.24	48.70	48.23	47.89
$Var(Y_2^*)$	0.5	0.13	0.15	0.14	0.14	0.14
	1.0	0.43	0.52	0.57	0.58	0.57
	2.0	1.71	1.72	1.89	2.07	2.21
	5.0	11.66	10.91	10.57	10.56	10.78
	10.0	48.68	46.65	44.91	43.62	42.76
$Var(Y_3^*)$	0.5	0.12	0.14	0.13	0.13	0.13
	1.0	0.40	0.49	0.53	0.54	0.54
	2.0	1.57	1.59	1.77	1.94	2.07
	5.0	11.00	10.03	9.68	9.71	9.97
	10.0	47.26	44.00	41.65	40.10	39.18
$Var(Y_4^*)$	0.5	0.12	0.13	0.13	0.13	0.13
	1.0	0.38	0.47	0.52	0.53	0.53
	2.0	1.47	1.52	1.70	1.88	2.00
	5.0	10.44	9.42	9.13	9.21	9.51
	10.0	45.81	41.78	39.22	37.69	36.85

**Table 4.10.** Gamma Distribution  $r = 2$ ,  $\beta = 0.95 Var(Y^*)$

		Values of $\delta$				
$\mu$		1	2	3	4	5
0.5		0.15	0.18	0.18	0.19	0.18
1.0		0.48	0.58	0.68	0.72	0.74
2.0		1.91	1.94	2.10	2.33	2.54
5.0		12.28	12.00	11.86	11.90	12.10
10.0		49.70	49.11	48.50	48.00	47.63

**Table 4.11.** Gamma Distribution  $r = 2$ , Standard Deviations of  $Y^*$

$\mu$	Values of $\delta$				
	1	2	3	4	5
0.5	0.38	0.42	0.43	0.43	0.43
1.0	0.70	0.76	0.82	0.85	0.86
2.0	1.38	1.39	1.45	1.53	1.59
5.0	3.50	3.46	3.44	3.45	3.48
10.0	7.05	7.01	6.96	6.93	6.90

**Table 4.12.** Gamma Distribution  $r = 2$  Ratios of Standard Deviations

(Standard Deviations of  $Y^*$  divided by  $\frac{\mu}{\sqrt{2}}$ )

$\mu$	Values of $\delta$				
	1	2	3	4	5
0.5	1.08	1.20	1.22	1.22	1.22
1.0	0.98	1.08	1.16	1.20	1.21
2.0	0.98	0.98	1.03	1.08	1.13
5.0	0.99	0.98	0.97	0.98	0.98
10.0	1.00	0.99	0.98	0.98	0.98

**Table 4.13.** Gamma Distribution  $r = 2$ ; Relative Error ( %) in Standard Deviations of  $Y^*$

Values of $\delta$					
$\mu$	1	2	3	4	5
0.5	8	20	22	22	22
1.0	-2	8	16	20	21
2.0	-2	-2	3	8	13
5.0	-1	-2	-3	-2	-2
10.0	-0	-1	-2	-2	-2

A summary of the findings is located in Table 4.14.

## 4.2 The Gamma Distribution for Sojourn Times for $r = 4$

The summary for the comparison of the screen-detected sojourn times with the over-all sojourn times using gamma distribution for  $r = 4$  is given in 4.15. A comparison of the three summary tables for  $r = 1, r = 2,$  and  $r = 4$  is given in the conclusion.

**Table 4.14.** Summary of  $Y^*$  when  $r = 2$ : The Means, Relative Increase, Variances, Standard Deviations, and Relative Errors

**The Preclinical Durations of Cases Eligible for Detection by Screening**

$E(Y^*)$	Values of $\delta$					
	$\mu$	1	2	3	4	5
Sampled Means	0.5	0.75	0.84	0.90	0.95	1.00
	1.0	1.26	1.49	1.61	1.68	1.74
	2.0	2.22	2.53	2.79	2.99	3.13
	5.0	5.13	5.39	5.70	6.01	6.32
	10.0	10.07	10.25	10.49	10.78	11.08
Relative Increase	$\mu$	1	2	3	4	5
$(\frac{E(Y^*)}{\mu} - 1)100$	0.5	49	68	79	90	100
	1.0	26	49	61	68	74
	2.0	11	26	40	49	56
	5.0	3	8	14	20	26
	10.0	1	3	5	8	11
$Var(Y^*)$	$\mu$	1	2	3	4	5
Variances of sampled sojourn times $E(Y^{*2}) - (E(Y^*))^2$	0.5	0.15	0.18	0.18	0.19	0.18
	1.0	0.48	0.58	0.68	0.72	0.74
	2.0	1.91	1.94	2.10	2.33	2.54
	5.0	12.28	12.00	11.86	11.90	12.10
	10.0	49.70	49.11	48.50	48.00	47.63
$sd(Y^*)$	$\mu$	1	2	3	4	5
Standard Deviations of sampled sojourn times	0.5	0.38	0.42	0.43	0.43	0.43
	1.0	0.70	0.76	0.82	0.85	0.86
	2.0	1.38	1.39	1.45	1.53	1.59
	5.0	3.50	3.46	3.44	3.45	3.48
	10.0	7.05	7.01	6.96	6.93	6.90
Relative Error	$\mu$	1	2	3	4	5
$(sd(Y^*)/\frac{\mu}{\sqrt{2}} - 1)100$	0.5	8	20	22	22	22
	1.0	-2	8	16	20	21
	2.0	-2	-2	3	8	13
	5.0	-1	-2	-3	-2	-2
	10.0	-0	-1	-2	-2	-2

**Table 4.15.** Summary of  $Y^*$  when  $r = 4$ : The Means, Relative Increase, Variances, Standard Deviations, and Relative Errors

**The Preclinical Durations of Cases Eligible for Detection by Screening**

		Values of $\delta$				
$E(Y^*)$	$\mu$	1	2	3	4	5
Sampled Means	0.5	0.65	0.71	0.77	0.82	0.87
	1.0	1.15	1.31	1.37	1.43	1.48
	2.0	2.08	2.30	2.49	2.61	2.69
	5.0	5.02	5.11	5.29	5.51	5.74
	10.0	10.00	10.03	10.11	10.23	10.39
Relative Increase	$\mu$	1	2	3	4	5
$(\frac{E(Y^*)}{\mu} - 1)100$	0.5	31	43	53	64	74
	1.0	15	31	37	43	48
	2.0	4	15	24	31	34
	5.0	0	2	6	10	15
	10.0	0	0	1	2	4
$Var(Y^*)$	$\mu$	1	2	3	4	5
Variances of sampled sojourn times $E(Y^{*2}) - (E(Y^*))^2$	0.5	0.14	0.21	.028	0.36	0.43
	1.0	0.26	0.34	0.40	0.44	0.48
	2.0	0.96	0.96	1.06	1.18	1.26
	5.0	6.21	6.04	5.86	5.77	5.82
	10.0	24.98	24.83	24.53	24.14	23.74
$sd(Y^*)$	$\mu$	1	2	3	4	5
Standard Deviations of sampled sojourn times	0.5	0.26	0.28	0.28	0.28	0.27
	1.0	0.48	0.53	0.55	0.55	0.55
	2.0	0.97	0.96	1.01	1.06	1.09
	5.0	2.49	2.46	2.42	2.40	2.40
	10.0	5.00	4.98	4.95	4.91	4.87
Relative Error	$\mu$	1	2	3	4	5
$(sd(Y^*)/\frac{\mu}{2} - 1)100$	0.5	6	10	10	10	1
	1.0	-4	6	10	10	10
	2.0	-3	-4	1	6	9
	5.0	0	-2	-3	-4	-4
	10.0	0	0	-1	-2	-3

## 5. Conclusion

The effect of length-biased sampling on the survival among screen-detected cases is not trivial. The mean increase in survival time among cases detected by screening can be as large as 79% when the screening interval ( $\delta$ ) is twice the mean sojourn time ( $\mu$ ) in the general population. This increase is only 47% when the screening interval is equal to the mean sojourn time, and decreases to 0 as  $\delta \rightarrow 0$  (continuous screening). The standard deviations are less affected, but some increases, particularly with the exponential distribution, are observable, on the order of 13% when  $\delta = 2\mu$  and 4% when  $\delta = \mu$ .

When the sojourn time distribution is gamma with values of  $r \geq 2$  the effect of length-biased sampling is reduced. Studies suggest that  $r$  is likely to be less than 2 (Zelen and Feinlieb), so these results should be taken into consideration when evaluating the potential benefits of a screening program. Length-biased sampling can cause screening to appear more beneficial than it actually is and therefore can result in over-optimistic conclusions concerning the benefit of screening programs.

## NOTATION INDEX

**coefficient of variation:** measures the amount of variability relative to the

value of the mean:  $CV = \sigma/\mu$

$\delta$ : Delta, the length of time between screenings

$\mathbf{E}(\mathbf{Y}_j)$ : the mean preclinical duration among cases that started before  $\delta$  but

no later than  $j\delta$  the time of the  $j^{th}$  screening.

$\mathbf{E}(\mathbf{Y}^*)$ : mean length of preclinical durations of the sampled cases

$\mathbf{f}_\mathbf{Y}(\mathbf{y})$ : the probability density function (pdf) of all preclinical durations.

$\mathbf{F}_\mathbf{Y}(\mathbf{y})$ : the cumulative distribution function,(cdf), for the all preclinical durations

**HIP:** Health Insurance Plan of Greater New York; a prepaid comprehensive group medical plan insuring approximately 700,000 city, state and federal government employees and their families.

**incident cases:** those cases of disease occurring twelve or more months after the previous CNBSS screening examination

**interval cases:** those cases of disease occurring less than twelve months after a negative screening exam.

**j:** an integer representing the jth screening for a disease

**lead time:** forward recurrence time; length of time by which the diagnosis is advanced over clinical detection by virtue of the screening procedure

**length bias:** biased sample caused by periodic sampling; observations of longer duration are more likely to be detected than those of short duration.

**power of an hypothesis test:** The ability of a test to reject the null hypothesis when the alternative hypothesis is true is called the power of the test.

**S<sub>c</sub>:** the clinical phase of disease characterized by overt signs or symptoms

**sensitivity:** the proportion of individuals designated positive by the screening test among all individuals who have the disease

**significance-level  $\alpha$ :** The probability of rejecting the null hypothesis when the null hypothesis is true.

**S<sub>o</sub>:** disease-free state of a disease characterized by being either free of the disease or having disease characteristics of that are undetectable by a screening test

**sojourn times:** preclinical durations

$S_p$ : preclinical state of disease; Symptoms are asymptomatic.

**specificity:** the proportion designated negative by the test among all those  
who do not have the disease

$Y$ : the random variable that denotes the preclinical duration of a case of  
disease in the general population

$Y^*$ : the random variable that denotes the the preclinical duration for a  
screen-detected case

## **A. APPENDIX**

The values found in the following tables, A.1 - A.4, were used to calculate the Summary table 3.9.

The values found in tables A.5 - A.10 were used to calculate the Summary Table 4.15.

**Table A.1.** Exponential Distribution Ratios  $E(Y_j^*)/\mu$ Mean Sojourn Time through the  $j^{\text{th}}$  Screening /  $\mu$ 

$E(Y_j)/\mu$	Values of $\delta$				
	1	2	3	4	5
$E(Y_1)/0.5$	1.69	1.93	1.99	2.00	2.00
$E(Y_1)/1.0$	1.42	1.69	1.84	1.93	1.97
$E(Y_1)/2.0$	1.23	1.42	1.57	1.69	1.78
$E(Y_1)/5.0$	1.10	1.19	1.27	1.35	1.42
$E(Y_1)/10.0$	1.05	1.10	1.14	1.19	1.23
$E(Y_2)/0.5$	3.69	5.93	7.99	10.00	12.00
$E(Y_2)/1.0$	2.42	3.69	4.84	5.93	9.97
$E(Y_2)/2.0$	1.73	2.42	3.07	3.69	4.28
$E(Y_2)/5.0$	1.30	1.59	1.87	2.15	2.42
$E(Y_2)/10.0$	1.15	1.30	1.44	1.59	1.73
$E(Y_3)/0.5$	5.69	9.93	13.99	18.00	22.00
$E(Y_3)/1.0$	3.42	5.69	7.84	9.93	11.97
$E(Y_3)/2.0$	2.23	3.42	4.57	5.69	6.78
$E(Y_3)/5.0$	1.50	1.99	2.47	2.95	3.42
$E(Y_3)/10.0$	1.25	1.50	1.74	1.99	2.23
$E(Y_4)/0.5$	7.69	13.93	19.99	26.00	32.00
$E(Y_4)/1.0$	4.42	7.69	10.84	13.93	16.97
$E(Y_4)/2.0$	2.73	4.42	6.07	7.69	9.28
$E(Y_4)/5.0$	1.70	2.39	3.07	3.75	4.42
$E(Y_4)/10.0$	1.35	1.70	2.04	2.39	2.73

**Table A.2.** Exponential Distribution Variances:  $Var(Y_j^*)$ 

$\mu$	Values of $\delta$				
	1	2	3	4	5
0.5	0.32	0.42	0.48	0.49	0.50
1.0	1.08	1.28	1.50	1.70	1.83
2.0	4.08	4.32	4.67	5.10	5.56
5.0	25.08	25.33	25.74	26.29	26.98
10.0	100.08	100.33	100.75	101.32	102.06

**Table A.3.** Exponential Distribution Standard Deviations for  $Y_j^*$

Values of $\delta$					
$\mu$	1	2	3	4	5
0.5	0.56	0.65	0.69	0.70	0.71
1.0	1.04	1.13	1.23	1.30	1.35
2.0	2.02	2.08	2.16	2.26	2.36
5.0	5.01	5.03	5.07	5.13	5.19
10.0	10.00	10.02	10.04	10.07	10.10

**Table A.4.** Exponential Distribution (Standard Deviations ( $Y_j^*$ ))/ $\mu$

Values of $\delta$					
$\mu$	1	2	3	4	5
0.5	1.13	1.30	1.38	1.41	1.411
1.0	1.04	1.13	1.23	1.30	1.35
2.0	1.01	1.04	1.08	1.13	1.18
5.0	1.00	1.01	1.01	1.03	1.04
10.0	1.00	1.00	1.00	1.01	1.01

**Table A.5.** Gamma Distribution Means  $r = 4$  for  $j = 1, 2, 3, 4$

		Values of $\delta$				
$E(Y_j^*)$	$\mu$	1	2	3	4	5
$E(Y_1^*)$	0.5	0.61	0.62	0.63	0.63	0.63
	1.0	1.12	1.23	1.25	1.25	1.25
	2.0	2.05	2.23	2.38	2.46	2.49
	5.0	5.01	5.08	5.21	5.39	5.58
	10.0	10.00	10.02	10.07	10.15	10.27
$E(Y_2^*)$	0.5	1.34	2.30	3.28	4.27	5.27
	1.0	1.73	2.67	3.62	4.59	5.57
	2.0	2.49	3.47	4.41	5.34	6.28
	5.0	5.14	5.79	6.70	7.69	8.66
	10.0	10.03	10.29	10.83	11.58	12.46
$E(Y_3^*)$	0.5	2.29	4.27	6.27	8.26	10.26
	1.0	2.63	4.59	6.56	8.55	10.62
	2.0	3.25	5.25	7.22	9.18	11.15
	5.0	5.53	7.17	9.13	11.14	13.13
	10.0	10.16	11.06	12.55	14.34	16.27
$E(Y_4^*)$	0.5	3.28	6.27	9.26	12.26	15.26
	1.0	3.58	6.56	9.54	12.53	15.53
	2.0	4.14	7.16	10.15	13.12	16.10
	5.0	6.12	8.85	11.86	14.90	17.91
	10.0	10.42	12.23	14.81	17.69	20.69

**Table A.6.** Gamma Distribution Ratio:  $\frac{E(Y_i^*)}{\mu}$  when  $r = 4$

		Values of $\delta$				
$E(Y_i^*)/\mu$	$\mu$	1	2	3	4	5
$E(Y_1^*)/\mu$	0.5	1.23	1.25	1.25	1.25	1.25
	1.0	1.12	1.23	1.25	1.25	1.25
	2.0	1.03	1.12	1.19	1.23	1.24
	5.0	1.00	1.02	1.04	1.08	1.12
	10.0	1.00	1.00	1.01	1.02	1.03
$E(Y_2^*)/\mu$	0.5	2.67	4.59	6.56	8.55	10.54
	1.0	1.73	2.67	3.62	4.59	5.57
	2.0	1.25	1.73	2.21	2.67	3.14
	5.0	1.03	1.16	1.34	1.54	1.73
	10.0	1.00	1.03	1.08	1.16	1.25
$E(Y_3^*)/\mu$	0.5	4.59	8.55	12.53	16.52	20.52
	1.0	2.63	4.59	6.56	8.55	10.62
	2.0	1.63	2.63	3.61	4.59	5.57
	5.0	1.11	1.43	1.83	2.23	2.63
	10.0	1.02	1.11	1.25	1.43	1.63
$E(Y_4^*)/\mu$	0.5	6.56	12.53	18.52	24.52	30.52
	1.0	3.58	6.56	9.54	12.53	15.53
	2.0	2.07	3.58	5.07	6.56	8.05
	5.0	1.22	1.77	2.37	2.98	3.58
	10.0	1.04	1.22	1.48	1.77	2.07

**Table A.7.** Gamma Distribution  $r = 4$   $\frac{E(Y^*)}{\mu}$

		Values of $\delta$				
$\mu$		1	2	3	4	5
0.5		1.31	1.43	1.53	1.64	1.74
1.0		1.15	1.31	1.37	1.43	1.48
2.0		1.04	1.15	1.24	1.31	1.34
5.0		1.00	1.02	1.06	1.10	1.15
10.0		1.00	1.00	1.01	1.02	1.04

**Table A.8.**  $r = 4$  Variances for  $E(Y_j^*)$  for  $j = 1, 2, 3, 4$

$Var(Y_j^*)$	Values of $\delta$					
	$\mu$	1	2	3	4	5
$Var(Y_1^*)$	0.5	0.07	0.08	0.08	0.08	0.08
	1.0	0.23	0.28	0.31	0.31	0.31
	2.0	0.96	0.93	1.02	1.13	1.20
	5.0	6.22	6.08	5.90	5.81	5.84
	10.0	24.99	24.90	24.67	24.33	23.95
$Var(Y_2^*)$	0.5	0.05	0.04	0.04	0.04	0.04
	1.0	0.19	0.20	0.18	0.17	0.16
	2.0	0.79	0.76	0.81	0.80	0.77
	5.0	5.92	5.17	4.75	4.68	4.78
	10.0	24.80	23.70	22.11	20.67	19.63
$Var(Y_3^*)$	0.5	0.04	0.04	0.04	0.03	0.03
	1.0	0.16	0.17	0.16	0.15	0.14
	2.0	0.64	0.64	0.67	0.66	0.64
	5.0	5.35	4.26	3.90	3.89	4.01
	10.0	24.18	21.38	18.78	17.05	16.07
$Var(Y_4^*)$	0.5	556.04	1153.12	1752.08	2351.56	2951.24
	1.0	258.19	556.04	854.14	1153.12	1452.50
	2.0	106.86	258.19	407.40	556.04	704.95
	5.0	22.33	76.93	137.19	197.90	258.19
	10.0	4.21	22.33	48.13	76.93	106.86

**Table A.9.**  $r = 4$  Ratio of Sampled Standard Deviations/ $\frac{\mu}{2}$

$E(Y_j^*)$	$\mu$	Values of $\delta$				
		1	2	3	4	5
$E(Y_1^*)$	0.5	1.06	1.12	1.12	1.12	1.12
	1.0	0.97	1.06	1.11	1.12	1.12
	2.0	0.98	0.97	1.01	1.06	1.10
	5.0	1.00	0.99	0.97	0.96	0.97
	10.0	1.00	1.00	0.99	0.99	0.98
$E(Y_2^*)$	0.5	0.90	0.82	0.79	0.77	0.76
	1.0	0.87	0.90	0.86	0.82	0.80
	2.0	0.89	0.87	0.90	0.90	0.88
	5.0	0.97	0.91	0.87	0.87	0.87
	10.0	1.00	0.97	0.94	0.91	0.89
$E(Y_3^*)$	0.5	0.82	0.77	0.75	0.74	0.73
	1.0	0.80	0.82	0.79	0.77	0.76
	2.0	0.80	0.80	0.82	0.82	0.80
	5.0	0.92	0.83	0.79	0.79	0.80
	10.0	0.98	0.92	0.87	0.83	0.80
$E(Y_4^*)$	0.5	0.78	0.75	0.74	0.73	0.72
	1.0	0.76	0.78	0.76	0.75	0.74
	2.0	0.75	0.76	0.78	0.78	0.77
	5.0	0.87	0.77	0.74	0.75	0.76
	10.0	0.96	0.87	0.81	0.77	0.75

**Table A.10.** Gamma Distribution  $r = 4$  Ratio Standard Deviations of  $Y^*/\sigma$

$\mu$	Values of $\delta$				
	1	2	3	4	5
0.5	1.06	1.10	1.10	1.10	1.10
1.0	0.96	1.06	1.10	1.10	1.10
2.0	0.97	0.96	1.01	1.06	1.09
5.0	1.00	0.98	0.97	0.96	0.96
10.0	1.00	1.00	0.99	0.98	0.97

## REFERENCES

- [1] Jay Devore. *Probability and Statistics for Engineering and the Sciences*. Brooks/Cole Publishing Company, Pacific Grove, California, 1991.
- [2] M Feinlieb and M. Zelen. On the theory of screening for chronic diseases. *Biometrika*, 56(3):601–613, 1969.
- [3] M. C. Jones. Kernal density estimation for length biased data. *Biometrika*, 78(3):511–519, 1991.
- [4] Karen Kafadar and Philip C. Prorok. Effect of length biased sampled sojourn times on the survival distribution from screen-detected diseases. Work performed for the National Cancer Institue, Biometry Branch, Division of Cancer Prevention and Control. This work is still in progress, May 1999.
- [5] Anthony B. Miller, Teresa To, Cornelia J. Baines, and Claus Wall. Canadian national breast screening study-2: 13-year results of a randomized trial in women aged 50-59 years. *Journal of the National Cancer Institute*, 92(18):1490–1499, 2000.
- [6] Alan Morrison. Sequential pathogenic components. *American Journal of Epidemiology*, 109(6):709–718, 1979. A formula for the mortality rate is derived. The components of increased survival time can be separated that is the benefit, time, lead time and length bias can be separated.
- [7] Joseph Petrucci, Balgobin Nandram, and Minghui Chen. *Applied Statistics for Engineers and Scientists*. Prentice Hall, Upper Saddle River, New Jersey, 1999.
- [8] Philip C. Prorok and Robert J. Conner. Screening for the early detection of cancer. *BioStatistics Cancer Investigation*, 4(3):225–238, 1986.

- [9] Sam Shapiro, Wanda Vent, Phillip Strax, and Loius Venet. *Periodic Screening for Breast Cancer the Health Insurance Plan Project and Its Sequelae*. John Hopkins University Press, Baltimore, Maryland, 1988.