U.S. DEPARTMENT OF ENERGY

PNNL-18913

Prepared for the U.S. Department of Energy
under Contract DE-AC05-76RL01830

Assessment of Uncertainty in the Radiation Doses for the Techa River Dosimetry System

BA Napier
MO Degteva
LR Anspaugh
NB Shagina

October 2009

Pacific Northwest
NATIONAL LABORATORY

Proudly Operated by Battelle *Since 1965*

DISCLAIMER

PACIFIC NORTHWEST NATIONAL LABORATORY
operated by
BATTELLE
for the
UNITED STATES DEPARTMENT OF ENERGY
under Contract DE-AC05-76RL01830

Printed in the United States of America

Available to DOE and DOE contractors from the
Office of Scientific and Technical Information,
P.O. Box 62, Oak Ridge, TN 37831-0062;
ph: (865) 576-8401
fax: (865) 576-5728
email: reports@adonis.osti.gov

Available to the public from the National Technical Information Service,
U.S. Department of Commerce, 5285 Port Royal Rd., Springfield, VA 22161
ph: (800) 553-6847
fax: (703) 605-6900
email: orders@ntis.fedworld.gov
online ordering: http://www.ntis.gov/ordering.htm

This document was printed on recycled paper.
(9/2003)

ASSESSMENT OF UNCERTAINTY IN THE RADIATION DOSES FOR THE TECHA RIVER DOSIMETRY SYSTEM

B.A. Napier, M.O. Degteva, L.R. Anspaugh, N.B. Shagina

Pacific Northwest National Laboratory
Richland, Washington, USA

Urals Research Center for Radiation Medicine
Chelyabinsk, Russian Federation

University of Utah
Salt Lake City, Utah, USA

Status Report

US–Russian Joint Coordinating Committee on Radiation Effects Research
Project 1.1:
"Further Studies on Uncertainty and Validation of Doses in the Techa River Dosimetry System"

October 2009

TABLE OF CONTENTS

Abstract

In order to provide more accurate and precise estimates of individual dose (and thus more precise estimates of radiation risk) for the members of the ETRC, a new dosimetric calculation system, the Techa River Dosimetry System-2009 (TRDS-2009) has been prepared. The deterministic version of the improved dosimetry system TRDS-2009D was basically completed in April 2009. Recent developments in evaluation of dose-response models in light of uncertain dose have highlighted the importance of different types of uncertainties in the development of individual dose estimates. These include uncertain parameters that may be either shared or unshared within the dosimetric cohort, and also the nature of the type of uncertainty as aleatory or epistemic and either classical or Berkson. This report identifies the nature of the various input parameters and calculational methods incorporated in the Techa River Dosimetry System (based on the TRDS-2009D implementation), with the intention of preparing a stochastic version to estimate the uncertainties in the dose estimates. This report reviews the equations, databases, and input parameters, and then identifies the author's interpretations of their general nature. It presents the approach selected so that the stochastic, Monte-Carlo, implementation of the dosimetry System - TRDS-2009MC - will provide useful information regarding the uncertainties of the doses.

Introduction

Population exposure in the Urals region occurred as a result of failures in the technological processes in the Mayak plutonium facility in the middle of the 20th century. A major source of environmental contamination was the discharge of about 10^{17} Bq of liquid wastes into the Techa River in 1949–1956. Residents of many villages downstream from the site of release were exposed via a variety of pathways; the more significant included drinking of water from the river and external gamma exposure due to proximity to bottom sediments and the shoreline. There are known to be additional sources of exposure for the Urals population. The most important was an explosion in the radioactive waste-storage facility in 1957 (the so-called Kyshtym accident) that formed the East Urals Radioactive Trace (EURT) due to dispersion of 7.4×10^{16} Bq into the atmosphere. Other sources of exposure include the gaseous aerosol releases from the Mayak facility in 1949–1957 and windblown contamination from Lake Karachay, when this contaminated lake dried out in 1967.

The series of radioactive releases that occurred in the same region in different years and the intensive migration of the population within the contaminated area are specific features of the Urals situation. This determined the approach to follow-up: selecting a fixed cohort and tracing all places of residence for each subject in the cohort since the beginning of radioactive contamination. The Extended Techa River Cohort (ETRC) includes approximately 30,000 members and represents an unselected population consisting of two distinct ethnic groups. The members of the ETRC were exposed to chronic radiation over a wide range of doses, but at low-to-moderate-dose rates.

Russian and US scientists have been working together to perform dose reconstruction and epidemiologic follow-up for the ETRC since 1995. Epidemiologic studies on cancer incidence and mortality (JCCRER Project 1.2b and NCI-URCRM Project) are ongoing; the investigators for these projects work together with those of JCCRER Project 1.1 who are charged to provide credible estimates of individual dose for the ETRC members.

In order to provide more accurate and precise estimates of individual dose (and thus more precise estimates of radiation risk) for the members of the ETRC, a new dosimetric calculation system, the Techa River Dosimetry System-2009 (TRDS-2009) has been prepared. The

deterministic version of the improved dosimetry system TRDS-2009D was basically completed in April 2009 (Degteva et al. 2008; Degteva et al 2009).

Recent developments in evaluation of dose-response models in light of uncertain dose data (Stram and Kopecky 2003; Schafer and Gilbert 2006) have highlighted the importance of different types of uncertainties in the development of individual dose estimates. These include uncertain parameters that may be either shared or unshared within the dosimetric cohort, and also the nature of the type of uncertainty as aleatory or epistemic and either classical or Berkson. This report is an attempt to identify the nature of the various input parameters and calculational methods incorporated in the Techa River Dosimetry System (based on the TRDS-2009D implementation), with the intention of preparing a stochastic version to estimate the uncertainties in the dose estimates. This report reviews the equations, databases, and input parameters, and then identifies the author's interpretations of their general nature. It presents the approach selected so that the stochastic, Monte-Carlo, implementation of the dosimetry System - TRDS-2009MC - will provide useful information regarding the uncertainties of the doses.

Characterization of Uncertainty in Dose Reconstruction

Radiation doses cannot be measured directly. They must necessarily be estimated from other information. The kinds of information used to estimate doses vary widely, depending on the exposure scenario and the measurements, if any, that were made of quantities that influence or reflect the radiation dose received. Dose reconstruction therefore requires the use of mathematical models that express the doses of interest as functions of the available information.

Estimates of radiation doses are almost certain to differ to some degree from the doses actually received, and the differences between the estimated and true doses are of course unknown. To understand the accuracy of the estimated dose, it is important to be able to characterize how large or small the differences might be. This is equivalent to characterizing the plausibility of possible values of the true dose. Ordinarily, some possible true doses are more plausible than others; for example, doses near the estimated dose may be much more plausible than doses which differ by a large amount from the estimate. Therefore, the most useful representation of the plausibility of various true doses assigns levels of plausibility to different doses or sets of doses. The scale upon which plausibility is measured is arbitrary, and may therefore be chosen so that the total plausibility, integrated over all possible true dose values, is

1.0. The appropriately scaled "plausibility function" is, in effect, a probability distribution, and the variability represented by this distribution characterizes the uncertainty in the dose estimate. This distribution is referred to as the "uncertainty distribution" of the estimated dose.

The uncertainty in dose estimates arises from uncertainty about the quantities that determine the true but unobservable dose. Uncertainty distributions are also used to represent the uncertainty about those quantities. Uncertainty distributions can be said to represent the "state of knowledge" about the values of their corresponding doses or quantities.

For quantities that can be observed or estimated from observations, uncertainty distributions can be derived or deduced from statistical analyses of the observations. An uncertainty distribution estimated from data is sometimes described as "objective." This is in contrast to a "subjective" uncertainty distribution, which relies heavily, if not exclusively, on the state of knowledge of the dose reconstruction analyst or other individuals, possibly experts, about the relative likelihoods of possible values of the parameter. Many if not most uncertainty distributions are derived from some combination of empirical data and subjective judgment: when empirical data are limited in terms of availability or applicability, the uncertainty distribution will be largely subjective. When the assessment endpoint involves estimation of the inter-individual variability of true doses within a group or cohort of individuals, such as the ETRC, uncertainties representing random variability of true dose must be separated from uncertainties representing lack of knowledge about parameters that are single true quantities. The assessment objective is the estimate of a frequency distribution of true doses in a cohort of n individuals; the state of knowledge about inter-individual variability of true dose may be represented by numerous alternative realizations of unique vectors of n doses.

Whether or not the uncertainty distribution of any particular parameter needs to be interpreted as a probability distribution in either the classical (frequentist) or Bayesian sense can be a matter of some contention. However, like the distinction between objective and subjective uncertainty distributions, this distinction is largely irrelevant to how uncertainty of dose estimates is assessed, represented, and interpreted. Therefore, distinctions between classical and Bayesian interpretations will not be highlighted here. In general it suffices to recognize that the uncertainty distribution characterizes the state of knowledge about the value of the parameter.

Before describing the general approach to estimating dose uncertainties, some basic concepts regarding the nature of uncertainties are introduced.

Independent Versus Correlated Uncertainties

In order to properly estimate and interpret uncertainty distributions of dose estimates, it is necessary to characterize how and to what extent uncertainties of doses and of input parameters are interrelated. Two uncertain quantities, X and Y, are said to be independent if the uncertainty distribution of X conditional on the fact that $Y = y$ is the same for any value of y. In other words, if X and Y are independent, then knowledge that $Y = y$ does not provide any information about the value of X. If X is independent of Y by this definition, it is also true that Y is independent of X. If X and Y are not independent, they are correlated; that is, knowing the value of one of the quantities provides some information about the likely value of the other.

Correlations can be either positive or negative. For example, mass balancing may induce negative correlation between estimates of environmental contamination in different geographical locations.

If input parameters X and Y of a dose estimation model are independent, the effects of their uncertainties on estimated doses can be dealt with separately. For example, if the uncertainty in X can be reduced by the collection of additional data, it is not necessary to make any adjustments in Y. However, if X and Y are correlated, then any adjustment in one variable must be appropriately reflected in the other. In an additive or multiplicative chain of terms, positive correlations will tend to increase uncertainty in the dose estimate. Negative correlations will tend to decrease uncertainties in the dose estimate. The opposite is true for terms of a quotient or ratio. In general, only strong correlations among parameters that are important contributors to the uncertainty in dose will have a noticeable effect on the uncertainty in the dose estimate. For dose reconstructions supporting an epidemiological analysis, it is important to account for correlations of dose variability and dose uncertainty among individual members of the cohort. If there is correlation between the parameters in a model affecting the uncertainty of more than one individual, the individual uncertainties may not be affected, but the correlations may impact later calculations that use the individual doses.

Correlations exist both between individuals and within components of one individual's calculations. For instance, if doses are calculated annually, the biokinetic parameters used for a specific individual should not vary widely from year to year. Similarly, a single person's dietary and exposure habits should not vary widely from year to year, assuming his residence and job type are constant. These types of correlation are known as autocorrelations.

There is a good deal of overlap in the concepts of correlated uncertainties and shared/unshared uncertainties, discussed below.

Classical Versus Berkson Error Models

When considering uncertainty in an estimate of quantity, such as a dose or a parameter in a dose calculation model, it is important to consider the probabilistic nature of the "error," (i.e., of the difference between that estimate and the unknown true value of the quantity). In many settings the estimate differs from the true value by an error that is stochastically independent of the true value, that is:

$$estimate = true\ value + measurement\ error,$$

where *measurement error* is a random variable with mean zero and is independent of *true value.* This is the classical error model, which arises naturally in settings where observations are made on the individual units for which the quantities are defined, and each observation is subject to random perturbation due to such factors as instrument imprecision and recording errors. It is important to recognize that classical measurement error leads to an observed variability of individual dose estimates that will be larger than the variability of true dose among individuals. Inflation of the inter-individual variability of true dose caused by classical measurement error will bias the epidemiological dose-response function towards the null (i.e., the value of the slope of the dose-response will be smaller than if the dose-response were analyzed based on true values of dose). In contrast, the Berkson error model, in which the true value varies from the estimate by an error that is random and is independent of the estimate, can be written:

$$true\ value = estimate + individual\ peculiarity,$$

where *individual peculiarity* is a random variable with mean zero and is independent of *estimate*. The Berkson model arises, for example, when a single estimate that is unbiased with respect to the mean dose for a subgroup of individuals is applied as a surrogate for the true dose to each individual belonging to that subgroup. For example, suppose only one location in a village exposed to external dose was measured, and that this dose rate reading can be considered to be

an unbiased representation of the dose received by all other locations in the village. Using this single dose rate measurement to indicate the dose rate to all other locations in the village is a form of Berkson error. Uncertainty is the difference between the individual's true dose and the assigned dose. The assigned dose is assumed to be unbiased with respect to the true mean dose for the residents of the village. In this case, it is assumed that the single dose rate measurement would not be biased with respect to the true mean rate dose received by averaging true dose rates among all residents of the village.

Note that both the classical and Berkson error models concern the difference, on an additive scale, between the true and estimated values (for a strictly positive quantity with a multiplicative uncertainty, the classical or Berkson model might apply to true and estimated values of the logarithm of the quantity). In the context of dose reconstruction, the distinction between additive classical and Berkson errors is primarily of concern in statistical analyses of the relationships between uncertain doses and outcome variables such as disease incidence. This can be illustrated by the example of a simple linear regression model $Y = \alpha + \beta Z + \varepsilon$, where the true value of the predictor Z is unknown, and only an estimate of Z subject to uncertainty, say X, is available. Then the effect of regressing Y on X will depend on the nature of the uncertainty in X. If X is subject to classical error, then the estimated slope will be attenuated (*i.e.,* biased toward zero relative to the true slope β). But if X is subject to Berkson error, then regressing Y on X will produce an unbiased estimate of β, although the variance of the estimate of β will be increased by the Berkson error in X (Stram and Kopecky 2003; Hofer 2008; Carroll et al. 2006; Shafer and Gilbert 2006).

Although the distinction between classical and Berkson errors is primarily a concern in dose-response analyses, it is also useful to consider this distinction when evaluating the uncertainty distributions of the input parameters in a dose reconstruction model. This will help ensure that the uncertainties or states of knowledge about the parameters are properly represented in the uncertainties of the calculated doses. Frequently, "environmental parameters" (generally those that are shared by all individuals across a particular dose realization) have Berkson uncertainty structure, and those that are specific to the individual (if known) have a classical uncertainty structure. However, generic models and assumptions are frequently used to fill in for lack of information about specific individuals (e.g., dose conversion factors); in this

case, the parameter can be considered to be Berkson in nature, provided that it is unbiased with respect to the true mean for the group of individuals concerned.

In many settings, neither the classical error model nor the Berkson error model applies. Dose estimation models may include some parameters with uncertainties having classical error models, others Berkson uncertainties, and still other parameters with uncertainties that have neither classical nor Berkson error. The uncertainties of dose estimates from such models are themselves neither purely classical nor purely Berkson (Mallick et al. 2002; Stram and Kopecky 2003; Li et al. 2007; Hoffman et al. 2007).

Discussions of classical and Berkson error models are provided in NCRP (2007), Schaefer and Gilbert (2006), and Carroll et al. (2006).

Type-A Versus Type-B Uncertainties

Type A uncertainty refers to random variability of true values (i.e., variability due to stochastic processes). With respect to dose reconstruction, Type A uncertainty refers to processes that affect random variability of true dose among members of a defined population or cohort. This is the amount of inter-individual variability that cannot be explained by factors such as location, residence history, diet, age, gender, lifestyle, occupation, etc. Type A uncertainty is described by frequency distributions representing the random variability of true values (e.g., of individual dose). In contrast, Type B uncertainty represents lack of complete knowledge about a unique quantity that has a true but unknown (or imperfectly known) value (e.g., the mean value of the milk transfer coefficient for cesium). Type B uncertainty is described by probability distributions composed of possibly true values for the unique unknown quantity. In the absence of experimental data composed of repeated measurements of the unique quantity, Type B probability distributions are subjectively derived; they reflect the investigator's state of knowledge about the true, unknown value. Type A uncertainty is related to aleatory uncertainty, while Type B uncertainty is related to epistemic uncertainty (IAEA 1989).

Often, both Type A and Type B uncertainty exist together. They can be combined when the dose reconstruction is focused on a specific individual, because there is only a single unique value of dose for that individual, (i.e., all uncertainty is effectively Type B). The probability distribution representing the state of knowledge of that individual's true dose is composed of alternative realizations of possibly true values of dose for that person. However, when the dose

reconstruction is focused on estimating inter-individual variability of *n* true doses in a defined population or cohort (such as the Techa River cohort), Type A and Type B uncertainty must be separated. In this case, the dose reconstruction will be composed of alternative realizations of vectors of *n* individual doses, with the *n* number of individuals in the vector equal to the number of persons in the defined population. Each vector of true dose is an expression of inter-individual variability of dose due to (1) factors known to determine differences in exposure and dose, such as location, diet, residence history, gender, etc., and (2) unexplained variability due to stochastic processes (Type A uncertainty). Type B uncertainty is expressed among the entire set of *m* realizations of vectors of *n* doses. The presence of Type B uncertainty will result in each vector of individual doses having a unique value for each quantity that is fixed but unknown. IAEA (1989), NCRP (1996), and NCRP (2009) provide additional discussion about and methods for separating Type A from Type B uncertainty when dose reconstruction is focused on the estimation of inter-individual variability of true dose.

When empirical data are used to define a Type A distribution, efforts should be made to remove uncertainty due to classical measurement error to ensure that the frequency distribution of true values reflects the variability of true values (Hofer 2008). If classical measurement error is not removed, the estimated variability of individual doses will likely exceed inter-individual variability of true doses. As mentioned previously, classical measurement error will inflate the estimate of inter-individual variability of true dose. When dose reconstruction is intended to support an epidemiological investigation, inflation of the estimate of inter-individual variability of true dose is undesirable because overestimation of the variability of individual doses in a cohort will produce a misleading overestimate of statistical power and will suppress the slope of the dose-response towards the null hypothesis (Hofer 2008; Carroll et al. 2006; Shafer and Gilbert 2006). Hofer (2008) gives an example of methods for removing classical measurement error from a Type-A distribution representing pure random inter-individual variability of true doses in a cohort.

There is a good deal of overlap in the concepts of Type A/Type B and shared/unshared uncertainties, discussed below.

Shared and Unshared Uncertainties

It can be important to determine whether each parameter in the model has a unique value for each person whose dose is to be estimated, or whether the parameter value is common to some or all individuals. Common or "shared" parameter values can occur, for example, when several individuals are exposed to radiation from a common environmental source for which the contamination level or exposure rate is uncertain. Uncertainty about a parameter that is common to multiple individuals can induce correlations in the uncertainties of those individuals' dose estimates.

Because each individual has only one uniquely true value of dose, uncertainties in fixed quantities shared among individual members of cohort subgroups are a type of systematic error or bias. The true dose among members of a subgroup may be either higher or lower than the central value. It will not be both higher for some persons and lower for others at the same time. The presence of systematic errors or bias can have a more profound effect on the interpretation of the statistical power and dose-response of an epidemiological study than sources of uncertainty that are random, such as those identified as classical or Berkson errors (Schafer and Gilbert 2006; Carroll et al. 2006).

Shared uncertainties may also affect different cohort subgroups to varying degrees. For example, uncertainty in the source term (or release of radioactivity from a facility) affects nearly all members of an exposed population to the same degree. In other words, if the true release is half the central, or "best estimate," all exposed individuals will be exposed to half of the amount calculated using the central estimate of the release. Likewise, if the true release is twice the central estimate, all exposed individuals would be exposed to twice the amount calculated using the central release estimate. On the other hand, uncertainty in the central estimate of the model parameters representing a transfer of radioactive substance from the cow's diet to fresh milk would be a systematic uncertainty affecting only those members of the cohort who consumed fresh milk. Furthermore, the degree of uncertainty in the central value of this model parameter could vary from dairy to dairy; thus, not every consumer of fresh milk would have a dose that is systematically over- or under- estimated to the same extent. Individuals who consumed low amounts of milk and, thus, had significant doses from other exposure pathways would also be less affected by systematic errors in the central estimate of the transfer of radioactive material from a cow's diet to fresh milk than would an individual who consumed large quantities of milk

and whose total organ dose was effectively dominated by the dose from the air-pasture-cow-milk pathway. Shared or systematic uncertainties are most likely to occur in model parameters that are fixed but imperfectly known quantities. The presence of shared uncertainties resulting from systematic uncertainties in model parameters are addressed using the same framework as that used to separate Type A uncertainties from Type B uncertainties.

Knowledge of the existence and likely magnitude of correlations arising from shared parameters is essential to proper interpretation and use of the dose estimates, especially in the context of epidemiological studies (Stram and Kopecky 2003; Li et al. 2007).

Calculation of Dose

The method being used for the TRDS-2009D dose calculations is relatively simple and can be written as a single equation in three parts as:

$$D_{o,Y,i} = \sum_{y=y_{\min}}^{P \le Y} \left[\sum_{L} M_{y,L,i} \left[\begin{array}{l} \left(\sum_{r} I^{*}_{y,r,L}(\tau_i) DF_{r,o,Y-y}(\tau_i) \right) + A_o D_{Riv,L,y} \left(T_1(\tau_i) + R_{out/Riv,L} \left(T_2(\tau_i) + R_{in/out} T_3(\tau_i) \right) \right) + \\ G_{Sr,L} \delta_y \left\{ \sum_{r} E_{r,y}(\tau_i) DF_{r,o,Y-y}(\tau_i) + A_o(\tau_i) D_{Sr,y} \left[(1 - T_3) + R_{in/out} T_3(\tau_i) \right] \right\} \\ + \sum_{e_i} X_{o,i}(e_i, y, \tau_i) \end{array} \right] \right]$$

Here the upper line in the internal brackets represents the dose from the Techa River, the middle line represents dose from exposure to fallout from the East Urals Radioactive Trace (EURT), and the lower line represents dose from medical x-ray examinations. (Note that doses from ingestion of iodine from Mayak releases are theoretically included in the TRDS, but the parameters will only be calculated and added to the system at a later date). The individual components are:

$D_{o,Y,i}$ = absorbed dose (Gy) in organ o accumulated through calendar year Y to individual i;

Y = the calculational endpoint for a particular individual (can vary according to the analyst's wishes within the range 1950–2015);

b_i = the year of birth of individual i;

y = year of environmental exposure (external irradiation and intake of nuclides). The minimum value of y in the summation is $y_{\min}$ = MAX{1949, b_i, year of first moving to the Techa River or EURT area};

P = the endpoint of external exposure and intake of radionuclides for a particular individual (can vary within the range 1950 – Y, $P \leq Y$).

L = location (settlement) identifier;

$M_{y,L,i}$ = fraction of year y spent in location L by individual i;

r = identifier of ingested radionuclide (^{89}Sr, ^{90}Sr, ^{95}Zr, ^{95}Nb, ^{103}Ru, ^{106}Ru, ^{137}Cs, ^{141}Ce, ^{144}Ce or ^{131}I);

τ_i = $y - b_i$, the age of individual i in year y (years);

$I^*_{y,r,L}$ = intake function (Bq) for year y, radionuclide r, and location L (function of age τ, related to y);

$I^* = I \times \xi_i$, where ξ_i is a modifier predetermined for individual i equal to the village average, IMR_i (individual to model ratio), or HSR_i (household-specific ratio), discussed below;

$DF_{r,o,Y\text{-}y}$ = conversion factor (Gy Bq^{-1}) for dose accumulated in organ o in year Y-y from intake of radionuclide r in year y (function of gender and age, related to y);

Y-y = time since intake, years;

A_o = conversion factor from absorbed dose in air to absorbed dose in organ o (function of age, related to y);

$D_{Riv,L,y}$ = absorbed dose in air near river shoreline at location L received in year y (Gy).

$R_{out/Riv,L}$ = ratio of dose rate in air outdoors at homes to the dose rate by the river at location L;

$R_{in/out}$ = ratio of dose rate in air indoors to that outdoors;

T_1 = time spent on river bank (relative to whole year) (function of age, related to y);

T_2 = time spent outdoors (relative to whole year) (function of age, related to y);

T_3 = time spent indoors (relative to whole year) (function of age, related to y).

$G_{Sr,L}$ = surface deposition (Bq m^{-2}) at location L of ^{90}Sr from fallout from the EURT;

δ_y = 0 or 1 depending on y. For the EURT, $\delta_{1,y} = 0$ for $y < 1957$;

$E_{r,y}$ = intake function (Bq per Bq m^{-2}) for EURT for year y, radionuclide r, (function of age, related to y), further described below;

$D_{r,y}$ = absorbed dose in air (Gy) received in year y per unit surface deposition of ^{90}Sr from fallout from the EURT; and

$X_o(e,y,\tau)$ = absorbed dose to organ o (Gy) from medical examination e in year y for age τ.

The intake function $I_{y,r,L}$ is a complex, time-dependent function derived from a combination of data from tooth beta counting and the whole-body counter. The village-average intake function $I_{y,r,L}$ for each year y is calculated as:

$$I_{y,r,L}(\tau) = I_R^{Sr90} \times \alpha_{\tau,R}^{Sr90} \times f_L^{Sr90} \times R_{y,r/Sr}^{L}, \quad (2)$$

where

I_R^{Sr90} = Annual ^{90}Sr intake for adult residents of the reference settlement (Muslyumovo);

$\alpha_{\tau,R}^{Sr90}$ = Annual ^{90}Sr intake for other age groups relative to that for adults living in the reference settlement;

f_L^{Sr90} = Annual ratio of ^{90}Sr intake for location *L* to ^{90}Sr intake for residents of the reference settlement; and

$R_{y,r/Sr}^{L}$ = Annual ratio of radionuclide (r)-to-^{90}Sr intake for location *L*.

The TRDS calculation of uncertainty will be based on a Monte Carlo approach to implement calculation of the basic dose equation. The required inputs for these analyses have been developed over the course of Project 1.1. The actual results vary depending on the analysis being undertaken, i.e., the specific individual, the particular calculation endpoint year *Y*, organ of interest *o*, and route of exposure (internal or external). Because the epidemiologists have expressed the desire to separately analyze the internal, external, and medical doses, the calculations will be separated into their component parts:

$$D_{o,y,i}^{T,E} = \sum_{y=y_{min}}^{P \le Y} \sum_{L} M_{y,L,i} \left[A_0 D_{Riv,L,y} \left(T_1(\tau_i) + R_{\frac{out}{Riv},i} \left(T_2(\tau_i) + R_{\frac{in}{out}} T_3(\tau_i) \right) \right) \right]$$

$$D_{o,y,i}^{T,I} = \sum_{y=y_{min}}^{P \le Y} \sum_{L} M_{y,L,i} \sum_{r} \left[I_{y,r,L}^{*}(\tau_i) DF_{r,o,Y-y}(\tau_i) \right]$$

$$D_{o,y,i}^{E,E} = \sum_{y=y_{min}}^{P \le Y} \sum_{L} M_{y,L,i} \left[G_{Sr,L} \delta_y \left\{ A_0(\tau_i) D_{Sr,y} \left[(1 - T_3) + R_{\frac{in}{out}} T_3(\tau_i) \right] \right\} \right]$$

$$D_{o,y,i}^{E,I} = \sum_{y=y_{min}}^{P \le Y} \sum_{L} M_{y,L,i} \left[G_{Sr,L} \delta_y \left\{ \sum_{r} E_{r,y}(\tau_i) DF_{r,o,Y-y}(\tau_i) \right\} \right]$$

$$D_{o,y,i}^{X,E} = \sum_{y=y_{min}}^{P \le Y} X_{o,i}(e_i, y, \tau_i)$$

Here, the leading superscripts *T, E*, and *X* symbolize Techa, EURT, and x-ray, respectively, while the trailing superscripts *E* and *I* symbolize Internal and External dose contributions. While this separation into five simpler equations may seem to simplify each individual calculation, it increases the complexity of the calculation process because of the need to ensure the retention of the various inter-individual correlations (shared and Type B parameters) and the intra-individual correlations (autocorrelations) as discussed above.

In the basic equation, the parameters b_i, $y_{\min}$, P, $M_{y,L}$, and τ for each individual come from individual-life-history information and are a series of constants (although there is some uncertainty associated with move dates $M_{y,L}$). All of the other parameter values are either calculated or approximated and have associated uncertainty.

It is possible to calculate a village-average intake function for every member of the ETRC. For about one-quarter of the cohort, an individual dose based on one or more whole-body counter measurements may be estimated. For these individual dose estimates, the general intake function is normalized by the whole-body count(s). The ratio between the generic estimate and the individual estimate is called the Individual to Model Ratio (*IMR*). In addition, for many people, IMR values are available for others within their personal household. These may be used to scale the generic intake function for everyone within the family or household, as the average of the household IMR values. This is called a Household-specific Ratio (*HSR*). Every member of the ETRC has been evaluated and the best type of intake function (that which minimizes the uncertainty based on use of the whole-body counts through Individual-to-model ratios (*IMR*), Household-specific ratios (*HSR*), or village averages) has been assigned (Shagina et al. 2007); these assignments are available in a database linked to the individual identification code. The advantage of the assignment is that a unique uncertainty distribution is associated with each assignment. Because the TRDS-2009D deterministic calculations have already been performed, point estimates of the IMR, HSR, and village-IMR-distributions are already available.

A recent and stable derivation of the key radionuclide intake term $I_{y,r,L}$ is described in detail in Tolstykh et al. (2001) and updated in Tolstyk et al. (2008a). It has a very complex uncertainty structure (Tolstykh et al. 2002; 2008a). The variation of intake levels within a single village and age cohort depends mainly on the source of drinking-water supply. In the TRDS-2009D system, the village-average WBC-determined body burdens of ^{90}Sr are used to derive the

deterministic estimate of accumulated dose. The village average is derived from the entire distribution of measured body burdens of residents of that village. An individual's measurements are used if they are available and appropriate (the *IMR*), if not but the individual has measured relatives in the same household, an average is taken of those (the *HSR*), or if neither are available, then the village average is used. The relation of the actual measurements to the model predictions is described using Individual-to-Model Ratios (*IMR*) (Degteva et al. 1999). For a person of age τ at the beginning of intake and who was measured by WBC at the year t_m, the value of *IMR* is determined as the ratio of an individual-body-burden measurement, $A_{ind}(\tau,t_m)$, to the value derived from the reference model (representing a permanent resident adult in Muslyumovo), $A_{mod}(\tau,t_m)$:

$$IMR = A_{ind}(\tau,t_m)\left[A_{\mathrm{mod}}(\tau,t_m)\right]^{-1}$$

In the case of repeated measurements, the value of *IMR* is determined as the average of all ratios of WBC measurements-to-the respective reference-model values. *IMRs* serve as age- and time-normalized values that permit the analysis of the entire set of individual data on ^{90}Sr in members of the ETRC.

The uncertainty in intake and retention of ^{90}Sr for any one individual for whom a village-average estimate is used is defined by the actual distribution of *IMR* developed for that village (Degteva et al. 1999). The *IMR* includes all the TRDS-2009D parameters that go into estimation of term $I_{y,r,L}$, except the location factors f_L. As defined and presented in Degteva et al. (1999), the *IMR* is the ratio of the measurement for a specific village to the *prediction made as if that individual lived in Muslyumovo*. As a result, if the *IMR*s are used to estimate the intake, it is not necessary to adjust the basic intake function for the location because this is already accounted for in the IMR estimation. (In essence, the term is used to account for inter-village differences in water concentration, and this same term would then be divided back out when assigning the generic intake.)

The normalized *IMRs* are time-integrated quantities, in that they reflect the deviation of total lifetime intake and retention from that predicted by the TRDS environmental and exposure models. However, it is reasonable to assume that particular individuals would have similar behavior from one year to the next, and that the inter-annual variation is captured in the total normalized *IMR*. Thus, the distribution of normalized *IMRs* for each village can be used to estimate the annual distribution of intakes and retentions for residents of that village. Because of

these considerations, it is not necessary to model explicitly the various components of drinking-water source, diet, uptake, or metabolism that go into estimation of the radionuclide-intake term, $I_{y,r,L}$, and this greatly simplifies the uncertainty analyses. The distribution shape and range of the term $I_{y,r,L}$ is defined for each village by the village-specific normalized *IMR*. The dose to any one individual can vary an order of magnitude up or down from the average model prediction, when based on the village-average values.

Dose-conversion factors, $DF_{r,o,Y\text{-}y}$, are calculated using biokinetic models, and their uncertainties are determined mainly by the variability of metabolic parameters (Shagina et al. 2000). However, for ^{90}Sr, the individual variability in uptake and metabolism is actually captured in the *IMR* values, because the *IMR's* reflect not only intake but also long-term retention. The remaining uncertainties in the dosimetric model are embodied within the specific effective energy quantity and are associated mainly with variations in masses, shapes and locations of the organ and tissue of the human body and with oversimplifications of the representations of certain complex anatomical structures in the body when calculating the energy deposition (NCRP 1998). Thus, the uncertainties in the dose-conversion component for ^{90}Sr and ^{89}Sr are relatively low. The uncertainties in the dose-conversion factors for other radionuclides are larger, reflecting the lack of available measurements and the potential for individual variations in uptake and retention. Because individual variations in uptake and retention will vary less from year to year than the variation among individuals, the dose-factor variability is held constant from year to year for a single realization of the dose estimate and only varied for additional realizations.

One additional uncertainty term is needed for the non-^{90}Sr radionuclides to address the ratio of intakes of these nuclides to ^{90}Sr. This is the term $R^{L}_{y,r/Sr}$, the annual ratio of nuclide-to-^{90}Sr intake. Because the intakes were primarily from drinking water, the intakes are proportional to the estimated concentrations of these radionuclides in river water (with the exception of ^{137}Cs). These ratios are currently estimated based on the results of a Techa River Model. Thus, uncertainties in intake are directly proportional to uncertainties in predicted concentrations in river water. Based on the data presented in the two-compartment Techa River transport model, the predicted concentrations could vary by up to 50% with different selection of transport parameters based on available data. Sensitivity analyses for residents of Muslyumovo indicate that this uncertainty contributes very little to the total uncertainty for dose to red bone marrow,

because the internal doses are dominated by the contribution from ^{90}Sr. The uncertainties in this term are relatively more important for organs of the gastrointestinal tract, because more of the dose resulted from these other radionuclides. (A new approach has been used in TRDS-2009D for intakes of ^{137}Cs because of the unique contribution from consumption via cow's milk (Tolstykh et al. 2008b). This adds a degree of complexity to the $R^{L}_{y,r/Sr}$ term for ^{137}Cs.)

Recent rework of the Techa River source term by a combined team of URCRM, Mayak, and US collaborators has greatly refined the temporal resolution of the source term [Degteva et al. 2008]. Additional modeling has been performed for the period 1951-1952 to adequately describe the dynamics of the water concentration of the various radionuclides. As a result, the various parameters described herein have been refined to shorter time periods – into "time slices" of one month for this period. To accommodate this new information, it is necessary to interpret the equations above with the annual summations replaced by monthly ones for the years 1950 and 1951. This increases the database size, but not the overall approach.

In a similar manner, intake functions have been developed for exposures to the EURT fallout. Data Directories of ^{90}Sr-contamination density of Urals settlements ($G_{Sr,L}$) were created (Tolstykh et al. 2006) with an evaluation of existing data on radionuclide contents in food and the human body that supported development of the necessary input parameters on time- and location-dependent intake rates of radionuclides (Tolstykh et al. 2006). The basic approach considered by Tolstykh et al. (2006) for the reconstruction of internal doses employ conversion factors $E_{r,y}$, that is, dose per unit ground deposition (Gy per kBq m^{-2}). The approach is based upon measurements of radionuclides in local foodstuffs and humans. For the purposes of the EURT analysis we consider the intake of long-lived ^{90}Sr up to 1980, the intake of short-lived radionuclides essentially ended after 1959. For external dose calculations, dose rates in air per unit-deposition density of ^{90}Sr for the EURT area $D_{r,y}$ derived by Vorobiova et al. (2006) are used. External dose accumulation is considered for only the two first years after contamination, because dose-rate values decreased rapidly due to radioactive decay of the short-lived radionuclides. Other parameters of external exposure (such as typical life patterns and shielding, as well as conversion factors from dose in air to dose in organs) are the same as used for Techa River exposure.

The conversion factor from absorbed dose in air to absorbed dose in organ o, A_o, is a mild function of radiation energy. However, there is a large plateau in the energy-dependent response

between about 0.08 and 1.3 MeV (Eckerman and Ryman 1993), the energies of most interest for the radionuclides discharged to the Techa River. There is also a minor variation as a function of body mass for various ages. This parameter will be slightly variable for individuals of different weights.

The terms T_1, T_2, and T_3, while ideally coming from individual data, are currently assigned generic values, depending on the age of the individual in year y. A discussion of various lifestyle surveys is presented in Vorobiova et al. (1999). They are allowed to change from year to year to account for individual circumstances.

The external dose rates, $D_{Riv,L,y}$ and $D_{r,y,,}$ are derived from measurements, or alternatively, from the radionuclide contents of sediment as calculated from the Techa River transport model. Values for doses at housing locations outdoors and indoors are derived from $D_{Riv,L,y}$ using river-bank-to-residence-area dose-rate ratios and indoor-to-outdoor dose-rate ratios. Extensive efforts have been made to identify the exact house in which each individual lives, thus allowing detailed specification of this distance for each subject – however, this information is not yet included in TRDS-2009D and village averages (with their associated larger uncertainties) are used.

Doses to the cohort members from medical x-ray examinations have been estimated by Degteva et al. (2008). A detailed record of each exposure exists; in essence, x-ray doses are added to the individual's appropriate annual organ dose summary at the proper time. These individual values of $X_o(e,y)$ have an associated uncertainty found in the database.

The TRDS Databases

The TRDS-2009D relies on extensive databases in order to compute the doses for each cohort member. These databases were calculated and modeled during the course of research performed in the framework of Project 1.1. A full description of the models and data sets used for the production of the TRDS databases is provided in Degteva et al. (2009). Separate databases were created for the first two years of exposure (1950–1951) in order to take into account rapid changes in the source-term parameters. Also, separate databases were established for the village of Metlino, which is a special case because this settlement was the closest to the site of radioactive release and the time pattern of the intake for the residents of Metlino differed from that in other villages. In addition, special databases were established to give definition to

^{137}Cs intake. And also, gender-specific databases were prepared for ^{90}Sr and ^{137}Cs internal dose coefficients. Additional parameters characterizing exposure in the EURT were introduced. These data were described in detail in our reports for Milestones 11–13 (Tolstykh et al. 2006; Vorobiova et al. 2006). Data Directories of ^{90}Sr-contamination density of Urals settlements were created through an evaluation of existing data on radionuclide contents in food and human tissues that supported development of the necessary input parameters for time- and location-dependent intake rates of radionuclides (Tolstykh et al. 2006). The approach for the reconstruction of internal doses employs conversion factors based on the dose per unit ground deposition. The approach is based upon measurements of radionuclides in local foodstuffs and humans. For external dose calculations, dose rates in air per unit-deposition density of ^{90}Sr for the EURT area derived by Vorobiova et al. (2006) are used.

The database entries are summarized in Table 1. These components of the database essentially provide the input data from which the dosimetry system runs. The first three files are codes for specific villages in the Techa River and EURT areas. The next four files are lists of external dose rate for each of these villages by year (VURS is the Russian acronym for EURT). The file REGIME is the times T_1, T_2, and T_3 for each age group. The file DOS-F is the internal dose conversion factors by age for each radionuclide. The series of files designated REPER contain the generalized intake function $I_{y,Sr}$ for the reference villages of Metlino and Muslyumovo. The series of files designated CHILD are the function $\alpha_{\tau,R}^{Sr90}$ (note the specialized file for ^{137}Cs). The series of files designated NUCL_STC are the functions f_L^{Sr90} and $R_{y,r/Sr}^{L}$ for each village; the series Cs137-STC are the equivalent functions for ^{137}Cs. (These incorporate the evaluated source term and river transport information.) The remaining files are internal dose conversion factor annual increments by sex and age-at-intake.

Because of the numerous files and varied information, it is considered to be easier to develop "uncertainty factors" for most of the parameters rather than provide uncertainty ranges for each required value of each individual instance of each parameter.

Table 1. TRDS-2009D databases.

System DB name	Form	Content
TECH_LIST	File	List of Techa settlements and codes
VURS_LIST	File	List of non-evacuated EURT settlements with codes and ^{90}Sr-contamination densities
VURS_EVAC	File	List of evacuated EURT settlements with codes and ^{90}Sr-contamination densities
TECH_EXT	Library of 41 files	For each of 41 settlements annual doses in air near the river and within residence areas (since 1952)
TECH_EXT_50-51	Library of 41 files	For each of 41 settlements monthly doses in air near the river and within residence areas in 1950-1951
VURS_EXT	File	Annual doses in air per unit ^{90}Sr-contamination density
VURS_EXT_57-58	File	Monthly doses in air per unit ^{90}Sr-contamination density
REGIME	File	Periods of time spent near the river and within residence area (outdoors and indoors), for age groups
DOS_F	Library of 23 files	For each of 23 organs, age-dependent dose-conversion factors (air-to-organ)
REPER1	File	Annual ^{90}Sr intakes for adult residents of the reference settlement Muslyumovo since 1952
REPER1_50-51	File	Monthly ^{90}Sr intakes for adult residents of the reference settlement Muslyumovo in 1950-1951
REPER_701	File	Annual ^{90}Sr intakes for adults in Metlino since 1952
REPER_701_50-51	File	Monthly ^{90}Sr intakes for adults in Metlino in 1950-1951
REPER2	File	Annual ^{90}Sr intakes normalized per unit ^{90}Sr-soil contamination density (Ci km^{-2}) for residents of EURT settlements since 1959
REPER2_57-58	File	Monthly ^{90}Sr intakes normalized per unit ^{90}Sr-soil contamination density (Ci km^{-2}) for residents of EURT settlements from September 1957 until January 1959
CHILD	File	Annual relative ^{90}Sr intakes for children who lived in the reference settlement
CHILD_701	File	Annual relative ^{90}Sr intakes for children who lived in Metlino
CHILD_Cs137	File	Ratios intake of water and milk by children to adults
NUCL_STC	Library of 41 files	Annual ratios of intake of ^{90}Sr to that for reference settlement and intake of a nuclide-to-^{90}Sr for each of 41 settlements since 1952
NUCL_STC_50-51	Library of 41 files	Monthly ratios of intake of ^{90}Sr to that for reference settlement and intake of a nuclide-to-^{90}Sr for each of 41 settlements in 1950-1951
Cs137_STC	Library of 41 files	Annual intakes of ^{137}Cs with water and milk for each of 41 settlements since 1952
Cs137_STC_50-51	Library of 41 files	Monthly intakes of ^{137}Cs with water and milk for each of 41 settlements in 1950-1951
Sr90M, Sr90F, Sr89M, Sr89F	Four libraries, each of 56 files	Age- and time-dependent dose-conversion factors for different organs for males and females
Cs137M, Cs137F	Two libraries, each of 23 files	Age- and time-dependent dose-conversion factors for different organs for males and females
Ru103, Ru106, Zr95, Nb95, Ce144, Ce141, Ba140, I131	8 libraries, each of 23 files	Age- and time-dependent dose conversion factors for different organs

Assignment of Uncertainty Types

The approach to uncertainty analysis is use of Monte Carlo replications of the basic model, using uncertain input parameters. As noted in Stram and Kopecky (2003), the assumption that each replication of possible dose is a sample from the distribution of possible dose for the study subjects is based upon the adoption of what is known as a subjective Bayesian view of the meaning of incomplete information regarding the determinants of dose. This simply means that parameters in the dosimetry system that are incompletely known are assumed to be random quantities, which follow a subjective probability distribution, agreed upon by the experts who developed the system, conditional upon whatever information was available to the experts.

The dose-response use of the reconstructed doses leads to the need to differentiate uncertainties in the dose estimates from two separate sources; shared versus unshared uncertain parameters (of which the Type A/Type B difference is a more global description), and, more generally, classical versus Berkson error structures.

Careful design of the dosimetry system will address the issue of shared versus unshared uncertain parameters. Development is planned of realizations of dose such that the same vectors of "environmental parameters" are used for each individual at a particular location and time. (This approach was used, for example, in the Hanford Environmental Dose Reconstruction project doses supplied to the Hanford Thyroid Disease Study).

Again, as noted by Stram and Kopecky (2003), "Berkson error models are realistic only if the characteristics of the study population are considered… One cannot build a dosimetry system that will provide Berkson errors for a single subject independent of the population in which the dosimetry system is applied (at least not if errors in subject-specific input data are to be adequately dealt with)." In the TRDS, it is assumed that "environmental parameters" (generally those that are shared by all individuals across a particular dose realization) have Berkson structure, and those that are specific to the individual (if known) have a classical uncertainty structure. However, in the TRDS, generic models are frequently used to fill in for lack of information about specific individuals; in this case the parameter must be considered to be Berkson in nature.

An assignment of the uncertainty structure to the various TRDS parameters is given in Table 2. Note that "shared/unshared" and "Type A/Type B" are listed separately, but will be handled in a similar fashion.

Table 2. Assignment of TRDS parameter uncertainty structures.

Parameter	Definition	Sharing	Structure	A/B
Constants - used to define individual calculation				
Y	The calculational endpoint for a particular individual (can vary within the range 1950–2005)	Definition of case	Constant	
y	Year of environmental exposure (external irradiation and intake of nuclides)	Definition of case	Constant	
P	The endpoint of external exposure and intake of radionuclides for a particular individual (can vary within the range 1950–Y, P≤Y)	Definition of case	Constant	
L	River-location (village) identifier	Definition of case	Constant	
r	Identifier of ingested radionuclide	Definition of case	Constant	
Y-y	Time since intake, years	Definition of case	Constant	
e	Number of x-ray exposures per individual	Definition of case	Constant	
b	Individual birthday	Definition of case	Constant	
Common to internal and external				
$M_{y,L}$	Fraction of year y spent in location L;	Individual/unshared	Classical	B
$G_{Sr,L}$	Deposition of ^{90}Sr at location L for EURT fallout (Bq m^{-2})	Shared within village	Classical	AB
Internal dose parameters				
$I_{y,r,L}$	Intake function (Bq year^{-1}) for year y, radionuclide r, and location L (function of age, related to y)	Product of the following parameters:		
I_R^{Sr90}	Annual ^{90}Sr intake for adult residents of the reference settlement (Muslyumovo)	Shared	Berkson	AB
$\alpha_{Age,R}^{Sr90}$	Annual ^{90}Sr intake for other age groups relative to that for adults living in the reference settlement	Shared within ages	Berkson	AB
f_L^{Sr90}	Annual ratio of ^{90}Sr intake for location L to ^{90}Sr intake for residents of the reference settlement	Shared within village	Berkson	AB
$R_{y,r/Sr}^L$	Annual ratio of radionuclide (r)-to-^{90}Sr intake for location L	Shared within village; along river	Berkson	BB
IMR	Individual-to-Model Ratio; the deviation of an individual for the $I_{y,r,L}$ (based on WBC)	Unshared if available	Classical	A

Parameter	Definition	Sharing	Structure	A/B
IMR	This is also used for those without WBC measurements; set to 1.0 with a wide uncertainty bound based on village IMRs	Shared if not available for ind.	Berkson	B
HSR	Household Specific Ratio - the average of IMR's for one household (based on WBC)	Shared within household	Classical	AB
$E_{r,L}$	Normalized intake function (Bq year^{-1} per Bq m^{-2}) for EURT for radionuclide *r* at location *L*	Shared within village w/ ind. variation	Berkson	B
$A_{ind}(\tau,t_m)$	Individual whole-body counter measurement made at age τ and time t_m	Unshared	Classical	A
$DF_{r,o,Y-y}$	Conversion factor (Gy Bq^{-1}) for dose accumulated in organ *o* in year *Y-y* from intake of radionuclide *r* in year *y* (function of age, related to *y*)	Shared with autocorrelation	Berkson	AB
External dose parameters				
A_o	Conversion factor from absorbed dose in air to absorbed dose in organ *o* (function of age, related to *y*)	Shared	Berkson	AB
$D_{Riv,L,y}$	Dose rate in air near river shoreline at location *L* in year *y* (Gy year^{-1})	Shared within village with autocorrelation	Classical after 1950	B
$D_{r,y}$	Normalized dose rate in air outdoors in year *y* (Gy year $^{-1}$ per Bq m^{-2}) from EURT fallout	Shared within village with autocorrelation	Berkson	B
$R_{out/Riv,L}$	Bank to residence ratio (function of distance of individual's home from river)	Unshared	Classical	A
$R_{in/out}$	Indoor/Outdoor ratio (function of building type)	Unshared	Berkson	A
T_1	Time spent on river bank (relative to whole year) (function of age, related to *y*)	Unshared	Berkson	A
T_2	Time spent outdoors (relative to whole year) (function of age, related to *y*)	Unshared	Berkson	A
T_3	Time spent indoors (relative to whole year) (function of age, related to *y*)	Unshared	Berkson	A
$X_o(e,y,\tau)$	Dose (Gy) from medical exposure *e* in year *y*	Shared by procedure type	Berkson	AB

Discussion of Parameter Uncertainty Assignments

The parameters that describe the contamination of the environment in which the subjects live generally have a shared uncertainty. Only inputs that are exclusive to a single individual are unshared; in the TRDS system, there are actually quite few of these. In general, it is assumed that inputs that are themselves the products of models (such as the dose conversion factors) have a shared Berkson uncertainty structure, because they are not really specific to any one individual even if an individual modifying factor is applied (because the individual modifiers are generic to all individuals of this type). Shared and Type B parameters are both selected by realization; unshared and Type A parameters are allowed to vary within a realization. Type AB parameters have components of each type (and the one Type BB is similar).

Many inputs are used to define a particular calculation; these are assumed to be constant and invariant. The uncertain, non-control parameter $M_{y,L}$ is common to both internal and external dose estimates. This parameter is derived from the individual's residence history, and should have a minor random uncertainty based on either individual recall or interpretation of tax records, etc. Calculations of radionuclide intake are highly dependent on a series of complex model calculations that are independent of any one individual; thus they are both shared and Berkson, because they are assigned to categories of individuals. The conversion factors from intake to dose are also derived from standardized models, some developed specifically for TRDS (such as the strontium metabolism model); thus, these also have shared Berkson uncertainties. The external doses for the first few years are based on models of source term, radionuclide transport in the river, and dose-rate-per-unit-deposition; thus for this period, the resulting parameters have shared Berkson uncertainties. In the later years, the dose rates in each village are based on actual measurements, and the dose rates at the river's edge have classical uncertainty structure. The estimation of dose at each individual's house is, however, based on a radiation transport model and probably has both an unshared classical (distance measurement related) and shared Berkson (model assignment) component. The dose rates within the homes are based on an assigned shielding factor, the distribution of which will again have shared Berkson structure. The exposure times on the river, in the neighborhoods, and in the houses could have unshared classical structure if it were based on individual questionnaire responses, but will usually be assigned from survey results and thus be shared and probably Berkson.

There are correlations between some variables. The terms $R^{L}_{y,r/Sr}$ in the internal dose and D_{riv} in the external dose are the terms in which the uncertain radionuclide "source term" enters the calculation. The structure of the computational model will account for this connection.

The IMR parameter is crucial to both the individuals who have a WBC and those who do not. For those with WBCs, this parameter is what differentiates them from the village, with a fairly small uncertainty based upon the number of WBCs. For those who do not have a WBC, the uncertainty in intake will be as wide as that of the entire village, because the village distribution is used.

Selection of Specific Parameter Distributions

Each of the parameters in Table 2 requires basic data and definition of one (or more) uncertainty distributions. The current sources of the individual parameters are given in the final column of Table 3; initial selections of distribution type and parameterization are provided in the other columns. Some discussion of the selection for each parameter is given here. Note that some parameters have both Type A and Type B (or shared and unshared) components. This means that there are overall group distributions with associated individual variabilities.

Note that, for most of the parameters, the "central estimate" of the selected distribution is being taken as the deterministic value derived for TRDS-2009D. For computational reasons, most distributions will be assigned to a multiplicative "uncertainty factor". The uncertainty factors are being defined as dimensionless adjustments with a mean of 1.0 and a distributional type and associated parameters such as a standard deviation or geometric standard deviation. It is important that the mean of the uncertainty factor be 1.0, rather than the median (or other central tendency); use of any other value would lead to a bias in the final result. Thus, while the usual tendency is to describe a lognormal distribution via a median and GSD, herein we use a mean and GSD, realizing that the median may appear to be significantly different from 1.

In the basic equation, the term $M_{y,L}$ comes from individual-life-history information and is a series of constants. However, there may be some uncertainty about dates associated with moving from one location to another; therefore, a uniform distribution allowing a range of up to 3 months on either side of the reported move date has been added.

Table 3. Initial Selection of Parameter Distributions

Parameter	Unshared Parameters (Uncertainty due to stochastic variation)			Shared Parameters (Uncertainty due to lack of knowledge)			Data Source
Constants - used to define individual calculation	Distribution Type	Mean or Geo. Mean	Std. Dev. Or GSD	Distribution Type	Mean or Geo. Mean	Std. Dev, GSD, or range	
Y							
y							
P							
L							
r							
Y-y							
e							
b							
Common to internal and external							
$M_{y,L}$	Uniform	± 3 mo in year of move					Individual: Sample.xls
$G_{Sr,L}$	Normal	1	0.1	Normal	1	0.05	File EURT_List.xls
Internal dose parameters							
$I_{y,r,L}$							
I_R^{Sr90}	Normal	1	0.25	Custom		Created from all village IMRs	File Repers.xls

Parameter	Unshared Parameters (Uncertainty due to stochastic variation)			Shared Parameters (Uncertainty due to lack of knowledge)			Data Source
Constants - used to define individual calculation	Distribution Type	Mean or Geo. Mean	Std. Dev. Or GSD	Distribution Type	Mean or Geo. Mean	Std. Dev, GSD, or range	
$\alpha^{Sr\,90}_{Age,R}$	Normal	1	0.2	Normal	1	0.1	File Child.xls
$f_L^{Sr\,90}$	Not used directly in the modeling: incorporated in IMR and HSR						File IMRs.xls
$R^L_{y,r/Sr}$				Lognormal Uniform	0.7864 1	2 0.5-1.0	File NUCL_STC.xls
IMR	Normal	1	Individual				"Delta" in Sample.xls
IMR				Custom	1	Based on Village IMR distribution, about 0.1 to 10 range	File IMRs.xls
HSR	Normal	1	0.1	Custom	1	Individual	"90CI" in Sample.xls
$E_{r,L}$	Lognormal	0.7864	2	Lognormal	0.5469	3	File repers.xls
$A_{ind}(\tau, t_m)$	Not used directly in the modeling: incorporated in IMR and HSR						
$DF_{r,o,Y\text{-}y}$	Lognormal	0.6572; 0.9754	2.5; 1.25 for Sr90	Normal	1	0.1	Series of files, "nuclide".xls
External dose parameters							
A_o	Uniform	1	0.9-1.1	Uniform	1	0.9-1.1	See Table 3
$D_{Riv,L,y}$				Custom	Village ranges		Files in directory

Parameter	Unshared Parameters (Uncertainty due to stochastic variation)			Shared Parameters (Uncertainty due to lack of knowledge)			
Constants - used to define individual calculation	Distribution Type	Mean or Geo. Mean	Std. Dev. Or GSD	Distribution Type	Mean or Geo. Mean	Std. Dev, GSD, or range	Data Source
							Tech_Ext
$D_{r,y}$				Uniform	1	0.9-1.1	See Table 5
$R_{out/Riv,L}$	Uniform	Village min to max					DBF files in directory Tech_Ext
$R_{in/out}$	Uniform	0.45	0.125-1.0				Vorobiova et al. 1999 p21
T_1	Uniform	1	0.7-1.3				Tabulated Vorobiova et al. 1999 (see Table 4)
T_2	Uniform	1	0.7-1.3				Tabulated Vorobiova et al. 1999 (see Table 4)
T_3	Uniform	1	0.7-1.3				Tabulated Vorobiova et al. 1999 (see Table 4)
$X_o(e,y,\tau)$	Uniform	Procedure min to max		Custom by procedure type			Individual; Sample.xls - but need more U/S info

The ground deposition of radionuclides from the EURT release, $G_{Sr,L}$, is based upon measurements made in each village. It is an overall average, there should be a small uncertainty based upon measurement error, so a shared value of a range of 5% has been assumed. The same value is assigned to each individual in the village. For each individual, some minor variability within the village is also likely, so an additional individual range of 10% is added.

The intake of radionuclides is described by the intake function (Tolstykh et al. 2008a). The intake function is a series of annual intake estimates for a reference person, derived through mathematical modeling based upon a large collection of tooth-beta counts, whole-body counts, autopsy evaluations, and other observations, along with some assumptions about the time-history of the releases into the Techa River. Recent significant revisions to the intake function, accounting for a large number of recent discoveries and other modifications, resulted in year-by-year changes from earlier estimates of about 25% for most years, with the overall intake (the area under the curve) changing by less than this amount. It is believed by the author that future revisions will, if anything, be smaller than these. Therefore, an uncertainty on the reference intakes represented by the generic curve is subjectively set to about 25%.

Individuals for whom an informative WBC is available have assigned an *IMR*. This is a normalizing value for that individual that adjusts the generic intake function and that accounts for all of their various sources of intake of radionuclides. Presence of an *IMR* greatly reduces the uncertainty in the individual's intake. However, there is some degree of measurement uncertainty associated with the WBC leading to the *IMR*, as well as individual peculiarities in biokinetics and other factors. As a result, there is some variation of their measurements around the idealized retention function. A measure of this is provided by the variance of their measurements about the predicted value. This variance is used as an uncertainty of the individual's *IMR*. The resulting correction factor is applied across years to account for autocorrelations.

Individuals without informative WBC data, but for whom measurements of individuals in the same household are available, have been assigned an *HSR*. It is assumed that intakes of individuals in the same household will be similar (to the extent predicted accounting for age and sex deviations). The aggregation of the householders' *IMR* values into an *HSR* results in a variance, which is used as a descriptor of the uncertainty in the *HSR* used. The resulting correction factor is applied across years to account for autocorrelations.

Individuals without an informative WBC, for whom no *IMR* or *HSR* is available, are assigned the village average intake (appropriately adjusted for age). The uncertainty in this village average is dealt with in the same manner as was done for the TRDS-2000 uncertainties (Napier et al. 2001). The uncertainty in intake and retention of ^{90}Sr for any one individual within the village is defined by the actual distribution of *IMR* developed for that village (Degteva et al. 1999). The *IMR* distribution includes all the TRDS-2009D parameters that go into estimation of term $I_{y,r,L}$ (except the location factors f_L). The normalized *IMR's* are time-integrated quantities, in that they reflect the deviation of total lifetime intake and retention from that predicted by the TRDS environmental and exposure models. However, it is reasonable to assume that particular individuals would have similar behavior from one year to the next, and that the inter-annual variation is captured in the total normalized *IMR*. Thus, the distribution of normalized *IMR's* for each village can be used to estimate the annual distribution of intakes and retentions for unmeasured residents of that village. The distribution shape and range of the term $I_{y,r,L}$ is defined for each village by the village-specific normalized *IMR*. However, because the *IMR*'s used include all permanent residents of each village (including children), there is some residual uncertainty associated with not de-convoluting the term $\alpha_{Age,R}^{Sr90}$ included in the children's values (see below). This is dealt with by artificially enhancing the distribution with a multiplier of an additional 10%. The selected value of *IMR* is applied across years to account for autocorrelations.

If an individual was under the age of 10 years during any part of the exposure, and does not have an IMR assigned, the generic intake function is modified by the term $\alpha_{Age,R}^{Sr90}$ to account for reduced intakes during younger years. This term consists of a table of age-and-year corrections to the generic intake rates. It is assumed that the derivation of these correction factors on average has a normal uncertainty distribution with a standard deviation of about 10% (i.e., the tabulated value is within about 10% of the "true" average), and that individuals can also vary about the average by an additional 20%.

Intakes for individuals who did not live in the reference villages of Metlino or Muslyumovo are based on the generic intake function corrected for the village in which they did live. The correction parameter f_L is the average of the *IMR* values for permanent residents of the specific village. Because the dose calculations use the *IMR, HSR*, or village-average *IMR*, this parameter is not directly used in the uncertainty calculations. (Its use is more conceptual than functional.)

The parameter $R^{L}_{y,r/Sr}$ is a combination of the outputs of the source term analysis and transport modeling in the Techa River (Degteva et al. 2008; Vorobiova and Degteva 1999). Thus, it has components related to both the release of each radionuclide as well as village differences related to transport. Source term estimates have been made by several teams. Radioactive discharges into the Techa River were evaluated on the basis of radioecological monitoring and dosimetric modeling data in a joint project supported by the International Science and Technology Center (ISTC). The research team included staff of the URCRM, the Mayak Production Association, the Russian Federal Nuclear Center–All Russia Scientific Research Institute of Technical Physics (RFNC–VNIITF), the Institute of Plant and Animal Ecology of the Urals Division of the Russian Academy of Sciences, and the U.S. collaborators of Project 1.1, plus Dr. Owen Hoffman of SENES, Oak Ridge. Summary and analysis of the results of ISTC Project No.2841 on *Reconstruction of the Techa River source term* were presented in Degteva et al. (2008). The interpretation differed for some radionuclides by up to factors of two from those of Glagolenko et al. (2006; 2008). Thus, uncertainties on the quantities released into the Techa River are assumed to be within factors of two (a lognormal distribution with a mean of 1.0 with a GSD set such that the 90% confidence interval is from 0.5 to 2 times the median). This is applied independently to each radionuclide and held constant throughout the entire environmental realization. The transport within the Techa River of the released materials is also subject to uncertainty; the transport is currently modeled with a simple double-exponential fit of historical observations (Vorobiova and Degteva 1999). Uncertainties in intake are directly proportional to uncertainties in predicted concentrations in river water. Based on the data presented in Vorobiova and Degteva (1999), the predicted concentrations could vary by up to 50% with different selection of transport parameters based on available data. For the uncertainty analysis, a uniform distribution between 0.5 and 1.5 (as normalized to a value of 1.0) has been used for each radionuclide. This is also applied independently to each radionuclide and held constant for all individuals within a given village. This is an area for additional evaluation in the future. The resulting modification per realization is the product of these two uncertainty factors.

The EURT intake function $E_{r,L}$ was described by Tolstykh et al. (2006). Tolstykh et al. (2006) report that the range of intakes of 90Sr within a given village could be broad (approaching one order of magnitude above and below the central values), that there was substantial heterogeneity of the fallout within villages, and that estimates based upon food-chain

modeling were up to 3 times greater than those based on human measurements. In addition, individuals within a village could have also had individual peculiarities of diet that significantly impacted their intakes. Therefore, the overall generic function is assigned a lognormal distribution with a geometric standard deviation (GSD) of 3, used across time periods within one village, and individual variation is superimposed via a distribution with a GSD of 2.

Dose-conversion factors, $DF_{r,o,Y\text{-}y}$, are calculated using biokinetic models, and their uncertainties are determined mainly by the variability of metabolic parameters (Shagina et al. 2000). However, for ^{90}Sr, the individual variability in uptake and metabolism is actually captured in the *IMR* values, because the *IMR'*s reflect not only intake but also long-term retention. The remaining uncertainties in the dosimetric model are embodied within the specific effective energy quantity and are associated mainly with variations in masses, shapes and locations of the organ and tissue of the human body and with oversimplifications of the representations of certain complex anatomical structures in the body when calculating the energy deposition (NCRP 1998). Thus, the uncertainty in the dose-conversion component for ^{90}Sr is relatively low. For this assessment, its variability has been approximated as a lognormal distribution with a geometric standard deviation (GSD) of 1.25. The uncertainties in the dose-conversion factors for other radionuclides are larger, reflecting the lack of available measurements and the potential for individual variations in uptake and retention. Because of their derivation for radiation protection purposes, there is some uncertainty about the central value of the tabulated dose factors; this is a shared uncertainty. This is addressed as a minor normal distribution with a small standard deviation of about 0.1 for non-Sr radionuclides. The individual variations for the dose factors for other radionuclides can be considered as lognormal with GSD's of about 2.5 (see NCRP 2009). Because individual variations in uptake and retention will vary less from year to year than the variation among individuals, the dose-factor variability is held constant from year to year for an individual for a single realization of the dose estimate to account for this autocorrelation and only varied for additional realizations.

The conversion factor from absorbed dose in air to absorbed dose in organ o, A_o, is a mild function of radiation energy. However, there is a large plateau in the energy-dependent response between about 0.08 and 1.3 MeV (Eckerman and Ryman 1993; Petoussi et al. 1991), the energies of most interest for the radionuclides discharged to the Techa River. Table 4 exemplifies the data on conversion factors for different age groups. Because of the minor variability of this

factor, a shared uncertainty of 10% is assumed. There is also a minor variation as a function of body mass for various ages; this parameter will be slightly variable for individuals of different weights; the distribution for individuals is assumed to vary by about 10%.

Model behavior factors T_1, T_2, T_3 – the fraction of time spent in different locations for different age groups – were derived from observational data (Vorobiova et al. 1999) from the 1950s of typical life-style patterns for different age groups of Techa Riverside residents. Table 5 exemplifies the data on behavior factors for different age groups. The terms T_1, T_2, and T_3 are currently assigned generic values, depending on the age of the individual in year y. These times are assumed to vary by up to 30% for individuals, constrained to the total hr/year. They are allowed to change from year to year to account for individual circumstances.

Table 4. Conversion-factor ratios: Absorbed dose in organ-to-absorbed dose in air.

Organ	Age-dependent dose-conversion factors, Gy Gy^{-1}		
	<7 y	7–17 y	≥17 y
Red bone marrow	0.85	0.76	0.73
Bone surface	1.37	1.22	1.18
Large Intestinal wall	0.75	0.67	0.64
Small Intestinal wall	0.73	0.65	0.62
Stomach wall	0.78	0.69	0.66
Testes	0.94	0.83	0.80
Ovaries	0.71	0.63	0.61
Uterus	0.72	0.64	0.62

Table 5. Typical life patterns for different age groups of the Techa Riverside residents.

Age group, years	Period of time spent at specified site, hours per year			
	Shoreline (summer time)	Residence area (outdoors)	Residence area (indoors)	Far from the river (uncontaminated territory)
<7	45	2235	6480	0
7–15	150	2130	5760	720
16–59	150	1410	3960	3240
≥60	150	2490	6120	0

The external dose rates by the Techa River, $D_{Riv,L,y}$, are derived from measurements, or alternatively, from the radionuclide contents of sediment as calculated from the model of Vorobiova and Degteva (1999). The ranges of modeled or measured dose rates near the riverbank are presented for each village in the Appendix to Vorobiova et al. (1999). These ranges may be used to define the distribution of $D_{Riv,L,y.}$ Values for $R_{out/Riv,L}$ and $R_{in/out}$ are derived from $D_{Riv,L,y}$ using river-bank-to-residence-area dose-rate ratios and indoor-to-outdoor dose-rate ratios, ranges for which are also presented in the Appendix to Vorobiova et al. (1999). Because the actual distances of individual residences from the river are largely known but not yet included in the databases, it is necessary to assume that specific individuals could live in any house. Therefore, the bank-to-residence and indoor-to-outdoor dose-rate ratios are treated as uniform distributions between the lower and upper observed bounds within each village. These are held constant from year to year within a realization. These are currently treated as Type A parameters; when the upgrades to the external dose algorithm are completed they will become Type AB parameters.

The external dose rates per unit ^{90}Sr deposition, $D_{r,y}$, for the EURT area derived by Vorobiova et al. (2006). Dose accumulation is considered for only two calendar years after contamination, because dose-rate values decreased rapidly due to radioactive decay of the short-lived radionuclides (Table 6). These have only a minor variability (most uncertainty is associated with the actual deposition or exposure circumstances, treated above). A uniform range of plus-or-minus 10% is assigned.

The doses from medical procedures are a special case; these doses are pre-calculated using a separate computer program and the results stored in a database (Degteva et al. 2007). There are several types of medical x-ray, including various radiography and fluorography procedures. For the fluorography procedures, the imaging time is an important variable; for the radiography, the number of possible exposures is potentially important. Because these types of procedure have been performed at known times for known individuals, the accumulated dose is simply the sum of the doses per procedure. However, each individual procedure has its own type and magnitude of possible uncertainty that must be included.

Table 6. Dose in air on the EURT territory with ^{90}Sr contamination density 1 Ci per km^2.

Calendar year	Month	Monthly dose, 10^{-6} Gy per month
1957	October	525
	November	377
	December	278
1958	January	208
	February	160
	March	126
	April	100
	May	82
	June	68
	July	57
	August	49
	September	43
	October	38
	November	34
	December	30

Approach to Uncertainty Propagation

The database processor structure of the current TRDS systems will be helpful in the design of the stochastic dose calculations. There are numerous occurrences of "shared" uncertainties. These largely derive from shared environmental conditions in common residence locations. Thus, in any one realization, people living in a particular village should all see the same conditions at the same time. Water concentrations, external dose rates at the riverbank, and soil concentrations will all be the same for all individuals residing in the same location at the same time. Similarly, those residing in a specified household will all have the same effective shielding factor ($R_{in/out}$). Thus, the databases of environmental information will be used with variable multiplicative "correction factors" as multiple realizations of the possible values.

The relatively limited individual information is unshared. Multiple realizations of doses will be estimated – each will used one of the precalculated sets of environmental data and a random selection of the appropriate unshared stochastic parameters, such as relate to individual behavior or metabolism (T_1, T_2, T_3, and *DF* or *IMR or HSR*).

This general structure was used in the Hanford Environmental Dose Reconstruction Project (Farris et al. 1995; Gilbert et al. 1993) with success.

While such a calculation could conceivably be made with the existing TRDS system, copied and subtly adjusted 100 or 1000 times, it is more economical, practical, and quality-traceable to independently re-code and verify the algorithms in a faster, automated program.

Dosimetric Product Structures

The plan for application of the stochastic TRDS system, when it is complete, is to generate numerous realizations (one hundred to several thousand) of sequential annual organ absorbed dose for every cohort member. This set of outputs will be transferred to the epidemiologists as an input to the dose-response analyses. The individual output vectors will embody the overall uncertainty of the doses; the outputs may be mathematically manipulated to provide mean, median, geometric standard deviation, or other desired statistical parameters for individuals' doses. However, this naive set of outputs will incorporate all shared, unshared, classical, and Berkson uncertainties, because individual dose realizations for every member of the cohort would include the same "shared" data.

As noted above, proper implementation of the dosimetric calculation will appropriately account for shared uncertainties, such that the same realization of dose for every individual would use the same set of shared input parameters. Unshared parameters will be randomly used within each vector for each individual.

This approach combines the classical and Berkson components. In order to separate the contributions of the classical uncertainties from the Berkson uncertainties, an "on-off switch" will be implemented for each class of inputs. In this way, the relative contributions of various input parameters to the overall variance may be determined, and these results combined to state that some fraction of the variance, ν, was contributed by the classical uncertainties and the remaining fraction $(1-\nu)$ was contributed by the Berkson uncertainties.

This can be determined by replicate with specific parameters set as constants to their mean values. If the classical uncertainties were allowed to vary while the Berksons were held constant, a set of realizations might be generated from which a geometric standard deviation GSD_C could be generated; similarly, if the Berkson uncertainties were allowed to vary while the

classically-distributed parameters where held constant, a related GSD_B could be generated. There ought to be some relationship between these statistics and the υ parameter described above.

We understand that the current plan for biostatistical analyses of the dose uncertainty in risk analysis is that described in Stram and Kopecky (2003). These methods require a computer program that generates dose estimates from the distribution of true dose for all individuals given all the data available for the cohort members taking into account both shared and individual sources of error, as is described herein. The following methods of analysis can be considered

1) *Estimation of the average shared multiplicative error coefficient by analysis of the correlation structure of the realizations.*

   Here sample covariances, $C_{i,j}$ (over realizations) of the dose realizations are computed for all pairs of subjects, *i*, and *j*. Then these $C_{i,j}$ are related by simple linear regression to the product of the mean values of dose for subject *i* and *j*. The slope of the regression line estimates the variance of the overall shared multiplicative error (var(SME)), in the simplified model for dosimetry error described by Stram and Kopecky. This can be then be used to approximate the variance of the risk estimate (e.g. the excess relative risk per Gy) obtained according to an equation of form

   $$\text{Var}(\hat{\beta}) = \text{Var}(\hat{\beta} \mid \text{no dosimetry error}) + \beta^2 \text{Var(SME)}$$

   Where $\text{Var}(\hat{\beta} \mid \text{no dosimetry error})$ is calculated using the mean doses as if they are true dose and the standard sampling variance (inverse information matrix) is used. This expression, while based on an (over) simplification, often gives reasonable guidance as to the likely effect of dose errors on risk estimates and confidence intervals

2) *Monte-Carlo maximum likelihood estimation*

   In this analysis, a likelihood is computed as a function of risk estimates β for each of a large number of realizations of true dose provided by the dosimetry system. This likelihood function is averaged over all the realizations and then maximized to provide the MLE of β.

3) *Multiple imputation approach*

   In the multiple imputation method, the risk estimates are computed (by maximum likelihood) for each of the realizations of dose for the cohort, and then the variance of the observed estimates of dose-response is added to the nominal sampling variance of $\hat{\beta}$ as

Var($\hat{\beta}$)=Var($\hat{\beta}$|no dosimetry error)+Var($\hat{\beta}$ over simulations)

There are complications with each of these approaches; the first one depends upon a possibly very oversimplified "model" for the dosimetry system (the simple shared and unshared additive and multiplicative errors). The maximum likelihood method while technically giving the "true" likelihood can be very computationally demanding when there is more than one parameter to be estimated (e.g. additional nuisance parameters in the likelihood) or when true dose response is very strong. Technically, multiple imputation requires (Rubin 1991) that the samples from the distribution of dose be computed conditionally upon not only the input data for each individual in the cohort, but also upon the cancer (or other) outcome data as well. However in some cases this complication can be ignored. Both the MLE and imputation methods require several hundred or more true dose realizations

We are also following with interest the activities by others in dealing with the analysis of uncertainty in dosimetry and its transfer to the dose-response derivation, e.g., Stayner et al. (2007). The approach described above is directly compatible with the Stayner et al. methods.

Distribution History of this Report

A first draft of this report was originally prepared prior to the JCCRER International Meeting in Las Vegas of November 2007. At that time, it was shared and discussed within the JCCRER Project 1.1 team and with those epidemiologists and biostatisticians present who are familiar with Projects 1.1 and 1.2b. It has since been expanded to include the potential confounding pathways of medical x-ray and the EURT exposures for conceptual completeness. The recent improvement of knowledge about the Techa River source term (e.g., Degteva et al. 2008) adds computational complexity, but does not alter the basic structure proposed, by requiring monthly time increments in 1950-1951 and 1957-1958. The most recent update incorporates suggestions from reviewers to the basic equation and some of the definitions, as well as substantial additions considering the nature of correlations, shared/unshared, Berkson/Classical, and Type A/Type B uncertainties. The selection of uncertainty distributions for each key parameter has been based upon the successful implementation of TRDS-2009D.

References

Carroll RJ, Ruppert D, Stefanski LA, Crainiceanu C. *Measurement Errors in Non-Linear Models: A Modern Perspective.* Second Edition. Vol. 105. Chapman & Hall/CRC (2006)

Degteva, MO, Anspaugh LR, Napier BA, Tolstykh EI, Shagina NB, Kozheurov VP, Vorobiova MI, Tokareva EE, Shishkina EA. *Analysis of the main factors contributing to uncertainty in internal dose from ^{90}Sr and feasibility evaluation for reduction in uncertainty*. Chelyabinsk: Urals Research Center for Radiation Medicine; Final Report for Milestone 8; (1999)

Degteva MO, Shagina NB, Vorobiova MI, Golikov VYu, Barkovsky AN, Kozyrev AV, Anspaugh LR, Napier BA. *Reconstruction of individual medical doses for members of the Extended Techa River Cohort.* Chelyabinsk and Salt Lake City: Urals Research Center for Radiation Medicine and University of Utah; Final report for Milestone 19; (October 2007) (in Russian and English)

Degteva MO, Vorobiova MI, Shagina NB, Shishkina EA, Anspaugh LR, Napier BA. *A Review of Data on Releases of Radioactive Wastes from the "Mayak" Production Association into the Techa River In 1949-1956*. Chelyabinsk: Urals Research Center for Radiation Medicine; Appendix to Quarterly Report of ISTC Project No.2841 (January 2008)

Degteva MO, Tolstykh EI, Vorobiova MI, Shagina NB, Kozyrev AV, Anspaugh LR, Napier BA. *Structure of the Revised Techa River Dosimetry System: Exposure Pathways and System Databases*. Chelyabinsk and Salt Lake City: Urals Research Center for Radiation Medicine and University of Utah; Combined report for Milestones 20 and 21, Part 2. (April 2009)

Eckerman KF, Ryman JC. *External exposure to radionuclides in air, water, and soil*. Washington: U.S. Environmental Protection Agency; Federal Guidance Report No. 12, EPA 402-R-93-081; (1993)

Farris WT, Napier BA, Ikenberry TA, Shipler DB. "Radiation Doses from Hanford Site Releases",*Health Physics 71:588-601*; (1995)

Gilbert RO, Simpson JC, Napier BA, Haerer HA, Liebetrau AM, Ruttenbur AJ, Davis S. "Statistical Aspects of the Hanford Environmental Dose Reconstruction Project and the Hanford Thyroid Disease Study", *J. Radiation Research* 124:354-355; (1993).

Glagolenko YuV, Alexakhin AI, Stukalov PM, Ivanov IA, Mokrov YuG, Antonova TA. *Reconstruction of the Techa River source-term parameters on the basis of the data of historical and technological reports*. Ozersk: Mayak PA; Technical Report 2 for ISTC Project #2841; (July 2006) (in Russian)

Glagolenko YV, Drozhko EG, Mokrov YG, Rovny SI, Stukalov PM, Ivanov IA, Alexakhin AI, Vorobiova MI, Degteva MO, Akleyev AV. Reconstruction of parameters of the source of liquid radioactive waste discharges from the radiochemical plant into the Techa River. Report 1: Development of methods, main results. *Radiat Safety Problems (Mayak Production Association Scientific Journal)*, Special Issue 2008: 76-91 (in Russian) 72-86 (in English); (2008)

Hofer E. "How to account for uncertainty due to measurement errors in an uncertainty analysis

using Monte Carlo simulation," *Health Phys.* 95(3):277–290; (2008)

Hoffman FO, Ruttenber AJ, Apostoaei AI, Carroll RJ, Greenland S. "The Hanford Thyroid Disease Study: An alternative view of the findings." *Health Phys*. 92(2):99 –111; (2007)

IAEA (1989). International Atomic Energy Agency. *Evaluating the Reliability of Predictions Made Using Environmental Transfer Models*. IAEA Safety Series No. 100, STI/PUB/835 International Atomic Energy Agency, Vienna, Austria. (1989)

Kozheurov VP, Degteva MO. "Dietary intake evaluation and dosimetric modelling for the Techa river residents based on in vivo measurements of strontium-90 in teeth and skeleton." *Sci. Total Environ*. 14**:**63–72; (1994)

Li Y, Guolo A, Hoffman FO, Carroll RJ. "Shared uncertainty in measurement error problems, with application to Nevada Test Site fallout data." *Biometrics* 63:1226-1236; (2007)

Mallick B, Hoffman FO, Carroll RJ. "Semiparametric regression modeling with mixtures of Berkson and classical error, with application to fallout from the Nevada Test Site," *Biometrics* 58(1): 13–20. (2002)

Napier BA, Shagina NB, Degteva MO, Tolstykh EI, Vorobiova MI, Anspaugh LR. "Preliminary uncertainty analysis for the doses estimated using the Techa River Dosimetry System – 2000." *Health Phys* 81:395–405; (2001)

National Council on Radiation Protection and Measurements (NCRP). *A Guide for Uncertainty Analysis and Dose and Risk Assessments Related to Environmental Contamination*, NCRP Commentary No. 14, National Council on Radiation Protection and Measurements, Bethesda, Maryland. (1996)

National Council on Radiation Protection and Measurements (NCRP). *Evaluating the reliability of biokinetic and dosimetric models and parameters used to assess individual doses for risk assessment purposes.* NCRP Commentary No. 15, National Council on Radiation Protection and Measurements, Bethesda, Maryland. (1998)

National Council on Radiation Protection and Measurements (NCRP). *Uncertainties in the Measurement and Dosimetry of External Radiation*, NCRP Report No. 158, National Council on Radiation Protection and Measurements, Bethesda, Maryland.(2007)

National Council on Radiation Protection and Measurements (NCRP). *Uncertainties in Internal Radiation Dose Assessment*, NCRP Report No. 163, National Council on Radiation Protection and Measurements, Bethesda, Maryland. (2009)

Petoussi N, Jacob P, Zankl M, Saito K. « Organ doses for foetuses, babies, children and adults from environmental gamma rays." *Radiat Prot Dosim* 37:31–41; (1991)

Rubin DB, Schenker N. "Multiple imputation in health-care databases: an overview and some applications." *Stat Med* 10(4):585-594 (1991)

Schafer DW, Gilbert ES. "Some Statistical Implications of Dose Uncertainty in Radiation Dose–Response Analyses," *Radiat Res.* 166, 303–312 (2006)

Shagina NB, Degteva MO, Tolstykh EI. "Uncertainty analysis of strontium retention in humans resulting from individual variability in metabolic parameters." In: *Harmonization of radiation, human life and the ecosystem, Proceedings of 10th International Congress on*

Radiation Protection. Hiroshima, Japan, May 14-19, 2000. Hiroshima: International Radiation Protection Association; CD-ROM, Paper No. P-3a-128; (2000)

Shagina NB, Degteva MO, Tolstykh EI, Zalyapin VI, Krivoshchapov VA, Tokareva EE, Anspaugh LR, Napier BA. *Algorithm for Selection of the Best Individual Estimate of Internal Dose on the Basis of the Results of Calculations Performed with use of Three Different Protocols*. Chelyabinsk: Urals Research Center for Radiation Medicine; Final Report for Milestone 18; (2007)

Stayner L, Vrijheid M, Cardis E, Stram D, Deltour I, Gilbert S, Howe G. "Monte Carlo maximum likelihood methods for estimating uncertainty arising from shared errors in exposures in epidemiological studies." *Radiat Res*. 168, 757–763 (2007)

Stram DO, Kopecky KJ. "Power and uncertainty analysis of epidemiological studies of radiation-related disease risk in which dose estimates are based on a complex dosimetry system: some observations," *Radiat Res.* 160, 408-417 (2003)

Tolstykh EI, Zalyapin VI, Krivoshchapov VA, Shagina NB, Peremyslova LM, Degteva MO, Kozheurov VP, Safronova NG, Anspaugh LR, Napier BA. *Verification of referent-intake levels for strontium-90*. Chelyabinsk and Salt Lake City: Urals Research Center for Radiation Medicine and University of Utah; Final report for Milestone 3 Part 1; (September 2001) (in Russian and English).

Tolstykh EI, Zalyapin VI, Shagina NB, Krivoshchapov VA, Degteva MO, Tokareva EE, Anspaugh LR, Napier BA. *Estimation of individual-to-model ratios (IMR) and their uncertainty for the Techa River residents*. Chelyabinsk and Salt Lake City: Urals Research Center for Radiation Medicine and University of Utah; Final report for Milestone 4; (November 2002) (in Russian and English).

Tolstykh EI, Shagina NB, Peremyslova LM, Degteva MO, Anspaugh LR, Napier BA. *Methodological approaches to the reconstruction of radionuclide intake for residents of the East Urals Radioactive Trace and Karachay Trace.* Chelyabinsk and Salt Lake City: Urals Research Center for Radiation Medicine and University of Utah; Final report for Milestones 11 and 12; (April 2006)

Tolstykh EI, Shagina NB, Krivoschapov VA, Degteva MO, Shishkina EA, Anspaugh LR, Napier BA. *Improvement in Strontium-90 reference intake function for residents of the Techa Riverside settlements*. Chelyabinsk and Salt Lake City: Urals Research Center for Radiation Medicine and University of Utah; Unscheduled report; (October 2008a)

Tolstykh EI, Shagina NB, Peremyslova LM, Safronova NG, Degteva MO, Anspaugh LR, Napier BA. *Improvement in cesium-137 intake with milk contaminated as a result of soil→ grass→ milk transfer*. Chelyabinsk and Salt Lake City: Urals Research Center for Radiation Medicine and University of Utah; Unscheduled report; (October 2008b)

Vorobiova MI, Degteva MO, Kozyrev AV, Anspaugh LR, Napier BA. *External doses evaluated on the basis of the Techa River Dosimetry System Approach*. Final report for Milestone 6. Chelyabinsk: Urals Research Center for Radiation Medicine; (May 1999)

Vorobiova MI, Degteva MO. "Simple model for the reconstruction of radionuclide concentrations and radiation exposures along the Techa River." *Health Phys* 77:142–149; (1999)

Appendix: Generic TRDS-2009MC Output Structure

The general structure of the dosimetry output file is given here. The doses will be stored by source (Techa internal, Techa External, EURT Internal, EURT External, x-ray). Because the doses will be estimated one year at a time, the indexing structure also uses the annual format.

Individual 1:

Techa Internal Dose:

Realization 1:

Year 1: Organ 1, Organ 2, Organ 3, … Absorbed dose in Gy
Year 2: Organ 1, Organ 2, Organ 3, …
Year n: Organ 1, Organ 2, Organ 3, …

Realization 2:

Year 1: Organ 1, Organ 2, Organ 3, … Absorbed dose in Gy
Year 2: Organ 1, Organ 2, Organ 3, …
Year n: Organ 1, Organ 2, Organ 3, …

Realization n:…

Techa External Dose:

Realization 1:

Year 1: Organ 1, Organ 2, Organ 3, … Absorbed dose in Gy
Year 2: Organ 1, Organ 2, Organ 3, …
Year n: Organ 1, Organ 2, Organ 3, …

Realization 2:

Year 1: Organ 1, Organ 2, Organ 3, … Absorbed dose in Gy
Year 2: Organ 1, Organ 2, Organ 3, …
Year n: Organ 1, Organ 2, Organ 3, …

Realization n:…

EURT Internal Dose:

Realization 1:

Year 1: Organ 1, Organ 2, Organ 3, … Absorbed dose in Gy
Year 2: Organ 1, Organ 2, Organ 3, …
Year n: Organ 1, Organ 2, Organ 3, …

Realization 2:

Year 1: Organ 1, Organ 2, Organ 3, … Absorbed dose in Gy
Year 2: Organ 1, Organ 2, Organ 3, …
Year n: Organ 1, Organ 2, Organ 3, …

Realization n:…

EURT External Dose:

Realization 1:

Year 1: Organ 1, Organ 2, Organ 3, … Absorbed dose in Gy
Year 2: Organ 1, Organ 2, Organ 3, …
Year n: Organ 1, Organ 2, Organ 3, …

Realization 2:

Year 1: Organ 1, Organ 2, Organ 3, … Absorbed dose in Gy
Year 2: Organ 1, Organ 2, Organ 3, …
Year n: Organ 1, Organ 2, Organ 3, …

Realization n:…

Medical X-Ray External Dose:

Realization 1:

Year 1: Organ 1, Organ 2, Organ 3, … Absorbed dose in Gy
Year 2: Organ 1, Organ 2, Organ 3, …
Year n: Organ 1, Organ 2, Organ 3, …

Realization 2:

Year 1: Organ 1, Organ 2, Organ 3, … Absorbed dose in Gy
Year 2: Organ 1, Organ 2, Organ 3, …
Year n: Organ 1, Organ 2, Organ 3, …

Realization n:…

Individual 2:

Techa Internal Dose:

Realization 1:

Year 1: Organ 1, Organ 2, Organ 3, … Absorbed dose in Gy
Year 2: Organ 1, Organ 2, Organ 3, …
Year n: Organ 1, Organ 2, Organ 3, …

Realization 2:

Year 1: Organ 1, Organ 2, Organ 3, … Absorbed dose in Gy
Year 2: Organ 1, Organ 2, Organ 3, …
Year n: Organ 1, Organ 2, Organ 3, …

Realization n:…

Techa External Dose:

Realization 1:

Year 1: Organ 1, Organ 2, Organ 3, … Absorbed dose in Gy
Year 2: Organ 1, Organ 2, Organ 3, …
Year n: Organ 1, Organ 2, Organ 3, …

Realization 2:

Year 1: Organ 1, Organ 2, Organ 3, … Absorbed dose in Gy
Year 2: Organ 1, Organ 2, Organ 3, …
Year n: Organ 1, Organ 2, Organ 3, …

Realization n:…

EURT Internal Dose:

Realization 1:

Year 1: Organ 1, Organ 2, Organ 3, … Absorbed dose in Gy
Year 2: Organ 1, Organ 2, Organ 3, …
Year n: Organ 1, Organ 2, Organ 3, …

Realization 2:

Year 1: Organ 1, Organ 2, Organ 3, … Absorbed dose in Gy
Year 2: Organ 1, Organ 2, Organ 3, …
Year n: Organ 1, Organ 2, Organ 3, …

Realization n:…

EURT External Dose:

Realization 1:

Year 1: Organ 1, Organ 2, Organ 3, … Absorbed dose in Gy

Year 2: Organ 1, Organ 2, Organ 3, …
Year n: Organ 1, Organ 2, Organ 3, …
Realization 2:
Year 1: Organ 1, Organ 2, Organ 3, … Absorbed dose in Gy
Year 2: Organ 1, Organ 2, Organ 3, …
Year n: Organ 1, Organ 2, Organ 3, …
Realization n:…

Medical X-Ray External Dose:

Realization 1:
Year 1: Organ 1, Organ 2, Organ 3, … Absorbed dose in Gy
Year 2: Organ 1, Organ 2, Organ 3, …
Year n: Organ 1, Organ 2, Organ 3, …
Realization 2:
Year 1: Organ 1, Organ 2, Organ 3, … Absorbed dose in Gy
Year 2: Organ 1, Organ 2, Organ 3, …
Year n: Organ 1, Organ 2, Organ 3, …
Realization n:…

In this structure, all input parameters {a,b,c…} that are "shared" use value {a1,b1,c1…} in realization 1, value {a2,b2,c2…} in realization 2, etc. "Unshared" parameters are randomly selected for each individual in realization 1, 2, etc.