BIOASSAY EVALUATION ASSUMING MULTIPLE UNKNOWN PARAMETERS APPLYING OPTIMAL DESIGN

Guillermo Sánchez (gsl@fab.enusa.es), Agustin Pérez and Diego Ortíz

ENUSA Industrias Avanzadas S.A., Apdo 328, E-37008 Salamanca, Spain

ABSTRACT

Bioassays can be used to estimate the intake in accidental and routinely situations. To evaluate the effective dose, apart from the intake, we need to know other parameters such as Activity Median Aerodynamic Diameter (AMAD) or the fraction absorption (f1) in the blood from the GI tract. In an accident situation these parameters are often unknown. The bioassay measurement values can be used to estimate by fitting the parameters unknown. We have applied non linear regression for solving this kind of problem. Furthermore, a method to find the best moments where the bioassay measurements should be taken is described. This method is applied to optimal designs. The goodness of the design will depend on the number of samples and the measurement accuracy. It requires obtaining the analytical solution of the biokinetic model as a function of the parameters to be fitted. Different cases are studied using the computer program BIOKMOD (developed by one of the authors. A web version of Biokmod is available at
http://www3.enusa.es/webMathematica/Public/biokmod.html .
Key words: Bioassay, Internal dosimetry, optimal design, non linear regression;.

INTRODUCTION

A few computer codes have been developed to estimate intake and to calculate internal dose using bioassay data. They usually only assume a parameter unknown: the quantity intake. When more than one parameter is unknown (e.g. Intake, AMAD, f1) the ICRP, (Draft 2006), propose a method using the chi squared test. However, it requires a long process where the users must make some assumptions. One of the authors [1,2] has developed a code called BIOKMOD where non linear regression techniques and optimal design are used to estimate the unknown parameters.
We have applied BIOKMOD to the evaluation of internal exposures using bioassay data. The methods described are accompanied with examples.

The mathematical criteria applied by BIOKMOD [1,2,3] to obtain the content $q_i(t)$ at compartment i and to compute the intake retention functions $r_m(t)$ for different kind of bioassays m, with m = {lung retention, daily urine excretion, thyroid content, …can be expressed as follows.

The predicted value for a kind of bioassay m after an acute input "1" at $t = 0$ is obtained by the sum of the content of one or several compartments. It will also be a sum of exponentials

$$r_m(t) = \sum_{v=1}^{l} c_v \; e^{-d_v t} \tag{1}$$

where c_v and d_v are the coefficients obtained solving the model for the specific case. They depend on the characteristic of the particles intaken (AMAD, type of metabolism, element, etc)

In the case of inhalation eqn (1) can be written as [4]:

$$r_m(t) = \sum_{j,v} IDF_j(p)\, c_{j,v}\; e^{-d_{j,v} t} \tag{2}$$

being *IDF* the initial deposition factor, p the AMAD value in μm

Multiple constant intakes

If we assumed multiple single intakes $\{I_1, \ldots, I_i, \ldots, I_n\}$ where I_j represents the intake which happens on the day j, then the retention function $RM_{\mathrm{m}}(t)$ on the day t is given by

$$RM_m(t) = I_1 r_m(t) + I_2 r_m(t-1) + \cdots + I_t r_m(1) = \sum_{j=1}^{t} I_j \, r_m(t-j+1) \tag{3}$$

In many situations the intake I_j happens for a few hours every day.That is: $\{I_1, \ldots, I_i, \ldots, I_n\}$ are multiple constant intakes where each I_i occurs the day t_i during a time T_i (usually a shift). Then the retention function $RMC_{\mathrm{m}}(t)$ can be computed as follows (Sanchez 2007)

$$RMC_m(t) = I_1\, rc_m(t, T_1) + I_2\, rc_m(t-1, T_2) + \cdots + I_t\, rc_m(1, T_t) = \sum_{j=1}^{t} I_j\, rc_m(t-j+1, T_j) \tag{4}$$

Random Intakes

In real situations, for workers being exposed to radioactive aerosols during the working days, the individual daily intake I is usually a random variable. In a previous article we found [2,3] that multiple random daily intake $\{I_1, \ldots, I_j \ldots, I_n\}$ happens then the retention function, called $RA_m(t)$, can be approximated by

$$RA_m(t) = \mu_I \sum_{j=1}^{t} r_m(j) \pm z\, \sigma_I \sqrt{\sum_{j=1}^{t} r_m^2(j)} \tag{5}$$

being $\hat{\mu}_I = \frac{1}{N}\sum_i I_i$; $\hat{\sigma}_I^2 = \frac{1}{N-1}\sum_i (I_i - \mu_I)^2$ and z is the 100 $(\gamma + 1)/2$ – the quantile of the standard normal distribution.

FITTING BIOASSAY DATA

Bioassays can be used to estimate the initial intake I for the case of an acute intake exposure for an individual worker. To evaluate the effective dose, apart from I, we need to know other parameters such as Activity Median Aerodynamic Diameter (AMAD) or the fraction absorption (f1) in the blood from the GI tract, but in an accident situation these parameters are often unknown. The bioassay measurement values can be used to estimate by fitting the parameters unknown. BIOKMOD applied optimal design to find the best moments where the bioassay measurements should be taken.

Let's suppose a single intake I_0 (unknown) in $t = 0$ of radioactive particles, whose characteristics (AMAD, solubility, etc) are known, by a worker with a metabolism that responds to the ideal model for the standard worker. At time t after the intake, a bioassay is made obtaining a measurement m, with negligible uncertainties. If we make a bioassay at time t obtaining a value m, then $I_0 = m/r(t)$. In this case the intake I_0 will be known with only one measurement. This is an unrealistic situation. In the real world the evaluation of internal exposures using the bioassay data involves a lot of uncertainties. In fact, in an intercomparison exercise where the same cases, using the same data, have been evaluated by different experts, large discrepancies have been obtained [5].

If all parameters (AMAD, absorption parameters, etc.) of the model, except the quantity intakes, are assumed to be known, the only uncertainties will be the ones of the measurements. The linear statistical model can be applied to estimate $\hat{I}$ and its associated uncertainty u_I (e.g.:[6, 7]) obtaining

$$\hat{I} = \frac{\sum_{i=1}^{N} r_{C,j}(t_i)\frac{m_i}{u_i^2}}{\sum_{i=1}^{N}\frac{r_{C,j}^2(t_i)}{u_i^2}}, \qquad u_I = \frac{1}{\sqrt{\sum_{i=1}^{N}\frac{r_{C,j}^2(t_i)}{u_i^2}}} \tag{6}$$

where
t_i is the time from the start of the intake to the measurement i.
m_i and u_i are the measurement and their associated uncertainties (calculated with the same confidence level that u_I).
$R_{C,j}(t)$, with $C = \{A$ (acute) or Cr (Chronic)$\}$ is the retention function, with $I_o = 1$ or $I_d = 1$, associated with measurement m_i; and j is the type of bioassay (note: different kinds of bioassays can be applied simultaneously)
Other authors recommend (ICRP Draft 2006) the maximum likelihood method which uses the eqn (7) instead of eqn (6)

$$\ln(\hat{I}) = \frac{\sum_{i=1}^{N}\left(\frac{\ln(m_i/r_{C,j}(t_i))}{(\ln SF_i)^2}\right)}{\sum_{i=1}^{N}\frac{1}{(\ln SF_i)^2}} \tag{7}$$

being SF_i the scattering factor for m_i. If the bioassay data are log normally distributed then SF is the geometric standard deviation (SG) of the log-normal distribution.

Most of the codes use eqn (6) or (7), even BIOKMOD. The chi squared test (χ^2) should be used to estimate the goodness of the fitted data (ICRP, Draft 2006). BIOKMOD has also other possibilities. It can be assumed that not only the intake I but also other parameters $\{k_1,\ldots, k_r\}$ are unknown (AMAD, f_1, etc.) then it applies eqn (8) for fitting the bioassay data (minimizing χ^2):

$$(\hat{I},\hat{k}_1,\ldots,\hat{k}_r) : Arg \underset{[I,k_1,\cdots,k_r]}{Min}\left[\sum_{i=1}^{N}\frac{\left(I\, r_C(t_i,k_1,\ldots k_r)-m_i\right)^2}{u_i^2}\right],$$

$$\text{Restrictions :} \qquad (8)$$

$$I>0,\ k_1(\min)\le k_1\le k_1(\max),\ldots,k_r(\min)\le k_r\le k_r(\max)$$

If the bioassay data are log normally distributed then is used the eqn (9).

$$(\hat{I},\hat{k}_1,\ldots,\hat{k}_r) : Arg \underset{[I,k_1,\cdots,k_r]}{Min}\left[\sum_{i=1}^{N}\frac{\left(\ln\left[I\, r_{C,j}(t_i,k_1,\ldots k_r)\right]-\ln\left[m_i\right]\right)^2}{SG_i^2}\right],$$

$$\text{Restrictions :} \qquad (9)$$

$$I>0,\ k_1(\min)\le k_1\le k_1(\max),\ldots,k_r(\min)\le k_r\le k_r(\max)$$

The minimization of eqn (8) or (9) is a problem of nonlinear optimization. BIOKMOD applies the algorisms available by *Mathematica*, (http://documents.wolfram.com/mathematica/functions/AdvancedDocumentationNMinimize [Accessed 17 june 2008]), these are probably the state of the art in optimization.

Optimal design

To obtain the moments where the bioassay data should be taken we will use optimal design.

Let's suppose that the retention function $R_m(I, \beta, t)$ after an acute input at t=0, for the bioassay m chosen, can be expressed as function of the unknown parameters. It has the following pattern

$$R_m(I,\beta,t)=I\sum_{r=1}^{q}F_r(\beta)\,e^{-G_r(\beta)\,t} \qquad (10)$$

being Fr(β) and Gv(β) expressions obtained solving the model for the specific case.

Now we apply the D-optimal design method[8,9]:

Given a model η(t; s), where s={I, $\beta_1,\ldots,\beta_q$}, that we rewrite s={$s_1,\ldots, s_p$}, is the vector of unknown parameters, the Fisher information matrix M for a specific design ξ={$t_1,\ldots,t_n$} (ti is the time when the i-th sample should be taken) will be

$$\mathbf{M}=E\left[\frac{\partial \log l}{\partial s_i}\frac{\partial \log l}{\partial s_j}\right] \qquad (11)$$

where l denotes the likelihood function for the regression residuals.

When the model is not linear with respect to the parameters, the information matrix (and then the optimal designs) will depend on the unknown parameters. In this case, initial values are needed for the "non-linear" parameters(8), and the designs computed will be locally optimal. A D-optimal design will be one that leads the determinant of the information matrix to a maximum. The information matrix is the main tool for computing optimal designs, since it is asymptotically proportional to the inverse of the covariance matrix of the estimators of the model parameters.
If the model $\eta(t; s)$ is differentiable with respect to the parameters with a continuous derivative, the information matrix for design ξ and normally distributed random errors can be written as

$$\mathbf{M} = \mathbf{X}^T \boldsymbol{\Sigma}^{-1} \mathbf{X} \qquad (12)$$

where
X is the n × p matrix whose i-th row is

$$\nabla \eta(t_i, s) = \left(\frac{\partial \eta(t_i, s)}{\partial s_1}, \ldots, \frac{\partial \eta(t_i, s)}{\partial s_p} \right) \qquad (13)$$

and Σ denotes the covariance matrix of the residuals.

Since all the bioassays are performed on an individual worker it is convenient to consider a non-trivial covariance matrix. One of the common choices used in literature is $\Sigma = \sigma^2 \Gamma$, where σ^2 is the standard deviation associated with the system of measurement, and $\Gamma=(l_{ij})$ with $l_{ij}=\exp(-\rho|t_i-t_j|)$, meaning that the relationship between samples decays exponentially with increasing time-distance between them. The parameter ρ is characteristic of the worker, being a typical value $\rho=1$ that will be used in the examples of the next section. The D-optimal design will be the set of values of t_i that leads $\det|M|$ to a maximum.

APPLICATIONS

In the following cases we will use BIOKMOD (Fig.1) to estimate the intake fitting bioassay data.

Case 1.- As a result of Chernobyl accident (26 april 1986) a 39 years old male and 80 kg (member of the public) has been exposed to continuous and unknown ingestion of Cs-137 (Ansoborlo et al 2003). The results of the whole body activity retention are given below :
WholeData {time after the accident (d), activity (Bq)} = {{39,300}, {58,671}, {75,737}, {130,1661}, {156,1846}, {170,1882},{198,2247}, {234,2493}, {263,2926}, {297,3224}, {325,3608}, {374,3883}, {408,3773}, {432,3723}, {494,3195}, {520,2740}, {556,2469}, {592,2375}, {625,1954}, {682,1614}, {744,1221}, {800,1174}, {880,739}};

These values are represented in fig,2. It can be observed that the retention was increasing until T. We can suppose that the caesium ingestion happened until T, when it ceased. BIOKMOD computes the best fit supposing a daily chronic ingestion I during

a time t1. The ingestion of caesium stop in $t = T$, for t>T, I = 0. So it is obtained that the accumulated intake was 13133 Bq

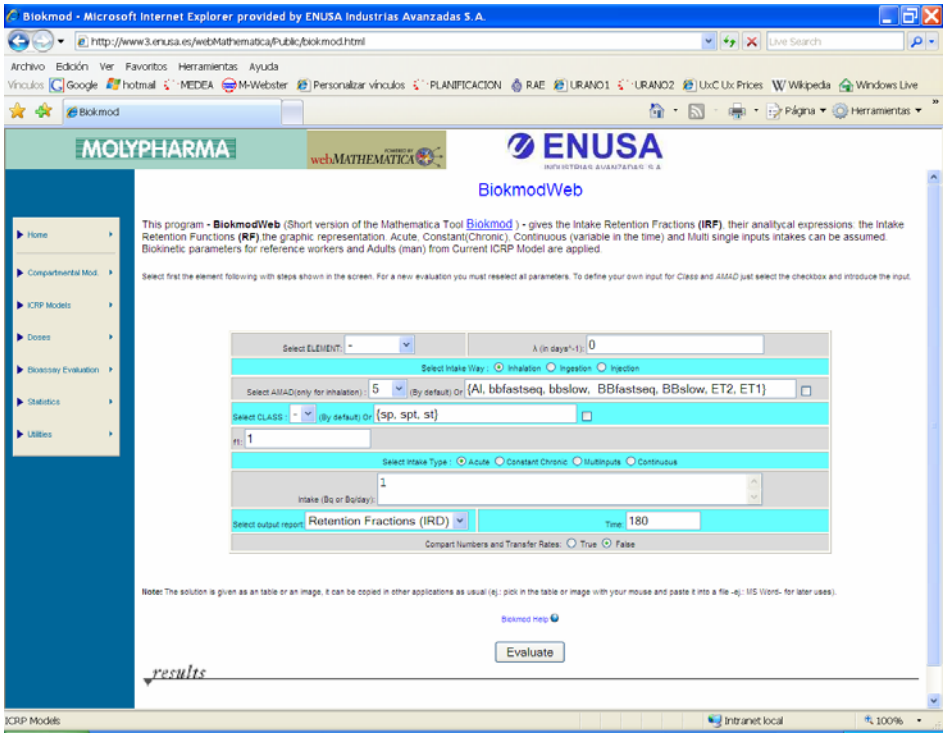


Figure 1 Web version of Biokmod (http://www3.enusa.es/webMathematica/Public/biokmod.html)

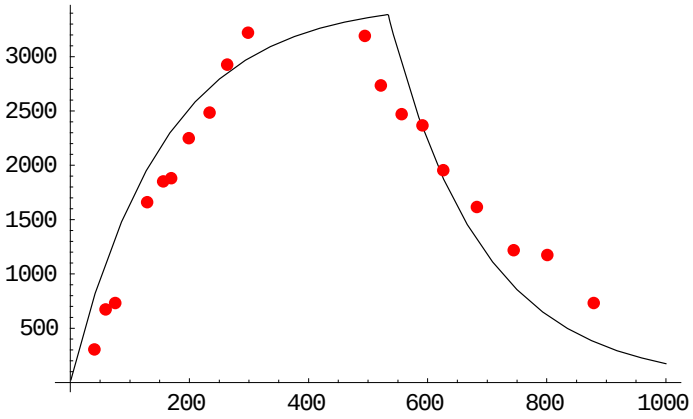


Figure 2 Whole body retention vs the best fit theoretical retention funtion

Table 1.- Bioassay data for Case 2 and Case 3.

Case 2			Case 3	
Time of measurement after intake (d)	Daily urinary excretion rate of ^{60}Co (Bq/d)	Whole body activity of ^{60}Co (Bq)	Time after the intake (d)	Lung activity of U (Bq)
10		$2.39\ 10^4$	1	186
14	709	$2.92\ 10^4$	5	181
17		$2.01\ 10^4$	30	161
20		$1.82\ 10^4$	70	149
27	64	$2.16\ 10^4$	120	143
40	71	$1.98\ 10^4$	250	113
60	37	$2.16\ 10^4$		
80	29	$1.75\ 10^4$		
190	11	$1.16\ 10^4$		
370	1.7	$8.1\ 10^3$		
747		$4.8\ 10^3$		
1010		$2.7\ 10^3$		

Case 2.- An operator has been exposed to a simple accidental intake by inhalation of ^{60}Co. The cobalt form was metal and oxide. A program (Table 1) of in-vivo monitoring was carried out ten days after the event and continued up to 3 years. Urine samples were also taken Additional information: It is recommended to assume that the whole body and urine measurements be approximated by a log-normal distribution with a geometric standard deviation of 1.07 Bq and 1.8 Bq respectively (Data from IM 2005 European workshop on individual monitoring of ionising radiation. Vienna April 2005, available at http://www.ideas-workshop.de [Accessed 15 June 2006]).
We have assumed that apart of the intake I, p (AMAD value in μm), the absorption rates :$\{s_{pt,}\ s_{p,}\ s_t\}$ and f_1 are unknown. The best fit obtained corresponds to 398.5 kBq with AMAD 5.5 μm, $\{s_{pt,}\ s_{p,}\ s_t\}$ ={10, 90, 0.0007} and f_1 = 0.1. The committed effective dose, E(50) calculated using these values is: 4.5 mSv.

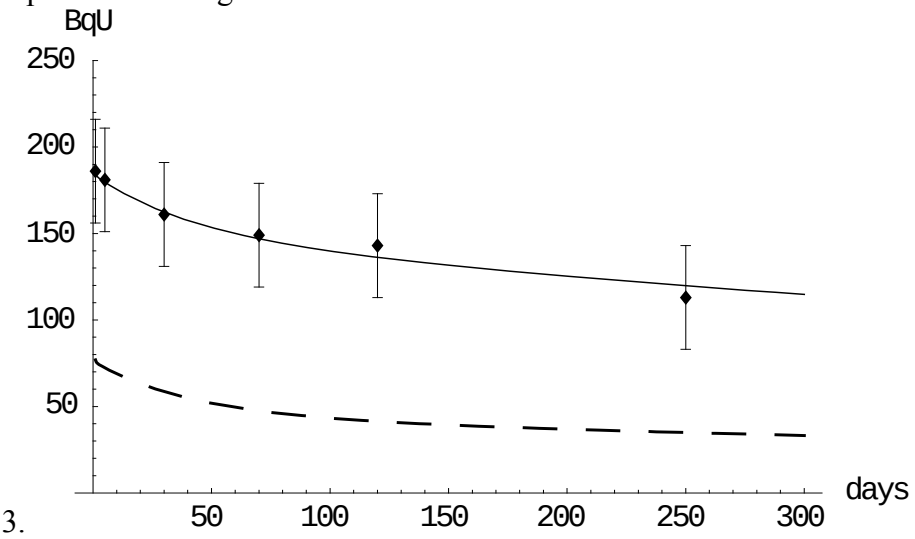
The above solution can be compared to the given in Annex B of ICRP (Draft 2006). The method applied requires a high participation by the user. The solution reported is: intake 404 kBq and committed effective dose, E(50) 5.0 mSv.

We have used optimal design in the previous example to decide the best moment when the measured should be taken The 3-point D-optimal designs for different values of f1 are shown in Table 2 proving that the optimal designs are very robust with respect to the election of the initial value for parameter f1.

Table 2.- Results of D-optimal designs for 60Co whole body retention varying the initial value of f1

f_1	*Days after the intake happened*		
0.05	0.5	4.00	7.04
0.10	0.5	3.97	7.01
0.15	0.5	3.94	6.99
0.20	0.5	3.91	6.96

Case 3.- A worker has been exposed from $t = 0$ to $t = 2000$ day to a chronic intake by inhalation of 3 BqU/day of UO_2 aerosols type S and AMAD 5 μm. On the day $t = 2000$ he accidentally has an intake by inhalation of an unknown I quantity of UO_2. The uranium lung content has been measured (Table 1) using a lung body counter with a standard deviation of 15 BqU. We want to know the accidental quantity of intake. The solution obtained is that the accidental intake was 1174 ± 247 BqU for AMAD 5μm 5 m (computed with a confidence interval of 95%, $z \approx 2$) . If it is supposed that the AMAD is unknown then eqn (19) is applied obtaining 1875 BqU and AMAD 7.8 μm. These are nearer to the "true" values (1700 BqU and AMAD 7 μm). The solution is represented in Fig.



3.

Figure 3 Predicted lung retention after an acute intake assuming a previous chronic intake (Case 2 of the main text). The dashed line represents the underlying contribution from the chronic intake..

Remark: If the AMAD value is not really known, the bioassay data should be fitted taken the AMAD as a parameter unknown to be fitted. This does not apply for substances type F and $f_1 = 1$.

Case 4.- A worker has accidentally intaken by inhalation an unknown I quantity of UO2 being the AMAD p also unknown. We wish to estimate I measuring the uranium lung content using a lung body counter with σ = 1.8 Bq. It will be assumed that the worker had not previously been exposed to significant uranium intakes.

Note.- The disintegration constants can be assumed "0" for all uranium isotopes because their half-lives are too long. The AMAD p value is expected in the range between 1 μm and 10 μm.

In this case the unknown variables are I and AMAD p. To define the optimal design the first step to obtain the lung retention function RLung(I,p,t) for these kinds of radioactive aerosols. It can be made with BIOKMOD choosing metabolism type S.
We have used p= p_0 = 5 μm as the initial value for this parameter when computing optimal designs. The designs obtained taking an initial value p0 = 5 μm have proved to be very robust with respect to this choice, giving very high efficiencies in every case. The designs, computed for different number of sample points, are shown in Table 4.

Table 3.- Results of the *D*-optimal designs for Uranium retention (solubility S and $\rho = 1$) in the lungs as function of *n* (number of points used for the estimation)

n	Days after the intake happened										
2	0.5	69									
3	0.5	65	73								
4	0.5	61	69	77							
5	0.5	58	66	73	81						
6	0.5	5	69	77	84	92					
7	0.5	5	66	74	81	88	95				
8	0.5	4	9	72	80	87	94	102			
9	0.5	4	8	12	77	84	91	98	106		
10	0.5	4	8	12	74	81	88	95	102	109	
11	0.5	4	8	65	72	78	85	91	97	104	111

CONCLUSION

The estimation of unknown intake using bioassay data requires a high participation of the radiologist that usually must make many assumptions. We have show by example that using BIOKMOD this evaluation can be made in an easy way.. The standard version of BIOKMOD is available for free download at the author web site: http://web.usal.es/guillermo. Furthermore there is a web version (available at http://www3.enusa.es/webMathematica/Public/biokmod.html , sponsored by ENUSA Industrias Avanzadas. S.A) and therefore it can be used everywhere where an internet connection exists.

Apart from accute intake BIOKMOD can be used in the evaluation of internal exposures using the bioassay data in cases that represents real situations: Multiple constant and random intakes in occupational exposures taking into account periods without intake (weekends, holidays, etc.);

REFERENCES

[1] Sánchez G; "Fitting bioassay data and performing uncertainty analysis with BIOKMOD" Health Physics. 92(1) :64-72. 2007. ISSN/ISBN: 0017-9078

[2] Sanchez, G. BIOKMOD: A Mathematica toolbox for modeling Biokinetic Systems. Mathematica in Education and Research 10(2): 50 – 70 ; 2005.

[3] Lopez-Fidalgo, J. Sanchez, G. Statistical criteria to establish bioassay programs, health phys. Health Phys. 89(4): 333–338; 2005.

[4] Sanchez G, Lopez-Fidalgo, J. Mathematical techniques for solving analytically large compartmental systems, Health Phys: 85(2):184 –193; 2003.

[5] Doerfel H. EULEP/EURADOS Feedback: 3rd European Intercomparison Exercise on Internal Dose Assessment. Proceedings of 3rd European ALARA Network Workshop. Munich 1999

[6] Potter Charles A. Intake Retention Fractions Developed from Models Used in the Determination of Dose Coefficients Developed for ICRP Publication 68. Special Issue. Health Phys: 83(5); 2002.

[7]Skrable, KW. French, CS. Chabot, GE. Tries, M. La Bone M. *Variance Models for estimating intakes from repetitive bioassay measurements*. Chapter 5 in: Both WE, ed. Practical applications of internal dosimetry: Madison, WI; Medical Physics Publishing; 2002.

[8] Lopez-Fidalgo, J. Rodriguez-Diaz, JM. Sanchez, G. Santos-Martin, MT. Optimal designs for compartmental models with correlated observations. Journal of Applied Statistics: 32(10) 1075–1088; 2005.

[9] G. Sánchez; J. M. Rodríguez-Díaz. Optimal design and mathematical model applied to establish bioassay programs": Radiation Protection Dosimetry. doi:10.1093/rpd/ncl499. 2007. ISSN/ISBN: ISSN 1742-3406.

[10]Ansoborlo, E. Berard, P. Eckermann, K. Berkovski, V. Birchall, A.. Fry, F. Guilmette, R. Miller, G. Ishigure, N. Lipsztein, J. and Nosske, D. Review of methods and computer codes for interpretation of bioassay data. Radiat Prot Dosimetry. 105: 341–346; 2003.

INTERNATIONAL COMMISSION ON RADIOLOGICAL PROTECTION. ICRP Database of Dose Coefficients: Workers and Members of the Public. Version 2.0.1 (CD-ROM), Oxford: Pergamon Press; 2001.

INTERNATIONAL COMMISSION ON RADIOLOGICAL PROTECTION. Draft guidance document: Bioassay data interpretation. http://www.icrp.org/news_guidance.asp [Accessed 15 June 2006]); 2006.