Abstract

Ghassan Hamra

Dose Reconstruction for an Occupational Cohort at the Savannah River Nuclear Facility:
Evaluation of a hybrid method
(Under the direction of Dr. David Richardson)

The Savannah River Site (SRS) is the sole nuclear facility in the US that has the expressed goal of producing tritium, a radioactive isotope of hydrogen that may be related to certain adverse health outcomes. Although a great deal is known about personal exposure to radiation at SRS via dosimetry records, a number of employees do not have recorded tritium dose measures, except as recorded in aggregate with external dose estimates. The purpose of this paper is to derive tritium dose estimates for SRS employees through development of a job-exposure matrix. Our method is unique in that along with qualitative information on job, area, and time of employment, our model utilizes quantitative estimates of annual whole body dose in order to estimate doses from tritium intakes. From our original cohort of 18,883 workers employed from 1951-1999, we calculated estimates for those individuals with missing tritium dose records who were employed from 1953-1978 with a positive whole body dose. Using information from 75,253 known tritium dose measures fitting our criteria above, we estimated tritium doses for 43,590 employment-years. The collective estimated tritium dose is 4,319 mSv compared to the collective known tritium dose of 17,382 mSv.

Acknowledgements

I would like to recognize a number of individuals who have been involved with this project. First, I would like to thank Dr. David Richardson for allowing me to pursue this research project, and for constant guidance throughout. I would also like to thank Dr. Leena Nylander-French for her recognition of the natural connection between exposure assessment and epidemiology, and for allowing me to pursue a research project, which bridges the gap between the two. I am grateful to Dr. Mike Flynn for agreeing to be a member of my research committee in the latter stages of the project, and for his advice on later drafts. Also, I would like to thank Dr. Dana Loomis for his guidance in the early stages of this project. Finally, I would like to thank Dr. Steve Wing, early discussions related to this project and our method of analysis, before any analysis even took place. This project was supported by the National Institute for Occupational Safety and Health, grant number T42 OH008673.

TABLE OF CONTENTS

	Page
LIST OF TABLES.....	vi
LIST OF FIGURES.....	viii
LIST OF ABBREVIATIONS.....	ix

Chapter

1. INTRODUCTION.....	1
1.1 Background.....	1
1.2 Savannah River Site.....	1
1.3 Tritium.....	2
1.4 SRS and Tritium.....	3
1.5 The Goal.....	3
2. Methods.....	5
2.1 Cohort.....	5
2.2 Occupation.....	5
2.3 Health Physics Area.....	6
2.4 Tritium dosimetry program at SRS.....	6
2.5 External dosimetry recording practice at SRS.....	7
2.6 Computerized dosimetry information.....	9
2.7 Employment-years lacking computerized dose estimates.....	10
2.8 Estimation of tritium dose.....	11
2.9 Model form and evaluation.....	12

3. Results.....	15
3.1 Cohort Data.....	15
3.2 Model Evaluation.....	16
3.3 Estimating tritium values for employment-years lacking computerized tritium records.....	19
4. Discussion.....	22
References.....	38
Appendix: SAS Code.....	40

LIST OF TABLES

Table

	Page
1. Description of employee records for entire SRS occupational cohort from 1951-1999.....	26
2. Distribution of known annual tritium doses (mSv) for employment-years between 1954 and 1978, Savannah River Site.....	27
3. Mean, standard deviation, and sum of known tritium dose by year (1954-1978). Also, ratio of annual mean tritium dose to estimated annual whole body dose (ANDS). Savannah River Site.....	28
4. Number of employment-years (N), the % of which are missing tritium dose, the mean and standard deviation (Std Dev) of tritium dose, mean of annual estimated whole body dose (ANDS), and the ratio of tritium dose to ANDS by area and occupation. (n=74,610) Savannah River Site 1954-1978.....	29
5. Mean and 95% of the distribution of estimated tritium dose for employees with known tritium dose by year (n=74,610), SRS 1954-1978.....	30
6. Mean and 95% of the distribution of estimated tritium dose for employees with known tritium dose by health physics area and occupation (n=74,610), SRS 1954-1978.....	31
7. Distribution of estimated tritium dose for employees with no known tritium dose (mSv), SRS 1954-1978.....	32

8. Mean and 95% of the distribution of estimated tritium dose for employees
without a known tritium dose records by year (n=43,381), SRS 1954-1978....33
9. Mean and 95% of the distribution of estimated tritium dose by area and
occupation for employees without a known tritium dose (n=43,381), SRS
1954-1978.....34

LIST OF FIGURES

Figure

Page

1. Plot of estimated vs observed tritium dose (1954-1978, SRS).....35
2. Box-plots of estimated vs observed tritium mean by group of observed values
(1954-1978, SRS).....36
3. Box-plots of difference between predicted and known tritium values (1954-1978,
SRS).....37

ABBREVIATIONS

ANDS	Annual Estimated Whole Body Dose
HPAREH	Health Protection Annual Radiation Exposure History
HTO	Tritiated Water
JEM	Job Exposure Matrix
MDL	Minimum Detection Limit
NIOSH	National Institute of Occupational Safety and Health
NTA	Nuclear Track Emulsion
ORAU	Oak Ridge Associated Universities
ORNL	Oak Ridge National Laboratory
PCB	Polychlorinated Biphenyl
SRS	Savannah River Site
TLD	Thermoluminescent Dosimeter

Chapter 1

INTRODUCTION

1.1 Background

Often in occupational and environmental epidemiological studies, quantitative information is available for some members of a cohort on an exposure metric that reflects aggregate exposure to several related hazards. The investigator may be interested in deriving estimates of exposure to one specific agent based on known levels of exposure to that agent.

In this paper we focus on radiological constituents of whole body dose; however, similar issues might arise in studies in which total PCBs were quantified (and the researcher is interested in exposure to a specific congener), or in studies in which total particulate matter was assessed (and the researcher is interested in exposures to particulate of specific dimensions).

1.2 Savannah River Site

For a subpopulation of our study cohort we have computerized records of each component of whole body dose: estimated deposition of beta/photon rays and internal deposition of tritium. Using the available data, we developed an approach to estimation of the tritium component of whole body dose for each dose record. We applied this method in the context of a cohort study of workers employed at the Savannah River Site

(SRS), a 315 square mile nuclear fuels facility located in Aiken, SC. Originally operated by E.I. duPont de Nemours & Company, SRS has produced nuclear fuels, mainly tritium and plutonium-239, for over 50 years. In 1989, management of SRS was turned over to Westinghouse Savannah River Company, who has managed the plant since. Since SRS is solely a nuclear fuels facility, a good deal of radiological hazard protection is integrated into everyday operations. Although this may help reduce occupational exposures, it does not eliminate them. About 85% of the total occupational dose is characterized as external exposures, while the remaining 15% are attributable to internally deposited radionuclides (Cragle et al. 1988).

1.3 Tritium

One of these radionuclides is tritium, an isotope of hydrogen that emits beta radiation as it decays into helium (Hill et al. 1993). Since it acts like hydrogen, tritium gas is capable of binding to oxygen molecules to form tritiated water (HTO). HTO may enter the body via inhalation, ingestion, or absorption through the skin (Okada et al. 1993). Once absorbed, HTO will readily diffuse through cellular membranes, uniformly integrating itself into the water present in the human body (Hill et al. 1993). While tritium has a physical half-life of 12.3 years, ingested HTO has a biological half-life of about 10 days (Elliot 2005). In that time, tritium is capable of producing genetic mutations via beta radiation or the energy release associated with transformation from ^3H to ^3He (Hill et al. 1993). Animal tests have shown that acute exposure to HTO can lead to malformations and death (Okada et al. 1993); however, this tells us little about the effects of chronic low-level exposure. One study by Joksic and Spasojevic-Tisma (1998)

found that low-level exposure to tritium caused chromosomal damage in human lymphocytes, but direct estimates of human cancer risk following tritium exposure are not available (Straume 1993).

1.4 SRS and Tritium

SRS is the only nuclear fuels facility in the US that produces tritium. Although tritium is a by-product of processes at most other nuclear facilities, the fact that SRS has the explicit task of producing tritium necessitates an examination of exposure to tritium among SRS employees. Tritium dose was measured at SRS via biological monitoring. Since the late 1970s dose measures have been maintained in an electronic format that facilitates their use for research purposes. However, dose measures for earlier time periods have only been computerized in summation with the other components of a worker's whole body dose (i.e., summed together with penetrating dose estimates from external irradiation).

1.5 The Goal

The goal of this work is to develop a predictive model of the tritium dose component of annual estimated whole body dose (ANDS). This model will be applied to workers who have computerized whole body dose estimates in order to derive tritium dose estimates. The project is unusual in that often in occupational settings a researcher has very little information with which to derive individual quantitative exposure estimates. In this case, we have annual tritium dose estimates for a large proportion of the workers in our cohort. In addition, for those with missing tritium dose estimates (i.e., the

tritium component of whole body dose was not computerized) we do have computerized estimates of ANDS. Therefore our project involves estimation of tritium dose using information on job characteristics and recorded whole body dose estimate.

Chapter 2

METHODS

2.1 Cohort

Our cohort includes 21,204 individuals hired by DuPont between 1950 and 1986. We excluded from analyses those workers without a known date of birth ($n=57$), date of first hire ($n=184$), or gender ($n=10$). In addition, we excluded individuals who were employed less than 90 days ($n=1355$) since they may differ compared to long-term employees with respect to mortality risk and cumulative dose estimates. Finally, SRS workers previously employed at another Department of Energy facility ($n=715$) were also excluded since we do not have information on occupational radiation exposure that occurred outside of the individuals employment at SRS. This leaves 18,883 SRS workers.

2.2 Occupation

A file containing work history information was created from du Pont payroll records that contain information about dates of employment and job title changes. Job titles were standardized and coded to 34 major occupational groups. Based on this information, a file was created that describes the number of days that a worker was employed during each calendar year (1951-1999). If a worker held more than one job in a calendar year, for simplicity, a single occupation for that worker was assigned based on the longest held occupation in that year.

In this paper, we used the term "employment-year" to describe the unit of observation contributed by a person each year he/she was employed at SRS, regardless of number of days employed. A worker who has computerized annual dosimetry information for his/her entire employment period would have one annual dosimetry record for each employment-year.

2.3 Health Physics Area

The term "health physics area" represents a system defined by health physicists for classifying workers at SRS based on location and similarity of procedure. In each "area," occupational exposure to radiation among employees was under the supervision of radiation monitors, or health physics staff. Health physics area may represent a single location (such as an administrative building) or a number of work locations, which are physically separated but take part in similar processes (such as "100-Reactors," which consists of five reactors at SRS, some of which are miles apart). Information on health physics area was ascertained from quarterly dosimetry logbooks for the years 1958-1989. If a worker was missing information on health physics area for a given employment-year, but had a known health physics area for an adjacent time period during which they were employed in the same job, then, for the purposes of exposure imputations, we assigned that health physics area to the employment-year.

2.4 Tritium dosimetry program at SRS

Tritium dose records at SRS represent the annual sum of internally deposited tritium measured via *in vitro* bioassay analysis (e.g. urinalysis). Calcium metal is added

to HTO in urine samples and the evolved hydrogen from this was passed through an ionization chamber. This practice was standard from the start of operations until 1958 and had a minimum detectable activity (MDA) level of 1 $\mu\text{Ci/l}$. Beginning in 1958, urinalysis was performed in a liquid scintillation counter (LSC). The MDA lowered over the time, and is currently 0.1 $\mu\text{Ci/l}$. Including urinalysis results, calculation of tritium equivalent dose (expressed in rem or sievert) took a number of factors into consideration, including: biological half-life, target tissue, default mass of body water, a quality factor for tritium, and mean energy of tritium beta particle (Taylor et al. 1995).

Tritium, along with externally estimated beta/photon rays, readily distributes across the body (via tritiated water); thus, annual tritium dose was recorded as a component of the estimated whole body dose and represents a fraction of the recorded annual whole body dose. Since tritium is measured via biological monitoring after it has been distributed in the body, we refer to tritium as a dose based on the ISEA definition of dose as "the amount of agent that enters a target after crossing an exposure surface" (Zartarian et al. 2005).

Modern radiation dose information is expressed in sievert (Sv), which represents the biological equivalent dose based on joules per kilogram multiplied by weighing factors for the exposed organism and radiological agent of interest. For the sake of consistency with contemporary nomenclature, dose estimates that were originally expressed in units of rem are discussed and reported in this article in units of sievert, where 1 Sv=100 rem.

2.5 External dosimetry recording practices at SRS

During the beginning of operations at SRS, dosimetry services were provided by Oak Ridge National Laboratory (ORNL). ORNL's dosimeter technology was developed at the Metallurgical laboratories at the University of Chicago. These dosimeters were two-element beta/photon dosimeters with an open window and 1 mm silver filter. Open windows measure "shallow" dose while the silver filter measures "deep" dose. ORNL processed filters for SRS, and dosimeter exchange took place on a weekly basis. In 1951, Personal Ionization Chambers (PICs) were used in addition to film dosimeters to measure exposure among SRS workers. Nuclear Track Emulsion (NTA) dosimeters were also issued to select SRS employees to measure neutron dose, and were exchanged on a weekly basis for processing by ORNL. On March 3rd 1952, SRS initiated an on-site dosimetry program. Dosimeters were the same two-element film dosimeters used by ORNL and were collected on a weekly basis. In March 1953, SRS began processing film using the ORNL film badge dosimeter. SRS began processing NTA films in August 1953. On 10/1/1957, SRS beta/photon dosimeter exchange practice was changed to occur on a biweekly basis. On 11/9/1959 SRS implemented use of a multi-element film dosimeter. This dosimeter allowed for individual analysis of beta, gamma, and x-ray exposures among personnel. In 1965, SRS implemented a four-week exchange program for beta/photon dosimeters, which was changed to a monthly exchange program in 1966, which continued for the remainder of our study period. On 4/1/1970, SRS thermoluminescent dosimeters (TLDs) replaced film as the means for recording beta/photon dose. The laboratory minimum detection limit (MDL) for this method was 0.15 mSv, as compared to 0.4 mSv for the previous method. On 7/1/1983, use of

commercial Panasonic beta/photon TLDs was implemented. This new dosimetry method, which was used through our time period of interest, reduced the MDL from 0.15 to 0.05 mSv (Taylor et al. 1995).

2.6 Computerized dosimetry information

In 1979 a computerized personnel dosimetry system, referred to as the Health Protection Annual Radiation Exposure History (HPAREH) system, was implemented at SRS. The HPAREH system was developed in order to produce a history file of annual radiation exposure data for all SRS employees who were actively employed in 1979. Historical dosimetry information was entered into the HPAREH system from hardcopy personnel folders and logbooks (1951-1964), magnetic tapes of logbooks (1965-1972), and HP Master File magnetic tapes (1973-1979). Since 1979 dosimetry information has been routinely entered into the HPAREH system. Recorded values include estimates of deep, shallow, neutron, and tritium doses. The HPAREH file includes some records for years in which workers were not monitored for external radiation exposure at SRS via a personal dosimeter; these records were entered into the HPAREH system in order to record information about offsite doses and internal doses from radionuclides other than tritium. If the only information for a monitoring year pertained to an estimate of offsite dose or an estimate of effective dose from an internal deposition (other than tritium) then the record was excluded from these analyses.

Dosimetry information for an additional 1,058 workers was identified and computerized during the course of an epidemiological cohort study of SRS workers conducted by Oak Ridge Associated Universities; we refer to this as the SRPABST file.

Recorded values include estimates of open (i.e., shallow), shielded (i.e., deep), neutron, and tritium doses.

2.7 Employment-years lacking computerized tritium dose estimates

Workers who terminated employment at SRS prior to 1979 are not included in the HPAREH system. An electronic file of annual radiation dose estimates for the period 1951-1979 was constructed for the purposes of epidemiological research conducted by the DuPont Corporation; we refer to this as the Fayerweather file. Recorded values in the Fayerweather file include estimates of shallow and deep doses; however, abstraction of annual tritium dose information was highly incomplete. If a non-zero tritium dose value is recorded in the Fayerweather file it tends to be a value that can be validated; however, relatively few such values are recorded.

As part of our current epidemiological cohort study, we abstracted and keypunched dosimetry information (shallow and deep dose estimates) from historical dosimetry logbooks for an additional 854 workers who were employed in the period 1964-1979; and, we identified 15,752 annual dosimetry records in the historical dosimetry logbooks that were not included in the HPAREH, Fayerweather, or SRPABST computerized files. We keypunched dosimetry information from historical logbooks for 5,686 of these employment-years. The recorded annual deep and shallow dose estimates were zero mSv for nearly all of the remaining 10,066 employment-years. These were dosimetry records for workers who terminated after 1/1/1979 and appear in the historical SRS logbooks but not in the HPAREH file. We assigned an estimated annual deep and shallow dose of zero mSv to these years. Lastly, a 'nearby' method was used to estimate

annual whole body dose for 13,812 employment-years for which dosimetry information was not available (Richardson et al. 2006).

2.8 Estimation of tritium dose

For the computerized annual dosimetry records for which only shallow and deep dose estimates are available, we wish to derive estimates of the tritium component of the annual whole body dose. The technique we have applied combines an industrial hygiene approach to development of a job-area-exposure matrix (JEM) with empirical methods for exposure prediction using regression modeling.

Typically, a JEM will utilize qualitative information about area of employment, occupation, and even time of employment in order to assign a level of exposure (or in our case, dose) based on expert knowledge of where and when an exposure was likely to occur (Hoar S. 1983, Kromhout et al. 1992, Loomis et al. 1994, Goldberg et al. 1993). This level of exposure may be dichotomous, with an employee assigned a "yes/no" to exposure, or ordinal, with exposure described as "low/medium/high" (McGuire et al. 1998). Our method of dose reconstruction differs in that we combine this qualitative information with quantitative data about estimated whole body dose in order to provide quantitative estimates of annual tritium dose.

Previous studies of workers at SRS focused on classifying dose level according to job and area description alone, and suggest that workers receive the highest tritium dose in one of three processes: Neutron irradiation of lithium-aluminum targets or heavy water (D_2O) and fission due to reprocessing of reactor fuels (Hickey & Cragle 1985). However, the combinations of job-area that may have lead to some tritium dose were not

necessarily consistent over time. For example, there were important changes over time in the type and number of reactors operating at SRS. In addition, employees in an occupation with little expected tritium dose but working in an area characterized by high levels of tritium related activities may not have a clear relationship of job to tritium dose, as may have been thought in previous studies. Thus, we have incorporated information on occupation, health physics area, and calendar year into consideration in developing our predictive model of tritium dose.

2.9 Model form and evaluation

We fit a linear regression model to describe the relationship between tritium dose and annual estimated whole body dose (ANDS). This linear regression model was stratified by occupational group, health physics area, and calendar year thereby allowing a different relationship between tritium dose and annual estimated whole body dose (i.e., the fraction of ANDS due to tritium intake) within each stratum defined by these factors.

In addition, we defined strata by categories of annual estimated whole body dose. These strata allow for potential differences in the relationship between tritium and ANDS within subgroups defined by occupation, area, and calendar year. A major concern when developing a JEM is that job titles may not provide substantive distinctions between tasks at a facility, since job titles and area codes may represent information used for administrative tasks, rather than research purposes (Loomis et al. 1994). Further dividing groups of workers who share similar occupational titles and areas into subgroups whose activities led to higher versus lower ANDS may serve to create subgroups that have

greater similarity in job activities (and therefore in relationships between tritium dose and ANDS).

A general model was developed assuming a linear relationship between tritium dose and annual estimated whole body dose.

$$Y_{ijkl} = \alpha_{ijkl} + \beta_{ijkl}X + \varepsilon_{ijkl}$$

for $i=1,2,\dots,G$ years, $j=1,2,\dots,F$ area, $k=1,2,\dots,M$ occupation, and $l=1,2,\dots,N$ annual estimated whole body dose group, where

Y_{ijkl} = tritium dose for the l th ANDS group in the k th occupation of the j th area in the i th year.

α_{ijkl} the intercept for the l th ANDS group in the k th occupation of the j th area in the i th year.

x the fixed effect for amount of ANDS exposure

β_{ijkl} amount of ANDS exposure for the l th ANDS group in the k th occupation of the j th area in the i th year.

ε_{ijkl} the random effect for the l th ANDS group in the k th occupation of the j th area in the i th year.

This model is fit using all workers who have known (i.e., computerized) tritium dose values. The parameter estimates obtained from this linear regression model are then used in conjunction with covariate patterns observed for employment-years with missing tritium dose values in order to derive a predicted tritium dose value for that year. This

predictive model allows us to distinguish between a variety of exposure scenarios (as characterized by occupation, area, year, and ANDS) while providing estimates of an individual's tritium dose based on the known levels of his/her coworker's who have known tritium dose values. We conducted analysis using Statistical Analysis Software (SAS v. 8.2 Cary, NC).

We evaluated the reliability of this estimation procedure by comparing observed tritium doses with estimated values. For these evaluations, all observed annual tritium dosimetry records utilized in the estimation were used. A predicted value was derived for each annual tritium dosimetry record using the estimation procedure described above. The correlation of observed and predicted values was calculated as a direct assessment of the model. Further evaluations of this estimation procedure were conducted in order to examine how reliably values were estimated when observed doses were of differing magnitudes and from different historical periods. Box-plots of observed versus estimated doses were created in order to examine how well estimated values compared to observed doses and in order to characterize the percentiles of these distributions. Additionally, box-plots were created of the difference between observed and estimated doses by level of observed dose.

Chapter 3

RESULTS

3.1 Cohort data

Table 1 provides information about the completeness of computerized tritium dosimetry information. In total there are 277,735 employment years over the period 1951-1999. There were recorded tritium doses for 224,357 employment years. This leaves 53,378 employment-years for which tritium dose information is unknown.

The bottom half of Table 1 reports results for those employment years in which the recorded whole body dose is greater than 0. By definition, if the annual estimated whole body dose estimate equals 0 then the annual tritium dose component is equal to 0. Therefore, for those employment-years in which the annual estimated whole body dose is equal to 0, the tritium dose can be taken as known, which leaves 44,604 employment-years for which tritium dose information is unknown.

Table 1 is divided into three columns, corresponding to the periods 1951-1953, 1954-1978, and, 1979-1999. While monitoring information during the period 1951-1953 was incomplete, radiological exposures were minimal for most workers during this period as the first production reactor at SRS went critical in December 1953. Therefore, we do not estimate tritium dose values for years of employment in the period 1951-1953.

From 1979 onward, tritium dose estimates have been routinely computerized at SRS via the HPAREH system. As indicated in Table 1, we have essentially complete

information on annual tritium dose estimates for workers employed in the period 1979-1999.

However, for the period, 1954-1978, computerized records for tritium doses are not complete. We observed 155,281 employment-years from 1954-1978 (Table 1). Computerized tritium doses are available for 71% (109,956) of these employment-years. Since estimation of tritium dose is not needed for those with an ANDS of 0, computerized tritium dose values need to be estimated for 43,381 employment-years.

Table 2 reports the minimum, 1st, 5th, 10th, 25th, 50th, 75th, 90th, 95th, 99th, and maximum values for known tritium dose. In total, the collective sum of recorded tritium doses from 1954-1978 is 17,382 mSv. Tritium values up to the 75th percentile are zero, showing the distribution of values is highly skewed. Table 3 provides a detailed account of the calendar year variation of the mean, standard deviation, and sum of recorded tritium dose.

As posited based on industrial hygiene information, there is variation in the average recorded tritium dose and its ratio to average recorded whole body dose by area and occupation (Table 4). The areas with the highest fraction of annual estimated whole body dose due to tritium include: 232-234-H Tritium processing, 400-D Heavy Water Plant, and 100-Reactors, respectively. The occupations with the highest fraction of tritium include: Heavy Water Operator, Carpenter, and Reactor Operator, respectively.

3.2 Model Evaluation

There were 74,610 recorded tritium dose values for employment-years in the period 1954-1978 on which to base estimation of annual tritium dose values for those

employment-years with missing tritium dose records. As evaluation of our predictive model, we compared the observed and estimated tritium values for these employment-years. Figure 1 illustrates the direct relationship between the observed and estimated tritium dose. Most observed values (all values up to the 75th percentile of the distribution) are zero, so most estimated values are similar, leading to a clustering near the lower range of values of the plot. Tested against our known values, our predicted tritium values produce an R^2 value of 0.69.

Figure 2 shows box-plots comparing predicted tritium dose to known tritium dose in groups of 0.3 mSv. This figure was produced via SAS PROC BOXPLOT with the 'clipping' option used to suppress a small number of outliers (in five of the 13 boxplots) in order to allow for better visualization of the relationship between observed and estimated dose. The mean and median for the category of observed tritium dose equal to zero are 0.06 and 0.00 mSv, respectively. In this group, 95% of the predicted values fall in the range 0 to 0.61 mSv. For the group of observed doses from 0 to 0.30 mSv, the mean and median of estimated tritium dose are 0.21 and 0.11 mSv, respectively (95% of predicted values fall in the range 0.00-1.10). Estimated tritium dose for the observed dose group from 0.31 to 0.60 mSv have a mean and median of 0.52 and 0.40 mSv, respectively (95% of predicted values fall in the range 0.04-1.95). For the observed tritium dose group from 0.61 to 0.90 mSv, mean and median values are 0.64 and 0.60 mSv, respectively (95% of predicted values fall in the range 0.05-2.042). Estimated tritium values for the observed dose group from 0.91 to 1.20 mSv have a mean and median of 1.03 and 0.98 mSv, respectively (95% of predicted values fall in the range 0.12-2.54). For the observed dose group from 1.21 to 1.50 mSv, the mean and median for estimated tritium are 1.28

and 1.30 mSv, respectively (95% of predicted values fall in the range 0.15-2.73). For the observed dose group from 1.51 to 1.80 mSv, the mean and median for estimated tritium dose are 1.52 and 1.56 mSv, respectively (95% of predicted values fall in the range 0.31-3.15). Estimated tritium dose for the observed dose group from 1.81-2.10 mSv have a mean and median of 1.73 and 1.84 mSv, respectively (95% of predicted values fall in the range 0.37-3.21). For observed tritium dose from 2.11 to 2.40 mSv, the mean and median of estimated tritium dose are 1.87 and 1.99 mSv, respectively (95% of predicted values fall in the range 0.34-3.21). Mean and median values of estimated tritium dose for the observed dose group from 2.41 to 2.70 mSv are 2.04 and 2.12 mSv, respectively (95% of predicted values fall in the range 0.42-3.43). For the observed dose group from 2.71 to 3.0 mSv, the mean and median of estimated tritium dose is 2.20 and 2.25 mSv, respectively (95% of predicted values fall in the range 0.46-3.95). Finally, the mean and median of estimated tritium dose for those employees with an observed dose greater than 3.0 mSv are 3.06 and 2.66 mSv, respectively (95% of predicted values fall in the range 0.52-6.90).

Figure 3 shows a box-plot of the error for our predictive model as observed tritium dose minus expected tritium dose for the same observed dose groups in Figure 2. Most of the estimated tritium values match well with observed tritium dose, since the boxes for each observed dose group are near zero; however, the model over predicted lower values and under predicted higher values. The mean and median errors for the lowest dose category (observed dose equal to zero) were -0.06 and 0 mSv, respectively (95% of values fall in the range -0.61-0.00). For the highest observed dose category

(observed dose greater than 3.0 mSv), the mean and median of error were 1.89 and 1.36 mSv, respectively (95% of values fall in the range -0.50-9.07).

Lastly, we examine error by year, occupation, and area. Table 5 shows the mean and 95% distributions for estimated tritium dose for employees with known tritium dose by year. The median value for all years is zero, and is not reported on this table. This is mentioned to highlight the nature of the distribution as being highly skewed towards zero values, which is also the case for known tritium dose records by year (Tables 2 and 3).

Table 6 displays mean and 95% distributions for estimated tritium dose for those employees with known tritium dose by occupation and area. The order of dose level for these estimates matches that of the known tritium dose. The 95% distributions for estimated tritium doses for the occupation and area groups show some workers in certain categories with a low estimated mean tritium dose are actually getting relatively high exposure (see Administration, Table 6). This 95% distribution of estimated values can be compared to that for tritium dose estimates for employment-year records with no known tritium dose.

3.3 Estimating tritium values for employment-years lacking computerized tritium records

We estimated 43,381 missing tritium values with a mean of 0.10 mSv and median of 0.00 mSv. The collective sum of our estimated tritium values is 4,319 mSv. Combined with our known collective tritium dose from 1954-1978, the collective sum of estimated tritium values represents 20% of the collective sum of tritium dose.

Table 7 reports the minimum, 1st, 5th, 10th, 25th, 50th, 75th, 90th, 95th, 99th, and maximum values for estimated tritium dose. The lower 75th percentile of estimated values

is equal to zero. Table 8 provides the mean, sum, and the range spanning 95% of the distribution of predicted tritium doses by year. Most of the estimation takes place for the years 1954-1969, as these years represent 91% (n=39,498) of the total missing tritium dose records.

Table 9 provides the mean and the range spanning 95% of the distribution of the estimated tritium dose values for employees with no known tritium dose. The levels of dose from our estimation do not match the order of the levels of dose for our known tritium dose. For example, estimated tritium dose was highest for the areas 100-Reactors and 232-234-H (Tritium Process/reservoir), respectively, while known tritium dose was highest for areas 100-Reactors and 400-D Heavy Water Plant. The fraction of tritium dose to ANDS is highest for 232-234-H (Tritium Process/reservoir), which may explain the higher estimated tritium dose.

In addition, the upper bounds of the range spanning 95% of the distribution of estimated tritium doses (i.e., the 97.5 percentile of the distribution) by area and occupation for employment-years without a known tritium dose are lower than the same boundary for the upper bounds of the 95% of the distribution of estimated tritium dose for area and occupation groups with known tritium dose records. This shows that many of the estimated tritium dose values are zero or near zero values, as is indicated in Table 7. A few occupations did not follow this trend, including: general service operator, engineering technicians and trainees, engineer, junior physicist, chemist, raw materials operator, students, laboratory supervisor, senior/process chemist, senior/process physicist, life sciences and medical services, clerical and kindred non-manual workers, and administrators and professionals (Tables 6 and 9). Some of these occupations had higher

upper bound estimates than others, and may be the result of overestimation due to a few high exposure events for individuals within the same exposure category (defined in our model) as a few other workers.

Chapter 4

DISCUSSION

This paper presents an innovative hybrid of traditional approaches to exposure assessment based upon the method of job exposure matrices in conjunction with quantitative dose estimation approaches using stratified regression models. Our evaluations suggest that overall this approach has performed extremely well (Figures 5-7).

A large proportion of the employment-years for which we estimated tritium dose were for workers employed in areas other than reactor operations and tritium loading/recycling. It is likely, therefore, that a large proportion of the workers with missing tritium dosimetry information had little or no occupational exposure to tritium. This is reasonably well reflected by the estimated tritium dose values for workers outside of the reactor and tritium areas: 79% ($n=34,334$) of estimated tritium doses are 0 mSv, and 90% of all estimated doses are less than 0.121 mSv. One aspect of the estimation method is important to recognize in order to understand why small positive dose values have been assigned to a large number of workers who were likely to have had little or no true exposure. If a worker's true dose was zero then our estimation approach will tend to assign a value that is a slight overestimate of the 'true' zero value. Based upon evaluation of our model when the true value for an annual tritium dose was zero, the mean and median estimated values derived were 0.057 mSv and 0 mSv, respectively.

The reason for this is that an estimated dose value may be greater than zero but it can not be less than zero. In contrast, if a true dose is high relative to the other dose records (i.e. the dose is in the 99th and higher percentile range) then the method tends to produce a positive estimate that will slightly underestimate the true dose. So, our model tends to lead to a small overestimate of the value for an unrecorded dose year in which the true dose was zero and a small underestimate of the value for an unmonitored year in which the true dose was greater than zero. If we had information that indicated that a worker was employed in an area in which the tritium dose rate was zero then we could assign a zero dose to that year. In effect, however, our estimation method does this: if the worker was employed in a occupation/area where the average recorded value was zero or near zero then the worker gets assigned a zero (or near zero) value for that year. Inclusion of annual whole body dose groups assists with this by clustering individuals based on their known ANDS, since with zero (or near zero) ANDS will tend to have a similar tritium dose, depending on the occupation/area combination. Our method offers an approach to imputing a distribution of estimated tritium dose values for employment-years with missing information using the available data from monitored workers in the year as well as information on occupation and area of employment.

Like the results of any dose reconstruction model, our predicted values are not perfect. However, for the ultimate goals of this project, which is to examine any possible associations between tritium exposure and adverse health effects in workers at SRS, this predictive model provides useful information about occupational exposures to tritium. In addition, our work provides a new level of understanding about a form of radiation exposure that has, thus far, received little attention.

Our model is essentially a job exposure matrix (JEM), and JEMs have always had a number of limitations to overcome. First, utilizing qualitative information to provide ordinal or dichotomous exposure classification is not always ideal for examination of exposure-disease relationships. Job and area classifications are not necessarily created for the purposes of distinguishing between tasks or exposures. Rather, they may represent distinctions created simply for administrative purposes (Loomis et al. 1994). In addition, since within-job variation cannot be taken into account, JEMs suffer from non-differential misclassification of exposure (De Vocht et al. 2005).

Our model attempts to overcome these obstacles by combining an industrial hygiene approach to evaluating exposure with an empirical method of exposure prediction. We found this to be ideal, and believe that it provides an understanding about exposure to tritium at SRS that may otherwise have been overlooked.

Firstly, providing quantitative estimates of tritium exposure will benefit further studies of workers exposed to this radionuclide. The estimation of tritium dose is based on known tritium and annual whole body dose exposure. Combining estimated and known tritium doses provides a complete exposure history for employees at SRS. This level of detail about exposure is more useful than that obtained in a typical JEM.

Also, our method attempts to overcome the obstacle discussed by Loomis et al. (1994) concerning codes that do not provide substantive distinctions between different jobs and areas. For example, one might assume "administration" is an area of employment that would not lead to high levels of tritium exposure. However, the average tritium dose and fraction of WBDS in this area was higher than the areas of "experimental fuel and target fabrication" and "200-H (H-Main Gate and H-Trit)," two

areas that may be expected to have a higher tritium exposure. This fact may have been overlooked, possibly leading to misclassification of the area administration as a low or no exposure area, or even 200-H (H-Main Gate and H-Trit) as a high exposure area.

Occupation and area descriptions for SRS employees may not provide an ideal distinction between different employees' true area and occupation. We can see this when comparing the number of area and occupation categories we know as compared to the ORAU Team Dose Reconstruction Project for NIOSH at SRS, which contains more specific and detailed descriptions of facilities and processes. Though our technique of dose reconstruction can help overcome this problem, our dataset is limited in the picture it paints of exposure scenarios for tritium at SRS.

Though there are limits to what dose reconstruction can tell us about tritium dose, it is essential that we estimate exposure for employees at SRS. Increasing our volume of knowledge of tritium exposure will help us better evaluate potential relationships between exposure and disease. In addition, we may be able to use estimated tritium dose records to adjust available estimates of annual whole body dose in order to take into account changes over time in tritium dose estimation methods, including International Commission on Radiological Protection models and quality factors for tritium. We hope this study provides useful information for future studies at SRS, and perhaps for other facilities where worker exposure to tritium is a point of concern.

Table 1. Description of employee records for entire SRS occupational cohort from 1951-1999.

	1951-1953 Number	1954-1978 Number	1979-1999 Number	Total Number
Total Employment Years	9,014	155,281	113,440	277,735
Known tritium	2,294	109,956	112,107	224,357
Missing tritium	6,720	45,325	1,333	53,378
Employment Years with ANDS > 0	1457	117,991	58,772	178,220
Known tritium	800	74,610	58,206	133,616
Missing tritium	657	43,381	566	44,604

Table 2. Distribution of known annual tritium doses (mSv) for employment-years between 1954 and 1978, Savannah River Site.

N	74,610
Mean	0.231
Percentile	
Minimum	0.000
1%	0.000
5%	0.000
10%	0.000
25%	0.000
50%	0.000
75%	0.000
90%	0.650
95%	1.650
99%	3.900
Maximum	86.450
Sum	17,382

Table 3. Mean, standard deviation, and sum of known tritium dose by year (1954-1978).
Also, ratio of annual mean tritium dose to estimated annual whole body dose (ANDS).
Savannah River Site.

Year	Tritium N	Mean	Sum	ANDS Mean	Ratio
1954	1294	0.00	1.743	0.40	0.0033
1955	2058	0.02	34.766	1.42	0.0119
1956	2966	0.07	211.256	2.15	0.0332
1957	3136	0.11	356.182	3.03	0.0374
1958	2907	0.13	377.894	2.55	0.0510
1959	2517	0.14	347.112	3.27	0.0422
1960	3084	0.36	1096.75	5.15	0.0691
1961	3138	0.28	869.723	3.46	0.0802
1962	3195	0.23	729.638	3.11	0.0735
1963	3077	0.19	576.483	3.64	0.0515
1964	2914	0.30	863.787	5.26	0.0564
1965	2932	0.26	758.456	4.74	0.0546
1966	2834	0.18	519.077	4.53	0.0404
1967	3158	0.13	399.015	5.17	0.0244
1968	3229	0.17	553.94	4.76	0.0360
1969	3336	0.30	1009.16	5.60	0.0540
1970	3004	0.27	825.183	5.62	0.0489
1971	2722	0.44	1192.68	6.52	0.0672
1972	2579	0.47	1221.21	4.79	0.0989
1973	2806	0.34	962.087	4.19	0.0818
1974	3155	0.29	914.096	3.46	0.0838
1975	3332	0.28	924.29	2.73	0.1015
1976	3510	0.25	890.972	2.54	0.1001
1977	3714	0.27	986.497	2.61	0.1018
1978	4013	0.20	822.62	2.30	0.0890

Table 4. Number of employment-years (N), the % of which are missing tritium dose, the mean and standard deviation (Std Dev) of tritium dose, mean of annual estimated whole body dose (ANDS), and the ratio of tritium dose to ANDS by area and occupation. (n=74,610) Savannah River Site 1954-1978.

	Tritium				ANDS	
	N	% Missing	Mean	Std Dev	Mean	Ratio
<u>Area</u>						
100 (Reactors)	25938	32.2	0.702	1.221	5.016	0.14
400-D (Heavy Water Plant)	2597	28.8	0.451	1.030	2.262	0.20
232-234-H (Tritium Process/reservoir)	975	1.4	0.404	1.059	1.468	0.28
Administration	3145	56.0	0.159	1.623	1.214	0.13
200-H (H-Main Gate and H-Trit)	13271	28.4	0.106	1.076	4.890	0.02
773-A (Experimental Fuel and Target Fabrication)	10916	26.0	0.088	0.848	3.979	0.02
T&T, E&I, other Plant Services	4675	34.9	0.079	0.291	2.735	0.03
Unknown	13869	74.0	0.075	0.798	1.549	0.05
200-F (F-Main Gate)	17090	27.5	0.058	0.516	5.784	0.01
777-M (Experiment Physics Lab) & CMX	1913	32.9	0.054	0.648	2.782	0.02
Administration and Services	8507	43.4	0.052	0.789	1.066	0.05
Physical Plant	6786	37.9	0.046	0.286	1.264	0.04
300-M (Raw Materials/Fuel and Target Fabrication)	8309	28.8	0.024	0.303	2.993	0.01
<u>Occupation</u>						
Reactor Operator	2414	12.1	1.298	1.271	7.560	0.17
Heavy Water Operator	410	22.4	1.280	1.557	2.037	0.63
Auxiliary Operator	3527	22.1	0.733	1.213	9.609	0.08
Radiation Monitor, Health Physicist	2145	9.1	0.539	1.243	10.876	0.05
Rigger	764	35.9	0.426	0.753	11.159	0.04
Other skilled manual	28553	39.2	0.311	0.939	4.115	0.08
Separations/Process Operator	3178	10.4	0.263	1.182	8.216	0.03
Unknown	2984	12.1	0.241	1.850	4.827	0.05
Carpenter	56	83.9	0.228	0.456	1.043	0.22
Technicians, Analysts, and Assistants	9068	25.5	0.198	1.413	6.704	0.03
Other Operator	10975	49.5	0.188	0.715	3.846	0.05
Production/Shift Supervisor	4443	24.5	0.186	0.610	3.146	0.06
Utility Operator	459	51.6	0.184	0.492	6.242	0.03
General Service Operator	4900	30.1	0.086	0.676	2.201	0.04
Production Operator	101	25.7	0.085	0.330	1.000	0.09
Other Supervisors	4697	33.4	0.085	0.367	1.554	0.05
Engineering Technicians and Trainees	2246	23.6	0.081	0.376	3.059	0.03
Managers, Specialist, and Associates	828	53.9	0.065	0.345	0.832	0.08
Engineers	6209	29.3	0.047	0.320	1.455	0.03
Junior Physicist	414	28.3	0.040	0.565	2.220	0.02
Chemists	1371	24.8	0.031	0.282	1.948	0.02
Raw materials Operator	1696	11.4	0.026	0.265	5.388	0.00
Students	1515	75.9	0.025	0.247	0.849	0.03
Crane Operator	961	23.3	0.023	0.149	3.870	0.01
Senior Engineers	853	15.1	0.021	0.139	1.036	0.02
Laboratory Supervisor	513	16.2	0.018	0.161	3.677	0.00
Senior chemist, process chemist	594	21.7	0.018	0.273	1.516	0.01
Senior Physicist, Process Physicist	574	17.1	0.016	0.100	1.689	0.01
Life scientists and medical services	442	58.1	0.008	0.077	0.519	0.02
Other semi-skilled workers	8211	51.0	0.006	0.073	0.582	0.01
Clerical and kindred non-manual workers	10571	68.4	0.005	0.076	0.426	0.01
Metallurgists	192	24.5	0.004	0.025	1.175	0.00
Power Operator	2044	31.2	0.003	0.030	0.426	0.01
Administrators and Professionals	83	60.2	0.000	0.000	0.354	0.00

Table 5. Mean and 95% of the distribution of estimated tritium dose for employees with known tritium dose by year (n=74,610), SRS 1954-1978.

Year	Mean	Percentile	
		2.5 th	97.5 th
1954	0.000	0.000	0.000
1955	0.017	0.000	0.147
1956	0.072	0.000	0.333
1957	0.114	0.000	0.997
1958	0.132	0.000	1.202
1959	0.138	0.000	1.258
1960	0.356	0.000	3.586
1961	0.278	0.000	2.630
1962	0.228	0.000	2.145
1963	0.188	0.000	2.115
1964	0.297	0.000	3.095
1965	0.259	0.000	2.790
1966	0.185	0.000	1.817
1967	0.128	0.000	0.762
1968	0.174	0.000	1.400
1969	0.310	0.000	2.141
1970	0.278	0.000	2.019
1971	0.441	0.000	2.860
1972	0.477	0.000	3.093
1973	0.343	0.000	2.585
1974	0.290	0.000	2.245
1975	0.278	0.000	2.353
1976	0.255	0.000	2.274
1977	0.267	0.000	2.201
1978	0.205	0.000	1.904

Table 6. Mean and 95% of the distribution of estimated tritium dose for employees with known tritium dose by health physics area and occupation (n=74,610), SRS 1954-1978.

Area	Mean	Percentile	
		2.5 th	97.5 th
100 (Reactors)	0.703	0.000	3.250
400-D (Heavy Water Plant)	0.454	0.000	3.565
232-234-H (Tritium Process/reservoir)	0.408	0.000	2.204
Administration	0.163	0.000	1.550
200-H (H-Main Gate and H-Trit)	0.109	0.000	0.859
773-A (Experimental Fuel and Target Fabrication)	0.092	0.000	0.935
T&T, E&I, other Plant Services	0.082	0.000	0.676
Unknown	0.076	0.000	0.476
200-F (F-Main Gate)	0.061	0.000	0.595
777-M (Experiment Physics Lab) & CMX	0.054	0.000	0.328
Administration and Services	0.053	0.000	0.299
Physical Plant	0.048	0.000	0.540
300-M (Raw Materials/Fuel and Target Fabrication)	0.025	0.000	0.157
<u>Occupation</u>			
Reactor Operator	1.299	0.000	3.303
Heavy Water Operator	1.285	0.000	5.266
Auxiliary Operator	0.735	0.000	3.355
Radiation Monitor, Health Physicist	0.548	0.000	3.951
Rigger	0.428	0.000	2.310
Other skilled manual	0.313	0.000	2.332
Separations/Process Operator	0.270	0.000	1.813
Unknown	0.243	0.000	2.381
Technicians, Analysts, and Assistants	0.228	0.000	1.150
Carpenter	0.204	0.000	1.856
Production/Shift Supervisor	0.189	0.000	2.011
Utility Operator	0.188	0.000	1.800
Other Operator	0.188	0.000	0.967
General Service Operator	0.088	0.000	0.550
Other Supervisors	0.088	0.000	1.050
Production Operator	0.086	0.000	0.929
Engineering Technicians and Trainees	0.082	0.000	1.055
Managers, Specialists, and Associates	0.065	0.000	0.797
Engineers	0.047	0.000	0.450
Junior Physicist	0.040	0.000	0.000
Chemist	0.031	0.000	0.200
Raw materials Operator	0.028	0.000	0.250
Students	0.025	0.000	0.032
Crane Operator	0.023	0.000	0.316
Senior Engineers	0.021	0.000	0.150
Laboratory Supervisor	0.018	0.000	0.015
Senior chemist, process chemist	0.018	0.000	0.019
Senior Physicist, Process Physicist	0.016	0.000	0.151
Life scientists and medical services	0.008	0.000	0.000
Other semi-skilled workers	0.007	0.000	0.051
Clerical and kindred non-manual workers	0.005	0.000	0.019
Metallurgists	0.004	0.000	0.044
Power Operator	0.003	0.000	0.021
Administrators and Professionals	0.000	0.000	0.000

Table 7. Distribution of estimated tritium dose for employees with no known tritium dose (mSv), SRS 1954-1978.

N	43,381
Mean	0.098
Percentile	
Minimum	0.000
1%	0.000
5%	0.000
10%	0.000
25%	0.000
50%	0.000
75%	0.000
90%	0.121
95%	0.508
99%	2.135
Maximum	74.533
Sum	4,319

Table 8. Mean and 95% of the distribution of estimated tritium dose for employees without a known tritium dose records by year (n=43,381), SRS 1954-1978.

Year	Mean	Sum	Percentile	
			2.5 th	97.5 th
1954	0.000	1.185	0.000	0.000
1955	0.018	54.024	0.000	0.133
1956	0.053	204.549	0.000	0.217
1957	0.080	312.021	0.000	0.818
1958	0.079	245.215	0.000	0.732
1959	0.091	238.481	0.000	0.836
1960	0.217	608.229	0.000	2.347
1961	0.163	458.074	0.000	1.907
1962	0.110	316.723	0.000	1.196
1963	0.121	300.299	0.000	1.513
1964	0.143	289.673	0.000	1.928
1965	0.131	225.442	0.000	2.087
1966	0.099	147.779	0.000	1.314
1967	0.055	85.428	0.000	0.591
1968	0.098	141.091	0.000	0.803
1969	0.125	153.551	0.000	1.831
1970	0.142	133.634	0.000	1.639
1971	0.280	169.195	0.000	2.654
1972	0.305	147.831	0.000	2.627
1973	0.057	23.446	0.000	0.495
1974	0.043	18.140	0.000	0.213
1975	0.046	16.008	0.000	0.300
1976	0.035	10.889	0.000	0.320
1977	0.039	8.786	0.000	0.432
1978	0.065	9.010	0.000	0.925

Table 9. Mean and 95% of the distribution of estimated tritium dose by area and occupation for employees without a known tritium dose (n=43,381), SRS 1954-1978

Area	Mean	Percentile	
		2.5 th	97.5 th
100 (Reactors)	0.344	0.000	2.595
400-D (Heavy Water Plant)	0.140	0.000	1.752
232-234-H (Tritium Process/reserv	0.297	0.000	2.195
Administration	0.040	0.000	0.300
200-H (H-Main Gate and H-Trit)	0.055	0.000	0.521
773-A (Experimental Fuel and Targ	0.020	0.000	0.111
T&T, E&I, other Plant Services	0.101	0.000	0.651
Unknown	0.038	0.000	0.195
200-F (F-Main Gate)	0.031	0.000	0.298
777-M (Experiment Physics Lab) & i	0.017	0.000	0.064
Administration and Services	0.016	0.000	0.088
Physical Plant	0.046	0.000	0.170
300-M (Raw Materials/Fuel and Tar	0.011	0.000	0.052
<u>Occupation</u>			
Reactor Operator	0.000	0.000	0.000
Heavy Water Operator	0.001	0.000	0.000
Auxiliary Operator	0.095	0.000	1.250
Radiation Monitor, Health Physicist	0.000	0.000	0.000
Rigger	0.001	0.000	0.000
Other skilled manual	0.010	0.000	0.080
Seperations/Process Operator	0.019	0.000	0.075
Unknown	0.000	0.000	0.000
Technicians, Analysts, and Assistani	0.001	0.000	0.000
Carpenter	0.011	0.000	0.177
Production/Shift Supervisor	0.000	0.000	0.000
Utility Operator	0.000	0.000	0.000
Other Operator	0.000	0.000	0.000
General Service Operati	0.130	0.000	1.422
Other Supervisors	0.019	0.000	0.206
Production Operator	0.001	0.000	0.000
Engineering Technicians and Trainee	0.762	0.000	4.739
Managers, Specialists, and Associat	0.028	0.000	0.096
Engineers	0.449	0.000	2.835
Junior Physicist	0.717	0.000	3.450
Chemist	0.104	0.000	0.837
Raw materials Operator	0.065	0.000	0.705
Students	0.734	0.000	3.170
Crane Operator	0.001	0.000	0.000
Senior Engineers	0.000	0.000	0.000
Laboratory Supervisor	0.053	0.000	0.537
Senior chemist, process chemist	0.004	0.000	0.024
Senior Physicist, Process Physicist	0.123	0.000	1.184
Life scientists and medical services	0.453	0.000	4.150
Other semi-skilled workers	0.000	0.000	0.000
Clerical and kindred non-manual wc	0.188	0.000	1.781
Metallurgists	0.003	0.000	0.006
Power Operator	0.001	0.000	0.000
Administrators and Professionals	0.047	0.000	0.298

Figure 1. Plot of estimated vs observed tritium dose (1954-1978, SRS).

Figure 2. Box-plots of estimated vs observed tritium mean by group of observed values (1954-1978, SRS).

Figure 3. Box-plots of difference between predicted and known tritium values (1954-1978, SRS).

References

- Cragle D, McLain R, Qualters J, Hickey J, Wilkinson G, Tankersley W, Lushbaugh C (1988). Mortality Among Workers at a Nuclear Fuels Production Facility. *Am J Ind Med* 14: 379-401.
- De Vocht F, Zock JP, Kromhout H, Sunyer J, Anto J, Burney P, Kogevinas M (2005). Comparison of Self-Reported Occupational Exposure With a Job Exposure Matrix in an International Community-Based Study on Asthma. *Am J Ind Med* 47:434-442.
- Elliot A (2005). Tritium dosimetry. *Nuclear Medicine Communications* 26:481-482.
- Goldberg M, Kromhout H, Guenel P, Fletcher A, Gerin M, Glass DC, Heederik, Kauppinen T, Ponti A (1993). Job Exposure Matrices in Industry. *Int J Epidemiology* 22 (Suppl. 2): S10-S15.
- Hickey J, Cragle D (1985). Occupational Exposures of Workers to Chemicals at the Savannah River Plant 1952-1984. Oak Ridge Associated Universities. Oak Ridge, TN.
- Hill R, Johnson J (1993). Metabolism and Dosimetry of Tritium. *Health Physics* 65(6): 628-647.
- Hoar S (1983-1984). Job Exposure Matrix Methodology. *J Toxicol Clin Toxicol* 21(1/2): 9-26.
- Joksic G, Spasojevic-Tisma V (1998). Chromosome analysis of lymphocytes from radiation workers in tritium-applying industry. *Int Arch Occup Environ Health* 71:213-220.
- Kromhout H, Heederik D, Dalderup L, Kromhout D (1992). Performance of Two General Job-Exposure Matrices in a Study of Lung Cancer Morbidity in the Zutphen Cohort. *Am J Epidemiology* 136: 698-711.
- Loomis D, Peipins L, Browning S, Howard R, Kromhout H, Savitz D (1994). Organization and Classification of Work History Data in Industry-Wide Studies: An Application to the Electric Power Industry. *Am J Ind Med* 26: 413-425.
- McGuire V, Nelson L, Koepsell T, Checkoway H, Longstreth WT (1998). Assessment of Occupational Exposures in Community-Based Case-Control Studies. *Annu Rev Public Health* 19:35-53.
- Okada S, Momoshima N (1993). Overview of Tritium: Characteristics, Sources and Problems. *Health Physics* 65(6):595-609.

Richardson D, Wing S, Daniels RD (2006). Evaluation of external radiation dosimetry records at the Savannah River Site, 1951-1999. *J Expo Anal Environ Epidemiol*: 1-12.

Straume T (1993). Tritium Risk Assessment. *Health Physics* 65(6): 673-682.

Taylor GA, Crase KW, La Bone TR, Wilkie WH (1995). A History of Personnel Radiation Dosimetry at The Savannah River Site. US Department of Energy. Springfield, VA.

Zartarian V, Bahadori T, KcKone T (2005). Adoption of an official ISEA glossary. *J Expo Anal Environ Epidemiol* 15:1-5.

Appendices

SAS code

```
libname trit 'C:\Documents and Settings\All Users\Desktop\Tritium project\tritium data';  
run;
```

```
proc contents data=trit.Ghassan; run;
```

```
*convert data to mSv, exclude years from model;
```

```
data tmp;  
  set trit.Paper;  
  newtrit=newtrit/100;  
  ands=ands/100;  
  if hparea=. then hparea=15;  
  else if hparea=14 then hparea=15;  
  if year > 1978 then delete;  
  if year < 1954 then delete;  
  if ands > 0;  
run;
```

```
proc univariate data=tmp; var newtrit; run;
```

```
*-----univariate by year-----;
```

```
proc sort data=tmp; by year; run;
```

```
proc univariate data=tmp; by year; var newtrit; run;
```

```
*-----ANDS groups for more precise estimation-----;
```

```
data tmp; set tmp; if ands =0 then wbds_gp=0;  
if 0< ands <= 2.50 then wbds_gp=1; if 2.50< ands <= 5.00 then wbds_gp=2;  
if 5.00< ands <= 7.50 then wbds_gp=3;  
if 7.50< ands <= 10.00 then wbds_gp=4; if 10.00< ands <= 12.50 then wbds_gp=5; if  
12.50< ands then wbds_gp=6; run;
```

```
*distribution of known tritium values;
```



```

proc univariate data=tmp; var newtrit;run;

*-----predictive model-----;

proc sort data=tmp; by hparea year wbds_gp occgp ; run;

proc glm data=tmp noprint;
  by hparea year wbds_gp occgp;
  class occgp wbds_gp;
  model newtrit= ands ;
  output out=out p=p;
  run; quit;

*organize output dataset;

data out; set out; pred= p; obs= newtrit; if pred < 0 then pred=0;
run;

*plot of estimated against observed, R2 value-----;

proc reg data=out ; model pred=obs;
run;
quit;

proc gplot data=out; plot pred*obs; title1 'Validation of tritium estimation procedure';
title2 'Estimated values for those with ANDS > 0 based upon';
title3 'Ratio of tritium to ANDS for strata defined by year, hparea, occgp,ands_gp ';
title4 'Observed versus estimated tritium doses n=74,610 '; title5 'SRS 1954-1978'; label
obs ='Observed Tritium Dose (mSv)'; label pred ='Estimated Tritium Dose (mSv)'; run;
quit;

*calculate % estimated dose = 0-----;

data zero;
  set pred;
  if pred>0 then pred=1;
  run;

proc freq data=zero;
  table pred;
  run;

*-----calculate distribution of estimated tritium for workers with a known tritium dose--
-----;

```

```

data pred1;
  set out;
  if newtrt >=0;
  run;

proc univariate data=pred1;
  var pred;
  run;

proc sort data=pred1; by year; run;

proc univariate data=pred1 noprint;
  by year;
  var pred;
  output out=year mean=mean pctlpre=P pctlpts=2.5 , 97.5 max=max;  run;

proc sort data=pred1; by hparea; run;

proc univariate data=pred1 noprint;
  by hparea;
  var pred;
  output out=hparea mean=mean pctlpre=P pctlpts=2.5 , 97.5 max=max;  run;

proc sort data=pred1; by occgp; run;

proc univariate data=pred1 noprint;
  by occgp;
  var pred;
  output out=occgp mean=mean pctlpre=P pctlpts=2.5 , 97.5 max=max;  run;

*-----distribution for predicted dose when known dose = 0;

data pred2;
  set out;
  if newtrt = 0;
  run;

proc univariate data=pred2;
  var pred;
  run;

*-----dist for predicted trit when trit dose is unknown-----;

```

```

data pred3;
  set out;
  if newtrit >= 0 then delete;
  run;

proc sort data=pred3; by year; run;

proc univariate data=pred3 noprint;
  by year;
  var pred;
  output out=year mean=mean pctlpre=P pctlpts=2.5 , 97.5 max=max; run;

proc sort data=pred3; by hparea; run;

proc univariate data=pred3 noprint;
  by hparea;
  var pred;
  output out=hparea mean=mean pctlpre=P pctlpts=2.5 , 97.5 max=max; run;

proc sort data=pred3; by occgp; run;

proc univariate data=pred3 noprint;
  by occgp;
  var pred;
  output out=occgp mean=mean pctlpre=P pctlpts=2.5 , 97.5 max=max; run;

*-----boxplot of prediction vs known-----;

data boxplot;
  set out;
  error = obs-pred;
  error2 = error*error;
  if obs = 0 then tritgrp=0;
  if 0 < obs <=0.3 then tritgrp=1;
  if 0.3 < obs <= 0.6 then tritgrp=2;
  if 0.6 < obs <= 0.9 then tritgrp=3;
  if 0.8 < obs <= 1.2 then tritgrp=4;
  if 1.2 < obs <= 1.5 then tritgrp=5;
  if 1.5 < obs <= 1.8 then tritgrp=6;
  if 1.8 < obs <= 2.1 then tritgrp=7;
  if 2.1 < obs <= 2.4 then tritgrp=8;
  if 2.4 < obs <= 2.7 then tritgrp=9;
  if 2.7 < obs <= 3.0 then tritgrp=10;
  if 3.0 < obs then tritgrp=11;
  run;

```

```

proc format;
  value tgrp 0='0'
            1='0-.30'
            2='.31-.60'
            3='.61-.90'
            4='.91-1.20'
            5='1.21-1.50'
            6='1.51-1.80'
            7='1.81-2.10'
            8='2.11-2.40'
            9='2.41-2.70'
            10='2.71-3.00'
            11='>3.00'
;
run;

proc sort data=boxplot; by tritgrp; run;

proc boxplot data=boxplot;
  where tritgrp ne 11;
  plot pred*tritgrp/ boxstyle=schematic;
  format tritgrp tgrp.;
  title1 'Validation of tritium estimation procedure';
  title2 'Estimated values for those with ANDS ne 0 based upon';
  title3 'Ratio of tritium to WBDS for strata defined by year, hparea, occgp,ands_gp';
  title4 'Observed versus estimated tritium doses n = 74,610';
  title5 'SRS 1954-1978';
  label tritgrp ='Observed Tritium Dose (mSv)';
  label pred='Estimated Tritium Dose (mSv)';
run;

proc boxplot data=boxplot;
  plot error*tritgrp / BOXSTYLE=SCHEMATIC skiplabels=0;
  format tritgrp tgrp.;
  title1 'Validation of tritium estimation procedure';
  title2 'Estimated values for those with ANDS ne 0 based upon';
  title3 'error=observed-expected (mSv)';
  title4 'n=74,610';
  title5 'SRS 1954-1978';
  label tritgrp ='Observed Tritium Dose (mSv)';
run;

proc sort data=boxplot; by tritgrp; run;

```

```

proc boxplot data=boxplot;
*   where tritgrp ne 11;
plot pred*tritgrp/ boxstyle=schematic
skiplabels=0
clipfactor=5 idsymbol=star
haxis=axis1
cframe    = vlight
cboxes    = dagr
cboxfill   = ywh
clipsymbol = dot
clipleghpos = top
clipllegend = '# Clipped Boxes'
clipsubchar = '#';
format tritgrp tgrp.;
title1 'Validation of tritium estimation procedure';
title2 'Estimated values for those with ANDS ne 0 based upon';
title3 'Ratio of tritium to WBDS for strata defined by year, hparea, occgp,ands_gp';
title4 'Observed versus estimated tritium doses n = 74,610';
title5 'SRS 1954-1978';
label tritgrp ='Observed Tritium Dose (mSv)';
label pred='Estimated Tritium Dose (mSv)';
run;

```

*calculate 95%CI for boxplots;

```

proc univariate data=boxplot noprint;
by tritgrp;
var pred;
output out=boxplot mean=mean median=median pctlpre=P pctlpts=2.5 , 97.5;
run;

```

```

proc univariate data=boxplot noprint;
by tritgrp;
var error;
output out=errorbox mean=mean median=median pctlpre=P pctlpts=2.5, 97.5;
run;

```

*----- Create Tables for mean by year for top 5 occ and area-----;

```

proc sort data=tmp; by year occgp; run;

```

```

proc means data=tmp noprint;
by year occgp;
output out=occ mean=mean;
var newtrit;

```

```

run;

proc sort data=occ; by occgp; run;

proc transpose data=occ out=occl;
  by occgp;
  id year;
  idlabel year;
  var mean;
  run;

proc sort data=tmp; by year hparea; run;

proc means data=tmp noprint;
  by year hparea;
  output out=area mean=mean;
  var newtrit;
  run;

proc sort data=area; by hparea; run;

proc transpose data=area out=area1;
  by hparea;
  id year;
  idlabel year;
  var mean;
  run;

*-----Create Tables of overall mean and median-----;

proc sort data=tmp; by occgp; run;

proc means data=tmp;
  by occgp;
  output out=occgp mean=mean median=median std=std;
  var newtrit;
  run;

proc sort data=tmp; by hparea; run;

proc means data=tmp;
  by hparea;
  output out=hparea mean=mean median=median std=std;
  var newtrit;

```

```

run;

proc sort data=tmp; by year; run;

proc means data=tmp;
  by year;
  output out=year mean=mean median=median;
  var newtrit;
run;

proc sort data=tmp; by occgp; run;

proc means data=tmp;
  by occgp;
  output out=occgp1 mean=mean median=median;
  var ands;
run;

proc sort data=tmp; by hparea; run;

proc means data=tmp;
  by hparea;
  output out=hparea1 mean=mean median=median;
  var ands;
run;

proc sort data=tmp; by year; run;

proc means data=tmp;
  by year;
  output out=year1 mean=mean median=median;
  var ands;
run;

*-----MEAN BY OCC AND AREA and YEAR FOR PREDICTED
VALUES-----;

data pred;
  set out;
  if obs >= 0;
run;

proc univariate data=pred; var obs; run;

```

```
proc sort data=pred; by occgp; run;
```

```
proc means data=pred;  
  by occgp;  
  output out=occgp mean=mean median=median std=std;  
  var pred;  
  run;
```

```
proc sort data=pred; by hparea; run;
```

```
proc means data=pred;  
  by hparea;  
  output out=hparea mean=mean median=median std=std;  
  var pred;  
  run;
```

```
proc sort data=pred; by year; run;
```

```
proc means data=pred;  
  by year;  
  output out=year mean=mean median=median std=std;  
  var pred;  
  run;
```

```
*-----CALCULATE SUMS OF TRIT EXPOSURE-----;
```

```
proc sort data=tmp; by year; run;
```

```
proc summary data=tmp noprint;  
  by year;  
  var newtrit;  
  output out=yearsum sum(newtrit)=newtrit;  
  run;
```

```
proc summary data=tmp noprint;  
  var newtrit;  
  output out=sum1 sum(newtrit)=newtrit;  
  run;
```

```
proc sort data=pred1; by year; run;
```

```
proc summary data=pred1 noprint;  
  by year;  
  var pred;  
  output out=yearpred sum(pred)=pred;
```



```

run;

proc summary data=pred noprint;
  var pred;
  output out=sum2 sum(pred)=pred;
run;

*-----;

data tmp2;
  set trit.Paper;
  if newtrit >=0 then newtrit=1;
  if ands >=0 then ands=1;
run;

*-----FIGURE OF MISSING WBDS AND TRIT-----;

proc sort data=tmp2; by year; run;

proc freq data=tmp2;
  by year;
  tables newtrit/ ALL;
  output out=out nmiss n;
run;

data percent;
  set out;
  total= nmiss + n;
  percent=nmiss/total;
run;

proc sort data=tmp2; by year; run;

proc freq data=tmp2;
  by year;
  tables ands/ ALL;
  output out=out1 nmiss n;
run;

data percent1;
  set out1;
  total= nmiss + n;
  percent=nmiss/total;
run;

```

```
*-----missing rates-----;
```

```
data tmp;  
  set trit.Paper;  
  if hparea=. then hparea=15;  
  else if hparea=14 then hparea=15;  
  if year > 1978 then delete;  
  if year < 1954 then delete;  
  if ands > 0;  
  run;
```

```
data tmp;  
  set tmp;  
  if trit >=0 then trit=1;  
  run;
```

```
proc sort data=tmp; by occgp; run;
```

```
proc freq data=tmp;  
  by occgp;  
  tables newtrit/ ALL;  
  output out=occout nmiss n;  
  run;
```

```
proc sort data=tmp; by hparea; run;
```

```
proc freq data=tmp;  
  by hparea;  
  tables newtrit/ ALL;  
  output out=areaout nmiss n;  
  run;
```

```
proc sort data=tmp; by year; run;
```

```
proc freq data=tmp;  
  by year;  
  tables newtrit/ ALL;  
  output out=yearout nmiss n;  
  run;
```

```
*-----information for table 1-----;
```

```
data tmp;  
  set trit.Paper;  
  if year > 1978 then delete;  
  if year < 1954 then delete;  
run;
```

```
data tmp1;  
  set trit.Paper;  
  if year > 1953 then delete;  
run;
```

```
data tmp;  
  set tmp;  
  if newtrit >= 0 then newtrit1 = 1;  
  if ands > 0 then ands = 1;  
  if ands >= 0 then newands = 1;  
run;
```

```
data tmp1;  
  set tmp1;  
  if newtrit >= 0 then newtrit1 = 1;  
  if ands > 0 then ands = 1;  
run;
```

```
proc freq data=tmp;  
  table newtrit1 ands;  
run;
```

```
proc freq data=tmp1;  
  table newtrit1 ands;  
run;
```