Risk of thyroid cancer following ^{131}I exposure in childhood –

Belarus/Russia/EU/IARC case-control study

Elisabeth Cardis, CREAL, Barcelona

centre for research in environmental epidemiology

List of collaborators

Belarus
Anoshoko, Y.
Astakhova, L.N.
Demidchik, E.P.
Drozdovitch, V.
Kroupnik, T.
Malakhova, I.V
Masyakin V.
Ostapenko, V.
Piliptsevitch, N.
Poliakov, S.
Shevchuk, V.
Sidorov, Yu.
Vorobey, S.

Russia
Abrosimov, A
Balonov, M.
Chekhin, S.
Khrouch, Y.
Ivanov, V.K.
Korobova, E.
Lushnikov, E.
Maksyoutov, M.
Parshin, A.
Parshkov, E.P.
Pitkevitch, V.A.
Stepanenko, V.
Tsyb, A.
Vlassov, O.
Zvonova, I.

Finland
Suonio, E.

Germany
Goulko, G.

Italy
Pacini, F.
Pinchera, A.

Japan
Hoshi, M.
Ito, M.
Shibata, Y.
Yamashita, S.

Sweden
Galanti, R.

UK
Thomas, G.
Williams, D.

US
Bouville, A.

IARC
Cardis, E.
Kesminiene, A.
Maceika, E.
Drozdovitch, V.
Tenet, V.

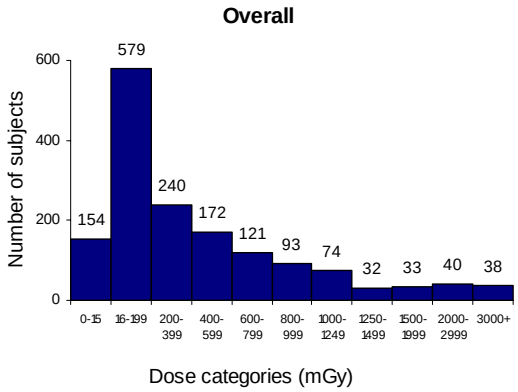
centre for research in environmental epidemiology

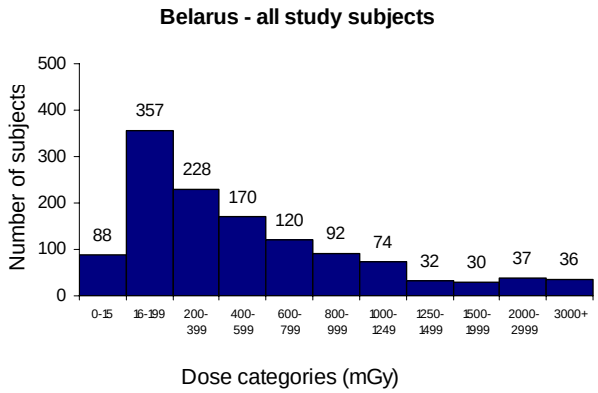
Number of cases and controls interviewed by region

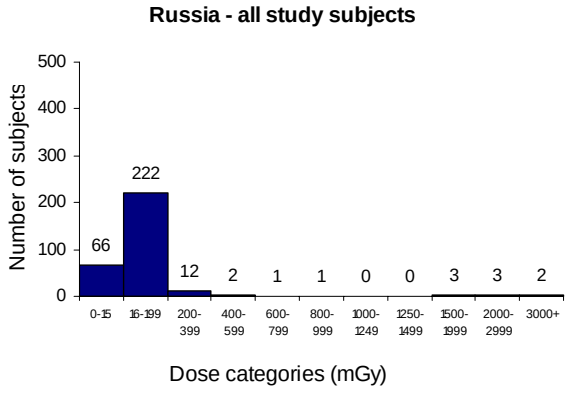
Status	Belarus		Russian Federation				Total
	Gomel	Mogilev	Bryansk	Kaluga	Orel	Tula	
No. of cases	188	32	11	10	18	17	276
No. of controls	877	167	49	39	87	81	1300
Age at exposure (cases only)							
<2 y	69	10	4	1	1	2	87
2-4 y	59	8	3	2	3	5	80
5-9 y	45	3	2	4	6	6	66
10-14 y	15	11	2	3	8	4	43
Sex (cases only)							
Boys	74	7	6	4	6	5	102
Girls	114	25	5	6	12	12	174

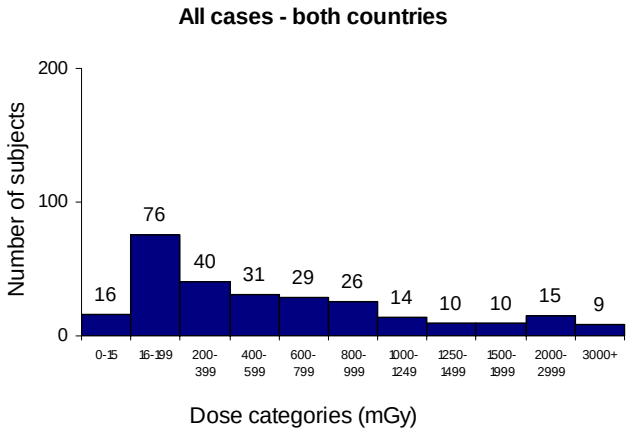
Note: results presented here are for subjects aged < 15 ATA – in Russia about 30 additional cases were also interviewed who were 15-18 ATA

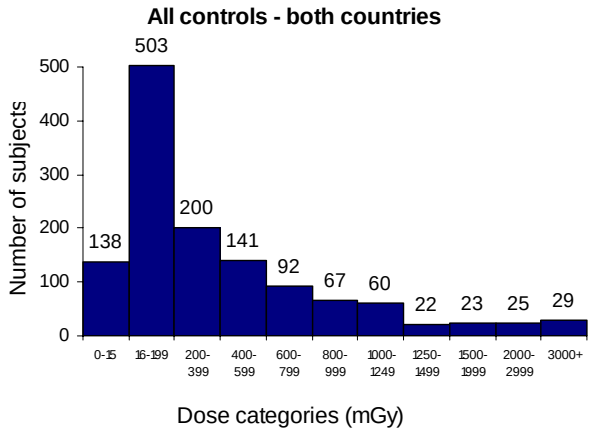
Total thyroid dose among study subjects *(mGy)*











centre for research in environmental epidemiology

Thyroid dose distribution (mGy)

		Median	Maximum
I-131	Belarus	355.7	9 528
	Russia	39.4	5 257
Short-lived	Belarus	1.6	534
	Russia	0.1	26
External	Belarus	2.4	98
	Russia	0.9	31
Long-lived	Belarus	1.2	42
	Russia	0.4	12
Total	Belarus	365.4	10 163
	Russia	40.4	5 314

centre for research in environmental epidemiology

Analyses

- Main analyses
  - Conditional logistic regression
  - Adjustment for confounders
    - ✓ By stratification
  - Effect modification
    - ✓ Modelled as interaction
    - ✓ Likelihood ratio test to evaluate significance
  - Reference dates for controls: date of diagnosis of matched case

… all time-dependent variables calculated up to reference date

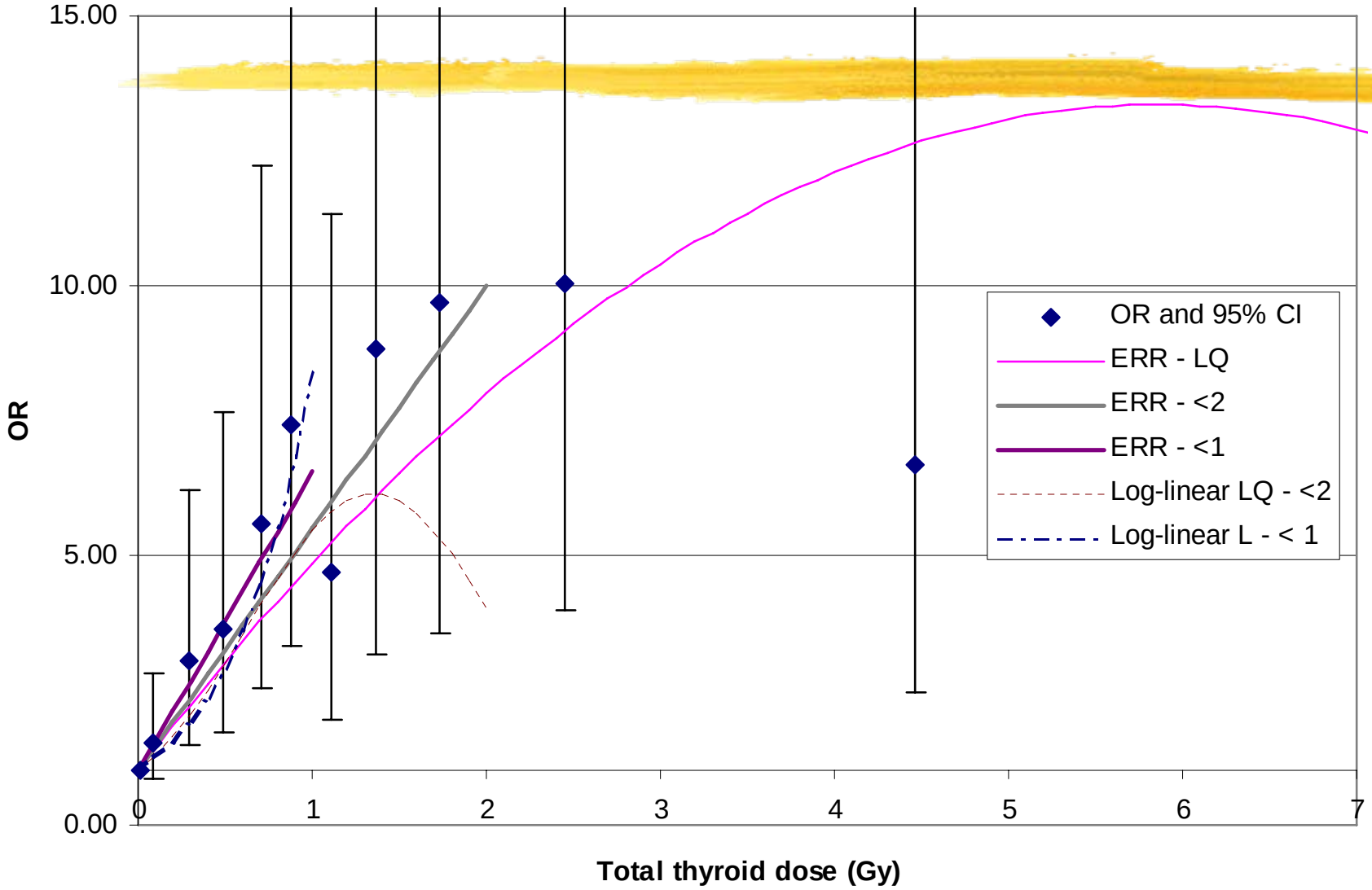
Analyses (cont'd)

- Main analyses
  - Log-linear risk model: OR = exp [β * f(d)]

    … also ERR model: OR = 1 + β * f(d)
  - Main dose-response analyses
    - ✓ Continuous exposure measures
      - Departures from linearity explored with polynomials
    - ✓ Lag of doses – 5 years (long lived isotopes, external)
    - ✓ Risk estimates – MLE ; Confidence intervals: likelihood-based ; p-values: 2-sided

centre for research in environmental epidemiology

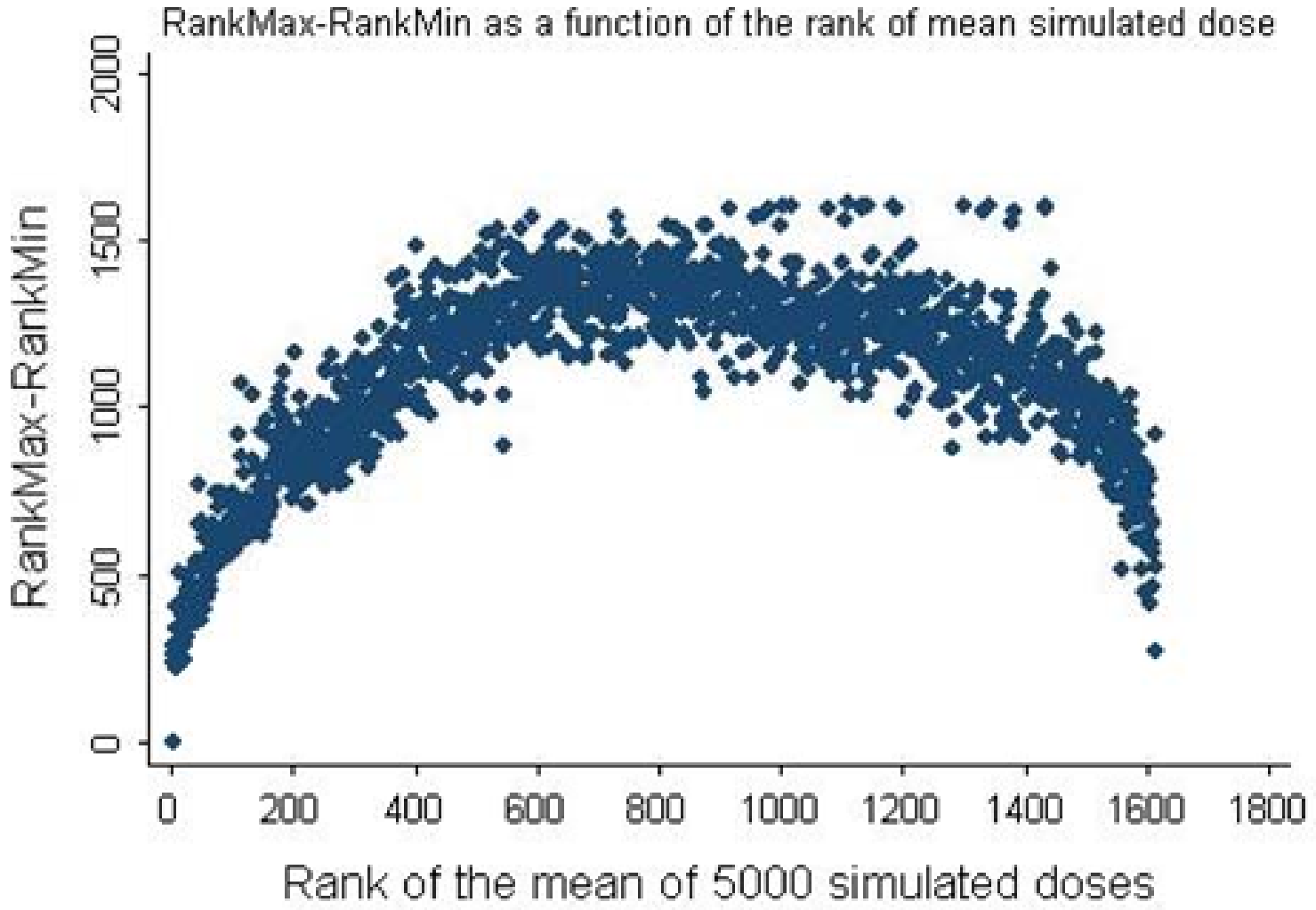
Summary of dose-response relationship



centre for research in environmental epidemiology

Approach for taking into account uncertainties

- Initial proposal – MCML method we have used before *(Stayner et al 2008)*
  - X thousand realisations of the doses taking into account (in a 2 stage fashion) the shared and unshared errors

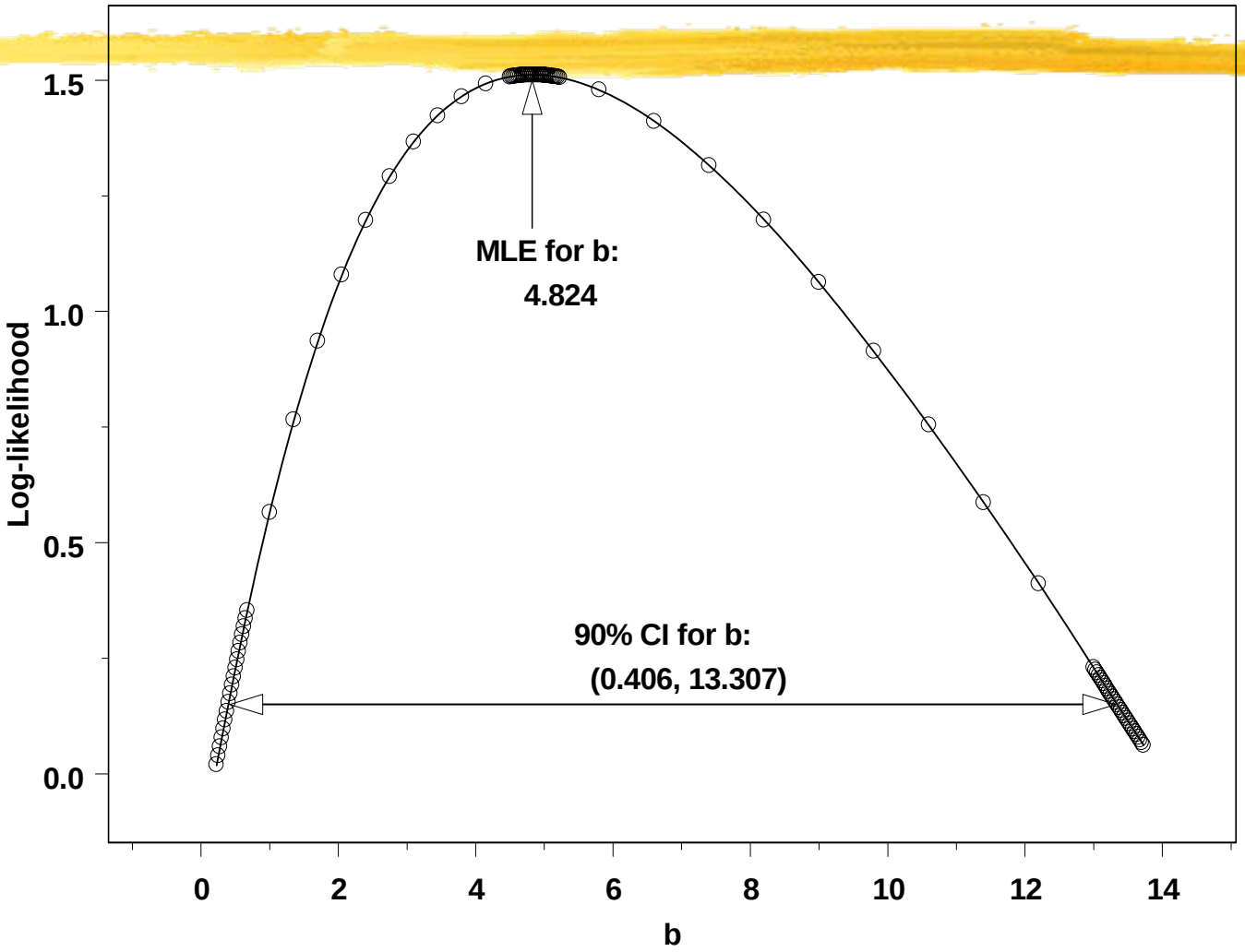

RankMax-RankMin as a function of the rank of mean simulated dose
RankMax-RankMin
0
500
1000
1500
2000
0
200
400
600
800
1000
1200
1400
1600
1800
Rank of the mean of 5000 simulated doses

centre for research in environmental epidemiology

Approach for taking into account uncertainties

- Initial proposal – cont´d
  - ⇨X thousand data sets –
    - ✓Derive the profile likelikhood for each
    - ✓Average the profile likelihoods over all the realisations
    - ✓Obtain the MLE and 95% CI

Figure 3: Profile likelihood, MLE and 90% CI for b from the 10,000 simulations



centre for research in environmental epidemiology

Problems

- Dose-response is not linear …
  - Depending on scale, the best fitting polynomial is either L-Q or L-Q-C ….
    - ✓ Calculate and maximise profile likelihoods on 2,3 or more dimensions …Not easy !
    - ✓ Or restrict analyses to subjects in dose-range where linear …
      - But subjects will change between simulations
  - The shape of the best fitting dose-response may actually vary in the different realisations
  - The magnitude and possibly the shape of the dose response may be influenced by modifiers (ID)…

centre for research in environmental epidemiology

Solution – choice of dose responses

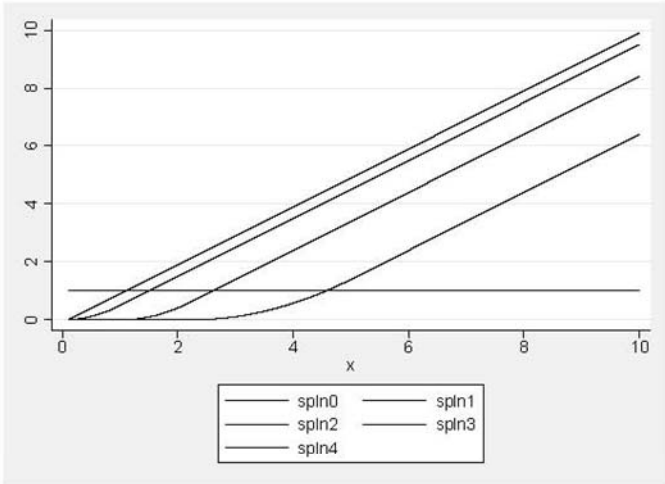
The work presented here is being conducted by Graham Byrnes (IARC)

- Objective
  - Use all the data
  - Avoid imposing a dose-response function that may not fit the data well

... and polynomials can provide very poor fit at extreme values, since data at small doses can strongly influence the predicted value at large doses

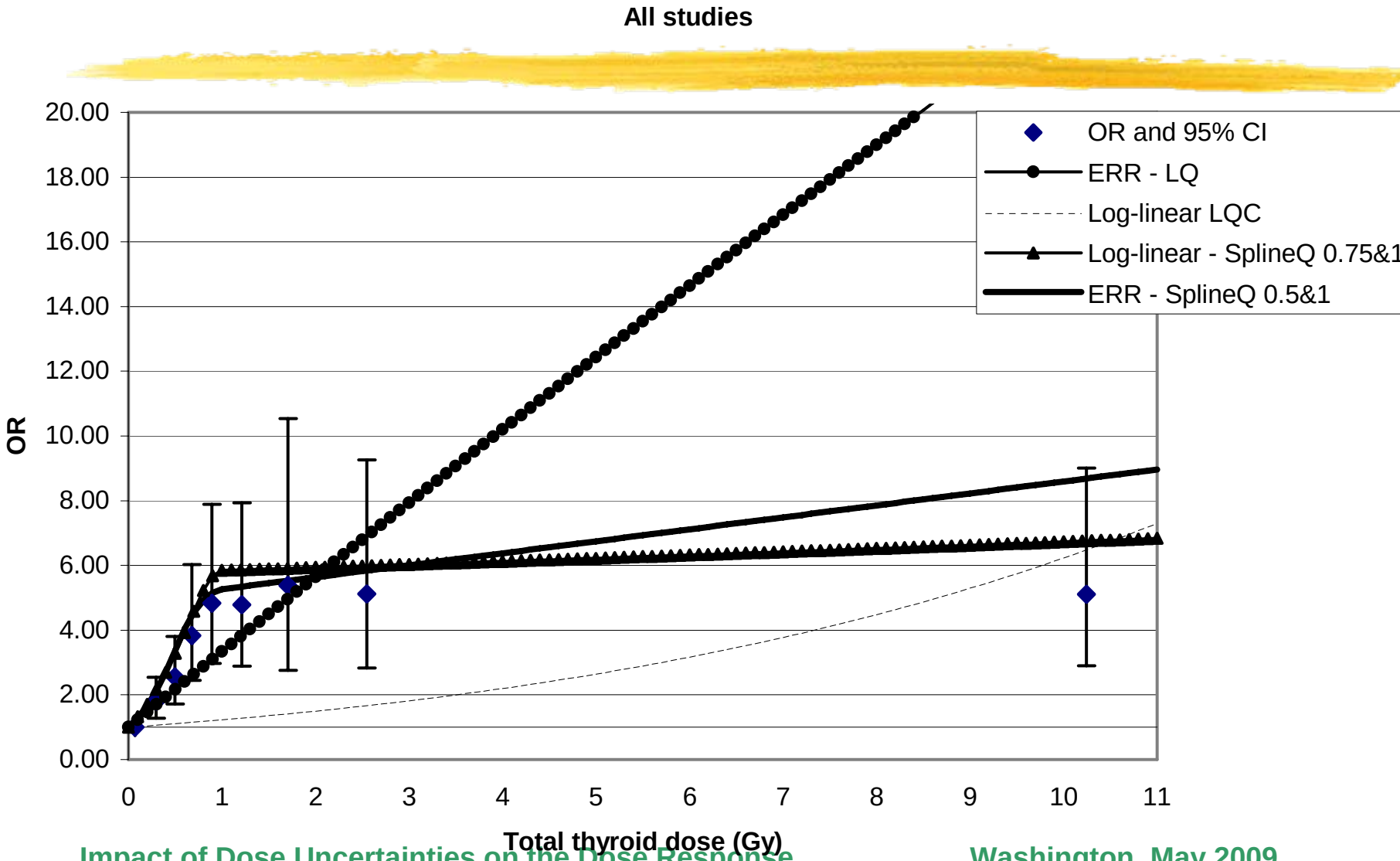
Natural Quadratic Splines

Any linear combination of the S_i will have continuous first derivatives



CREAL centre for research in environmental epidemiology

Seem to predict qualitatively similar response curves regardless of the regression framework (ie logistic vs linear ERR, log-transformed or raw dose).



centre for research in environmental epidemiology

Solution - averaging of k-dimensional likelihood

- As an alternative method for k parameters (after trying a lot of things):
  - use the inverse Hessian at the MLE to obtain a preliminary variance-covariance matrix V ;
  - apply a coordinate transform (orthogonal transform followed by scaling) to map V to the identity matrix;
  - Define a (k – 1)-sphere in these coordinates, centered at the MLE, with radius $\Phi^{-1}(1 - \alpha/2)$. Conventionally, this would define the boundary of the 95% confidence set;
  - distribute points uniformly on the surface of the sphere (deterministically for $k \leq 2$, at random for $k \geq 3$);
  - adjust the radial position of each point, via 1-d Newton's method (with jump size limited to half or double the current radius) to achieve the appropriate difference of likelihood relative to the MLE.

centre for research in environmental epidemiology

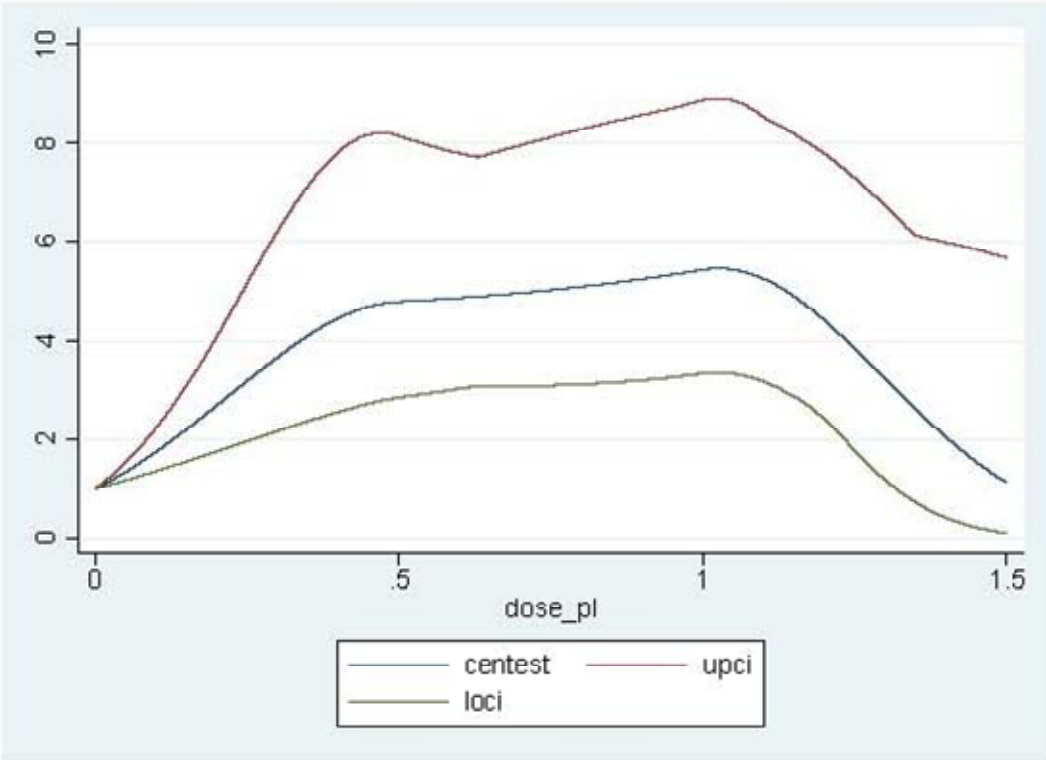
Parameter space to envelope of curves

- The real goal is not to define the confidence set B in the parameter space, but to find the corresponding family of dose response curves
  - For any give dose x and parameter vector $\beta 1,\ldots,\beta k$, the corresponding log odds-ratio curve has value

$$\Psi(x) = \sum_{j=1}^{k} \beta_j . S_j(x)$$

  - To find the boundary of the envelope of curves, we need to locate for each x, the β belonging to B which maximises or minimises $\Psi(x)$:
  - Ideally this means locating the points on ∂B where the tangent space T ∂B is orthogonal to $(S1(x),\ldots, Sk(x))$.
  - This is easy when the boundary is ellipsoidal, in general not.
  - Practical solution: for each of the uniform boundary points, generate the corresponding curve; then for a set of doses, find the max and min among these curves.

OR scale, restricted



Where we are …

- The approach can be implemented
- But:
  - width is based on the posterior likelihood having a chi-squared distribution … which appears not to be true….
    - ✓ It is the log of the sum over the Monte-Carlo simulations of the *product* of the likelihoods from the conditionally independent case-control sets… *and the usual assumption of asymptotic normality fails.*
  - confidence intervals appear to be too narrow …
    - ✓ and so are those in our previous 1-parameter analyses in nuclear workers and in liquidators …
  - simulations currently being conducted to determine the distribution of the posterior likelihood