National Cancer Institute

THE NCI SCHEMA FOR INCORPORATING UNCERTAINTY INTO ESTIMATES OF RADIATION DOSE IN A STUDY OF THYROID DISEASE IN KAZAKHSTAN

Steven L. Simon
Radiation Epidemiology Branch
National Cancer Institute

U.S. DEPARTMENT OF HEALTH AND HUMAN SERVICES

National Institutes of Health

Study Investigators

DOSIMETRY
- Steve Simon
- Andre Bouville
- Nick Luckyanov
- Vladimir Drozdovitch
- Harold Beck
- Robert Weinstock

STATISTICS
- Charles Land
- Deukwoo Kwon

DATA COLLECTION
- Sarah Schonfeld
- Nancy Potischman
- Michael Schwerin
- Zhaxybay Zhumadilov

EPIDEMIOLOGY
- Kiyo Mabuchi

KIBITZER & ADVISER
- Owen Hoffman

Brief Background of Study

Began in 1998 by NCI to study thyroid disease in Kazakhstan in villages near to the Semipalatinsk Nuclear Test Site (one test site of USSR).

Study has goal of deducing the relative effects of internal irradiation by ^{131}I compared to external exposure.

Thyroid medical examinations of about 3,000 persons conducted and interviews on lifestyle.

International initiative conducted over last decade to develop a joint U.S./Russian dosimetry methodology: collaboration between NCI and Institute of Biophysics in Moscow (plus collaborators)

Several sets of doses calculated over last 10 years. First dose response analysis published by Land et al. (2008) based on best estimates of total dose (external + internal).

Field study in Kazakhstan in 2008 to collect additional lifestyle information by focus group strategy.

Development of 2DMC computational strategy begun in 2008 to account for shared and unshared errors.

Location of Semipalatinsk test site and villages of medical examinations. Also showing fallout trajectories of major tests.

Study Population and Exposure Conditions

Study population: 2997 persons of both Kazakh and Russian ethnicity, 19 yrs of age or less at time of the primary exposure event in their village of residence (7 villages where subjects where thyroid examinations were held in 1998).

When considering all nuclear tests that subjects were potentially exposed to, 19 villages must be considered.

Study participants lived in villages of up to a few thousand persons and sustained themselves by local production of foods, primarily meat and milk products. Refrigeration was not available at the time.

Exposures took place 1949-1962 from 11 atmospheric nuclear tests.

Exposure Conditions and Pathways of Exposure

External irradiation from fallout deposition: exposure occurred while both outdoors and indoors in schools or homes. Two types of building constructions considered (Adobe and wood).

Internal thyroid irradiation from consumption of fallout contaminated fresh milk and certain dairy products of short shelf life (yogurt and sour milk and cream).

Milk was obtained from local sources from cows, goats, sheep, and horses. Consumption rates derived from group and individual data. Primary variations in consumption patterns were by age and ethnicity.

Type of dose calculated:

a) External absorbed dose to the thyroid (mGy) of all individual study subjects determined by location of exposure, age, and by lifestyle patterns based on group and individual data.

b) Internal absorbed dose (mGy) to thyroid of all individual study subjects from intake of radioiodines.

Note: This summary document primarily discusses the external dose calculation.

Simplified External Dose Model

$$D_{ext}(mGy) = \int_{TOA}^{t_2} \dot{E}(t) \cdot SF(t) \cdot DF \, dt = DF \int_{TOA}^{t_2} \left[\dot{E}_{12} \sum_{i=1}^{10} A_i e^{-\lambda_i t} \right] SF(t) \, dt$$

where,

$\dot{E}(t)$ is the exposure rate at time *t* in mR/h,

SF(t) is a building shielding factor for home or schools that can change with time having one of two values (1 for outdoors, or, for indoors, equal to the ratio: exposure rate indoors/exposure rate outdoors),

DF (mGy per mR) is an age-dependent dose factor for converting integral exposure to absorbed dose to the thyroid,

$\dot{E}_{12}$ is the exposure rate (mR/h) at 12-h post detonation,

A_i (mR/h per mR/h at H+12) and λ_i (h^{-1}) are parameters of a ten-term exponential fit to exposure rate as a function of time that accounts for both radioactive decay and weathering, specific for the type of nuclear material used in the nuclear device,

TOA is the elapsed time of fallout transit from the site of detonation to the site of deposition, termed for historical reasons as “time of arrival” (TOA), and

t_2 is the ending time of exposure.

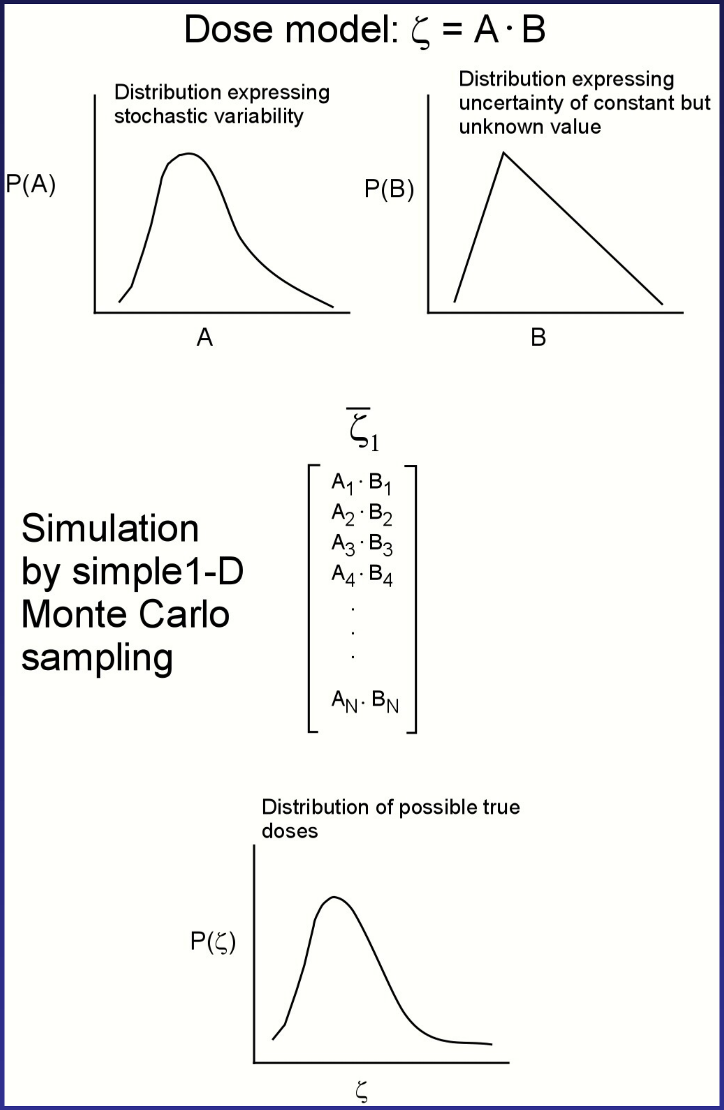
The calculation method used in the NCI schema is the two-dimensional Monte Carlo (**2DMC**).

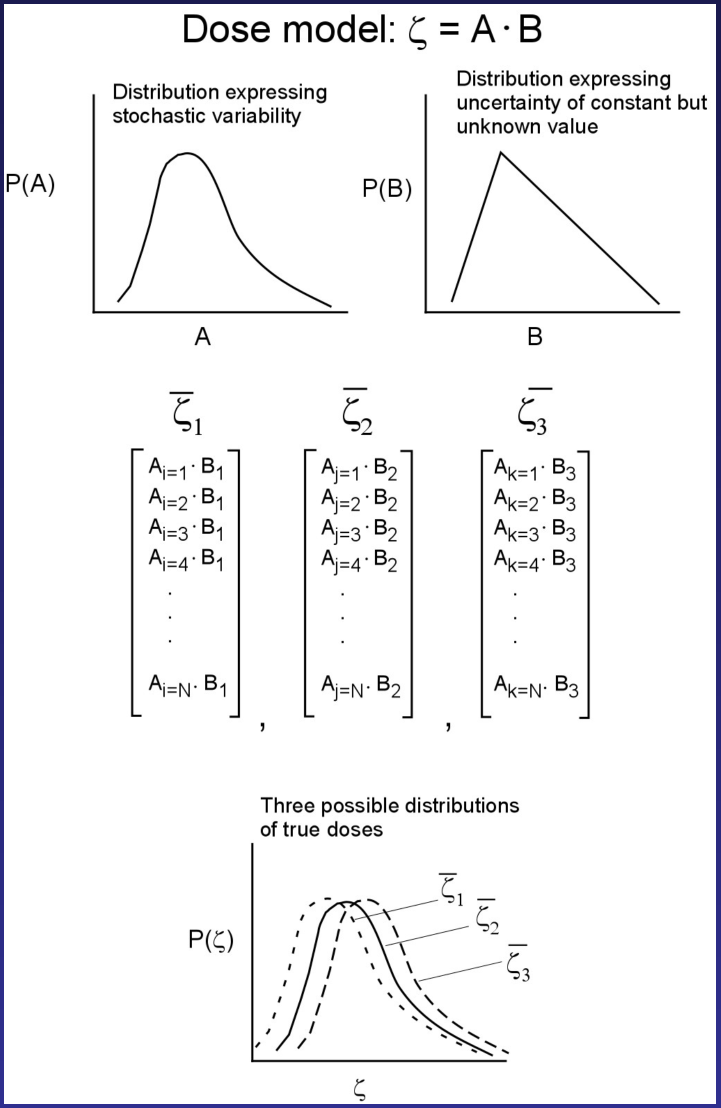
The 2DMC is a simulation technique in which the uncertainty of dose to each individual dose-related is captured among the multiple realizations of the entire set of cohort doses and the variation of doses to persons of similar characteristics is captured within each set.

The overall goal of this simulation method is to allow for multiple estimates of the dose-response to be made based on multiple realizations of the entire cohort dose distribution.

The multiple realizations of the cohort dose distribution are intended to reflect the uncertainties in individual doses due to uncertainty and possible biases in the values of model parameters while maintaining the correct subject-to-subject correlations within each set.

Comparison of 1D and 2DMC





Some terminology I will use:

Realization: a single estimate of dose for a single person.

Trial: a single estimate of dose for all subjects in the cohort where all estimates are internally consistent with one another (properly correlated).

By definition: The number of trials equals the number of independent sets of doses for the entire cohort that are generated in the Monte Carlo simulation as well as the number of dose-responses that can be fit.

Why the term "2 Dimensional MC?"

Variability is captured within each trial.
Uncertainty is captured among trials.

} 2D

Variability represents unshared errors.
Uncertainty represents shared errors.

} 2D

More important terms to understand:

Type A Parameter: Dose-related parameters that represent actual stochastic inter-individual variability of some attribute necessary for the estimation of individual doses.
Example: # hours spent indoors each day by identified individuals.

Type B Parameter: Dose-related parameters that only uncertain due to lack of knowledge of some attribute necessary for estimation of individual doses.
Example: Time of transit of fallout from detonation site to a village.

Type AB Parameter: Dose-related parameters that are variable but also uncertain with respect to estimation of individual doses.
Example: The distribution of deposited radioactivity per unit area derived from sparse measurements over the area of a village or city.

The purpose in assigning a “type” to each parameter is to unequivocally set the strategy of sampling the probability distributions for the model parameters.

This “type” assignment eliminates ambiguity in what otherwise might be a difficult and confusing programming problem

This assignment also allows dosimetrists to focus on understanding the available information and knowledge about each parameter to derive appropriate distributions that describe the variation and/or uncertainty of each.

These various steps are key to correctly implementing the 2DMC.

Rules of the 2DMC schema for parameter sampling:

- Within each trial, each Type B (uncertain) parameter is only sampled once for the entire group, while each Type A (stochastic variability) parameter is sampled once for each subject.

- In each successive trial, each Type B variable is again sampled once for the entire group and each Type A variable is again sampled for each person.

- Within each trial, each Type AB (uncertain and variable) parameter must be defined (or redefined) by sampling a new mean and variance from pre-assigned uncertainty distributions.

- Once done, individual values of the Type AB parameter are sampled in that trial for each subject.

- In each successive trial, a new mean and variance are sampled to determine the uncertainty distribution of the Type AB parameter and new values for each subject are sampled again.

How do you ensure proper sampling of probability distributions in a model with many parameters and/or a model that must account for many subsets of subjects (each with individual characteristics)?

- We considered various strategies including writing the dose model with extensive subscripts to delineate parameter “type”.
- We decided on a tabular system to keep track of parameter type and parameter values. This is the final key to correctly implementing the 2DMC.

THE BIG TABLE – Our book keeping tool.

Para-meter	Type A Parameters (Variability			Type B Parameters (Uncertainty due to lack of knowledge]		
	Frequency Distribution Type	Mean or Geo. Mean	Std. Dev. or GSD	Probability Distribution Type	Mean or Geo. Mean	Std. Dev. or GSD

Let's assume a dose model of this form:

$$D = X \cdot Y \cdot Z$$

X is a Type B parameter (a poorly known constant)

Y is a Type A parameter (variable, and based on lots of relevant data).

Z is a Type AB parameter (variable, but based on little data or data of nominal relevance).

EXAMPLE

Para-meter	Type A Parameters (Variability)			Type B Parameters (Uncertainty due to lack of knowledge)		
	Frequency Distribution Type	**Mean or Geo. Mean**	**Std. Dev. or GSD**	**Probability Distribution Type**	**Mean or Geo. Mean**	**Std. Dev. or GSD**
X				LN (uncertainty in the true fixed value of X)	10	2.0

EXAMPLE

Para-meter	Type A Parameters (Variability)			Type B Parameters (Uncertainty due to lack of knowledge)		
	Frequency Distribution Type	**Mean or Geo. Mean**	**Std. Dev. or GSD**	**Probability Distribution Type**	**Mean or Geo. Mean**	**Std. Dev. or GSD**
X				**LN (uncertainty in the true fixed value of X)**	**10**	**2.0**
Y	**LN**	**10**	**1.4**			

EXAMPLE

Para-meter	Type A Parameters (Variability)			Type B Parameters (Uncertainty due to lack of knowledge		
	Frequency Distribution Type	**Mean or Geo. Mean**	**Std. Dev. or GSD**	**Probability Distribution Type**	**Mean or Geo. Mean**	**Std. Dev. or GSD**
X				LN (uncertainty in the true fixed value of X)	10	2.0
Y	LN	10	1.4			
Z	LN	6	1.6		Uniform [4.0, 10.0]	Uniform [1.3, 1.8]

EXAMPLE

Para-meter	Type A Parameters (Variability)			Type B Parameters (Uncertainty due to lack of knowledge		
	Frequency Distribution Type*	Mean or Geo. Mean	Std. Dev. or GSD	Probability Distribution Type*	Mean or Geo. Mean	Std. Dev. or GSD
X				**LN (uncertainty in the true fixed value of X)**	**10**	**2.0**
Y	**LN**	**10**	**1.4**			
Z	**LN**	**6**	**1.6**		**Uniform [4.0, 10.0]**	**Uniform [1.3, 1.8]**

Completed table presents an unambiguous sampling strategy.

Reiterate :

- Within each trial, each Type B (uncertain) parameter is only sampled once for the entire group,
- Within each trial, each Type A (stochastic variability) parameter is sampled once for each subject,
- Within each trial, each Type AB (uncertain and variable) parameter is redefined by sampling a new mean and variance from pre-assigned uncertainty distributions.
- Once done, individual values of the Type AB parameter are sampled in that trial for each subject.

Parameters of NCI Kazakhstan External Dose Model and Parameter Type Assignments

Parameter	Values are specific for:	Type A Parameter (Variability)	Type B Parameter (Uncertainty)
A_i and λ_i terms of exposure rate functions	Each nuclear test, different levels of fractionation, with and without weathering corrections	Presently assumed as constants without uncertainty, may be modified in the future	
TOA (transit time of fallout, h)	Test and village		√
Exposure rate at TOA, corrected to 12 h post-detonation	Test and village	√	√
Number of hours per day spent indoors in house	Season, Ethnicity, Age, and Gender (>7y)	√	√
Number of hours per day spent indoors in school	Season, Ethnicity, Age, and Gender (>7y)	√	√
Construction materials of houses	Village and Ethnicity	√	√
Shielding factor for houses	Construction material	to be added	√
Shielding factor for schools	Construction material	to be added	√
Factors converting exposure (mR) to external thyroid dose (mGy)	Age	Presently assumed as constants, may be modified	

Actual table with data for all subgroups is 37 pages long.

Some brief results from 2DMC calculations of external dose

100 trials for 2997 subjects computed (~30,000 dose estimates).

The highest external doses were received in Dolon following the first test (copy of the U.S. Trinity device): about 500 mGy (± ~200 mGy depending on age and lifestyle).

Other nuclear tests in Dolon and other villages resulted in external doses of a few mGy to a few hundred mGy (most tests gave a few 10s of mGy).

Some brief results from 2DMC calculations of external dose

100 trials for 2997 subjects computed (~30,000 dose estimates).

Spatial variation of fallout deposition: GSD of 1.15 to 1.4 depending on test and village.

Coefficient of variation (CV) of external dose in a single village among Kazakhs or Russians of same age group: 5% to 8% (reflects similarity of lifestyle).

CVs of external dose (across trials) for individual infants in Dolon following test #1 was 80% and 38%, respectively (reflected uncertainty of time in and outdoors, type of home construction and building shielding).

CVs of dose for older subjects were typically 20% to 25% (can be equated to GSD of ~1.25).

Some brief comments about internal dose

Internal doses potentially much higher depending on age and consumption habits.

Internal dose model is much more complex (many more variables).

Parameter table for internal dose model is over 100 pages long.

Uncertainties expected for individual doses (expressed across trials) by internal irradiation will be much greater (probably greater than GSD of 2.5) than uncertainties for external dose (~GSD of 1.25).

QUESTIONS

Of the 100 (or 1,000) trials simulated, which have better fitting dose-responses?

Why do some trials fit better?

Can those trials that fit better tell us anything important, e.g., which values of parameters are likely closer to the truth?

Might the better fitting dose-responses be better indicators of the true risk?

Are there better ways to generate individual dose or sets of cohort doses for our purposes?

[end]