from a food, not understanding that even contact can result in trace contamination sufficient to induce an allergic reaction.

Funding: No additional funding.

Conflict of interest: None.

- 1 Fries JH. Peanuts: allergic and other untoward reactions. *Ann Allergy* 1982;48:220-6.
- 2 Yunginger JW, Sweeney KG, Sturner WQ, Giannandrea LA, Teigland JD, Bray M, Benson PA, et al. Fatal food-induced anaphylaxis. *JAMA* 1988; 260:1450-2.
- 3 Assem ESK, Gelder CM, Spiro SG, Baderman H, Armstrong RF. Anaphylaxis induced by peanuts. *BMJ* 1990;300:1377-8.
- 4 Sampson HA. Peanut anaphylaxis. *J Allergy Clin Immunol* 1990;86:1-3.
- 5 Smith A. Allergy to peanuts. *BMJ* 1990;300:1354.
- 6 Bock SA, Atkins FM. Patterns of food hypersensitivity during sixteen years of double-blind placebo-controlled food challenges. *J Pediatr* 1990;117:561-7.
- 7 Burks AW, Williams LW, Mallory SB, Shirrell MA, Williams C. Peanut protein as a major cause of adverse food reactions in patients with atopic dermatitis. *Allergy Proc* 1989;10:256-9.

- 8 Atkins FM, Steinberg SS, Metcalfe DD. Evaluation of immediate adverse reactions to foods in adult patients. A detailed analysis of reaction patterns during oral food challenge. *J Allergy Clin Immunol* 1985;75:356-63.
- 9 Donovan KL, Peters J. Vegetableburger allergy: all was nut as it appeared. *BMJ* 1990;300:1378.
- 10 Arshad SH, Malmberg E, Krapf K, Hide DW. Clinical and immunological characteristics of Brazil nut allergy. *Clin Exp Allergy* 1991;21:373-6.
- 11 Sampson HA, Scanlon SM. Natural history of food hypersensitivity in children with atopic dermatitis. *J Pediatr* 1989;115:23-7.
- 12 Bock SA, Atkins FM. The national history of peanut allergy. *J Allergy Clin Immunol* 1989;83:900-4.
- 13 Taylor SL, Busse WW, Sachs MI, Parker JL, Yunginger JW. Peanut oil is not allergenic to peanut-sensitive individuals. *J Allergy Clin Immunol* 1981;68:372-5.
- 14 Monet-Vautrin DA, Harahet R, Kanny G, Ait-Djafar Z. Allergenic peanut oil in milk formulas. *Lancet* 1991;338:1149.
- 15 Platts-Mills TAE. Type 1 or immediate hypersensitivity: hayfever and asthma. In: Lachmann PJ, Peters DK, eds. *Clinical aspects of immunology*. 4th ed. Oxford: Blackwell, 1982:570-686.
- 16 Warren JB, Doble N, Dalton N, Ewan PW. Systemic absorption of inhaled epinephrine. *Clin Pharmacol Ther* 1986;40:673-8.

(Accepted 9 January)

Glycophorin A biodosimetry in Chernobyl cleanup workers from the Baltic countries

William L Bigbee, Ronald H Jensen, Toomas Veidebaum, Mare Tekkel, Mati Rahu, Aivars Stengrevics, Ausrele Kesminiene, Juozas Kurtinaitis, Anssi Auvinen, Timo Hakulinen, Kristina Servomaa, Tapio Rytömaa, G Iris Orams, John D Boice Jr

See editorial by
Henshaw and p 1107

Center for Environmental and Occupational Health and Toxicology, Graduate School of Public Health, University of Pittsburgh and Molecular Carcinogenesis Program, University of Pittsburgh Cancer Institute, Pittsburgh, PA 15238, USA

William L Bigbee, associate professor

Department of Laboratory Medicine, University of California, San Francisco, CA, USA
Ronald H Jensen, professor

Department of Epidemiology and Biostatistics, Institute of Experimental and Clinical Medicine, Tallinn, Estonia
Toomas Veidebaum, director
Mare Tekkel, staff scientist
Mati Rahu, chair

Latvian Cancer Registry, Riga, Latvia
Aivars Stengrevics, head

Lithuanian Chernobyl Medical Center and Lithuanian Oncology Center, Lithuanian Cancer Registry, Vilnius, Lithuania
Ausrele Kesminiene, chief
Juozas Kurtinaitis, chief

The 1986 Chernobyl nuclear accident resulted in massive contamination of the area, necessitating evacuation of the population, extensive environmental cleanup of radioactive materials, and construction of a sarcophagus to isolate the reactor.¹ These operations were accomplished by 500 000 to 600 000 workers from all 15 republics of the former Soviet Union. To characterise the exposure to radiation and the potential adverse health outcomes in these populations, we have undertaken a comprehensive cohort study in the Baltic countries using record linkage techniques. Here we report estimates of physical doses and biodosimetry data for 782 of these workers.

Subjects, methods, and results

We identified three populations of Chernobyl workers who were male residents of Estonia (4836), Latvia (5709), and Lithuania (5446) and who were sent to the Chernobyl area primarily in 1986 or 1987. Estimates of their physical doses are based on dosimetry records obtained from Soviet military lists and individual Chernobyl passports. We derived biodosimetry data for 453 workers from Estonia (recorded physical doses: range 0.02-28.3 cGy, median 9.5 cGy, arithmetic mean (SD) 10.7 (6.4) cGy),

281 from Latvia (range 0-27.8 cGy, median 9.4 cGy, mean (SD) 9.6 (7.7) cGy), and 48 from Lithuania (range 2.5-36.0 cGy, median 16.2 cGy, mean (SD) 16.1 (7.7) cGy). Given the uncertainties of measurement and reporting surrounding these estimates, we wished to determine whether the radiation doses received by these workers resulted in a detectable biological response in an independent biodosimetric assay. We used the glycophorin A in vivo somatic cell mutation assay. This uses immunolabelling and flow cytometry to enumerate variant erythrocytes in peripheral blood expressing phenotypic loss of the glycophorin A allele resulting from mutations in the glycophorin A gene in bone marrow progenitor cells.² This assay has shown an association between exposure to ionising radiation and long term elevation of variants with loss of the glycophorin A allele in several populations, including those at Hiroshima, Japan,³ Chernobyl,⁴ and Goiânia, Brazil.⁵

We measured the frequency of such variants (per million erythrocytes analysed) in blood samples from the 782 workers and 60 male control subjects (27 from Estonia, 24 from Latvia, and 9 from Lithuania). These controls were from the same populations from which the cleanup workers were drawn and were group

Table 1—Frequencies of variant erythrocytes with loss of glycophorin A allele (per million erythrocytes analysed) in blood samples from Chernobyl cleanup workers and controls from Baltic countries

Population	No of subjects	Frequency of variants (×10 ⁻⁶)			P value*
		Range	Median	Mean (SD)	
Controls†	60	0.2-38.4	6.0	6.7 (5.5)	
	59‡	0.2-13.6	6.0	6.2 (3.6)	
Cleanup workers:					
Estonia	453	0.3-145.6	6.6	8.4 (9.1)	0.073
	444‡	0.3-24.6	6.6	7.6 (4.5)	0.071
Latvia	281	0.3-213.8	7.0	9.6 (14.6)	0.062
	274‡	0.3-32.0	7.0	8.0 (5.7)	0.068
Lithuania	48	2.0-96.2	6.6	9.9 (14.3)	0.11
	46‡	2.0-15.0	6.5	7.2 (3.1)	0.15
All	782	0.3-213.8	6.8	8.9 (11.7)	0.054
	762‡	0.3-25.4	6.6	7.7 (4.8)	0.062

* Mann-Whitney U test for workers v combined controls.

† Comprising 27 subjects from Estonia, 24 from Latvia, and 9 from Lithuania.

‡ Extreme outlier values (>3.0×distance between 25th and 75th centiles) omitted.

Finnish Centre for Radiation and Nuclear Safety and Finnish Cancer Registry, Helsinki, Finland

Anssi Auvinen, *senior scientist*
Timo Hakulinen, *professor*
Kristina Servomaa, *laboratory head*
Tapio Rytömaa, *research director*

Extramural Programs Branch and Radiation Epidemiology Branch, National Cancer Institute, Bethesda, MD, USA

G Iris Orams, *chief*
John D Boice Jr, *chief*

Correspondence to:
Dr Bigbee.

matched for age. The data, summarised in table 1, show a tendency for increased frequency of variants with loss of the glycophorin A allele in all three groups of workers compared with the controls, although none of the differences is significant. The pooled results, with an average increase of frequency of $\sim 1.2 \times 10^{-6}$ corresponding to radiation doses of $\sim 4-8$ cGy,⁴ indicate that these workers' average exposure was unlikely to greatly exceed 10-20 cGy, the approximate minimum radiation dose detectable by our assay.

Comment

We undertook this biodosimetry study to ascertain whether many Chernobyl cleanup workers received substantial radiation exposures that were either undocumented or inaccurately recorded. Our initial biodosimetry data strongly suggest that this is unlikely. It also seems that there is not a large subset of these workers who received doses substantially above the average physical doses. Thus the estimates of physical doses, while perhaps incomplete and imprecise, cannot be rejected as inadequately characterising the workers' exposures. Our results support the use of these estimates to assess possible health hazards and as the basis of power calculations for epidemiological studies of populations of Chernobyl cleanup workers. To strengthen this conclusion, we are now performing fluo-

rescence in situ hybridisation (FISH) based chromosomal translocation analysis in the peripheral blood lymphocytes of these workers. We are also studying the incidence of leukaemia and prevalence of thyroid cancer and are constructing assessments of exposure using combined physical dosimetry records, extensive questionnaire data, and biological dosimetry methods.

Funding: This work was supported by Contract No N01-CP-50520 from the Radiation Epidemiology Branch of the US National Cancer Institute.

Conflict of interest: None.

- 1 United Nations Scientific Committee on the Effects of Atomic Radiation (UNSCEAR). *Sources, effects and risk of ionizing radiation. 1988 Report to the general assembly, with annexes (E.88.IX.7)*. New York, NY: United Nations, 1988.
- 2 Jensen RH, Bigbee WL. Direct immunofluorescence labeling provides an improved method for the glycophorin A somatic cell mutation assay. *Cytometry* (in press).
- 3 Langlois RG, Akiyama M, Kusunoki Y, Bigbee WL, Grant SG, DuPont BR, et al. Analysis of somatic mutations at the glycophorin A locus in atomic bomb survivors: a comparative study of assay methods. *Radiat Res* 1993;136:111-7.
- 4 Jensen RH, Langlois RG, Bigbee WL, Grant SG, Moore D II, Pilinskaya M, et al. Elevated frequency of glycophorin A mutations in erythrocytes from Chernobyl accident victims. *Radiat Res* 1995;141:129-35.
- 5 Straume T, Langlois RG, Lucas J, Jensen RH, Bigbee WL, Ramalho AT, et al. Novel biodosimetry methods applied to victims of the Goiânia accident. *Health Phys* 1991;60:71-6.

(Accepted 9 April 1996)

Statistics Notes

Transformations, means, and confidence intervals

J Martin Bland, Douglas G Altman

This is the 18th in a series of occasional notes on medical statistics

When we use transformed data in analyses,¹ this affects the final estimates that we obtain. Figure 1 shows some serum triglyceride measurements, which have a skewed distribution. A logarithmic transformation is often useful for data which have positive skewness like this, and here the approximation to a normal distribution is greatly improved. For the untransformed data the mean is 0.51 mmol/l and the standard deviation 0.22 mmol/l. The mean of the $\log_{10}$ transformed data is -0.33 and the standard deviation is 0.17. If we take the mean on the transformed scale and back transform by taking the antilog, we get $10^{-0.33}=0.47$ mmol/l. We call the value estimated in this way the geometric mean. The geometric mean will be less than the mean of the raw data.

When triglyceride is measured in mmol/l the log of a single observation is the log of a measurement in mmol/l. The average of n such transformed measurements is also the log of a number in mmol/l, so the antilog is back in the original units, mmol/l.

The antilog of the standard deviation, however, is not measured in mmol/l. Calculation of the standard deviation of the log transformed data requires taking the difference between each log observation and the log geometric mean. The difference between the log of two numbers is the log of their ratio.² As a ratio is a dimensionless pure number, the units in which serum triglyceride was measured would not matter; the standard deviation on the log scale would be the same. As a result, we cannot transform the standard deviation back to the original scale.

If we want to use the standard deviation or standard error it is easiest to do all calculations on the transformed scale and transform back, if necessary, at the end. For example, the 95% confidence interval for the mean on the log scale is -0.35 to -0.31. To get back to the original scale we antilog the confidence limits on the log scale to give a

Fig 1—Serum triglyceride and $\log_{10}$ serum triglyceride concentrations in cord blood for 282 babies, with best fitting normal distribution

95% confidence interval for the geometric mean on the natural scale (0.47) of 0.45 to 0.49 mmol/l. For comparison, the 95% confidence interval for the arithmetic mean using the raw, untransformed data is 0.48 to 0.54 mmol/l. These limits are wider than those for the geometric mean. This is because with highly skewed data the extreme observations have a large influence on the arithmetic mean, making it more prone to sampling error. Lessening this influence is one advantage of using transformed data.

If we use another transformation, such as the reciprocal or the square root,¹ the same principle applies. We carry out all calculations on the transformed scale and transform back once we have calculated the confidence interval. This works for the sample mean and its confidence interval. Things become more complicated if we look at the difference between two means. We shall look at this in another Statistics Note.

- 1 Bland JM, Altman DG. Transforming data. *BMJ* 1996;312:770.
- 2 Bland JM, Altman DG. Logarithms. *BMJ* 1996;312:700.

Department of Public Health Sciences, St George's Hospital Medical School, London SW17 0RE
J Martin Bland, *professor of medical statistics*

ICRF Medical Statistics Group, Centre for Statistics in Medicine, Institute of Health Sciences, PO Box 777, Oxford OX3 7LF
Douglas G Altman, *head*

Correspondence to:
Professor Bland.

BMJ 1996;312:1079