Department of Geological Sciences,
University of Miami,
Coral Gables, Florida 33124, USA

1. Maddox, J. *Nature* **326**, 637 (1987).
2. Solomon, A.M. *et al.* in *Atmospheric Carbon Dioxide & the Global Carbon Cycle* (ed. Trabalka, J.R.) (USDOE, Washington, 1985).
3. Goreau, T.J. & de Mello, W.Z. in *Biogeochemistry of Tropical Rain Forests: Problems for Research* (eds Athie, D., Lovejoy, T.E. & Oyens, P. deM.) (WWF–CENA, Piricicaba, São Paulo, Brazil, 1985).
4. Goreau, T.J. & de Mello, W.Z. *Eos* **67**, 878 (1986).
5. Goreau, T.J. & de Mello, W.Z. *Ambio* (in the press).

Incubation time for AIDS

SIR—In "The sombre view of AIDS"[1], Rees suggests "a 15-year mean incubation and a high incidence of disease among carriers of the virus". In view of the seriousness of the implications, we offer some critical comments on the manner in which these startling conclusions have been derived and presented.

Assuming a normal distribution, the mean incubation time was estimated from US data[2] on 145 transfusion cases of AIDS (acquired immune deficiency syndrome). The maximum observed incubation time was 7 years and only 18 cases had incubation times greater than 4 years. This does not, of course, imply that longer incubation times are rare, but simply reflects the low prevalence of human immunodeficiency virus (HIV) infection in the United States in the early years of the epidemic. But it does mean that, if the author's estimate of 15 years was correct, it would have been obtained from data belonging exclusively to the extreme left of the distribution (more than 1.5 standard deviations from the mean) and representing less than 1 per cent of the total cases among those infected in the period 1978–84. In the absence of any indication of the associated confidence interval, the result would thus have very little credibility, however correct the estimation procedure.

More specific criticisms can be made regarding the methodology used to determine the best-fitting normal distribution. First, no systematic search appears to have been carried out to investigate the effect on the chosen criterion of varying, independently, the two parameters (the mean, μ, and the standard deviation, σ) to be estimated. In fact several quite different combinations of μ and σ may yield essentially the same value of the chosen criterion (in this case the "total absolute difference" (TAD) rather than the more usual sum of squared differences). For instance, TAD $= 14.67$ with $\mu = 15$ and $\sigma = 5$, TAD $= 14.47$ with $\mu = 5$ and $\sigma = 2.14$, and TAD $= 13.53$ with $\mu = 10$ and $\sigma = 3.75$. Although not explicitly stated, the examples presented by Rees give the impression that, in searching to minimize the TAD, an arbitrary constraint was imposed by fixing the ratio μ/σ at 3. If this is the case, there is

absolutely no reason to suppose that the unconstrained minimum has been found.

Among several classical optimization methods available, the method of maximum likelihood[3] has the advantage of being able to provide much needed estimates of confidence limits. This approach was used by Lui *et al.*[4], who fitted a Weibull distribution to a subset (83 cases) of the same data set and estimated the mean at 4.5 or 5 years (according to whether multiple transfused cases were included or not) with 90 per cent confidence intervals of (2.6 years – 14.2 years) or (2.3 years – 19.6 years) respectively. More recently, Iversen and Engen[5], also using a maximum likelihood method, have obtained similar estimates ($\mu = 5$ years, $\sigma = 1.6$ years) by fitting a normal distribution to the full data set. Regrettably, they did not give a confidence interval. Because they have used more data (150 cases), this could be expected to be smaller than that of Lui *et al.*, but would nevertheless remain very wide.

The very large disparity between the above estimates and those of Rees can be explained. Peterman *et al.*[2] classified their data by year of transfusion and length of incubation period (in months). For reasons which are unclear, Rees has attempted to classify by year of transfusion and year of diagnosis (his Table 1). As the exact time of transfusion is not stated, the year of diagnosis cannot be ascertained precisely, thus leading to the use of 2-year classes (1978/79, 1979/80, . . .) which are not mutually exclusive. The TAD is evaluated from the 'observed' and 'expected' numbers of cases in these overlapping classes. If, however, the original presentation of the data is retained, for both the observed and expected numbers (Table 1), the column

totals represent the cases observed with specified incubation periods (0–1 years, 1–2 years, . . .), these classes being mutually exclusive. When the TAD is calculated in this way, quite different results are obtained. In particular, it now appears that the combination $\mu = 5, \sigma = 1.6$ does indeed provide a better fit to the data than $\mu = 15, \sigma = 5$ (Table 2).

Iversen *et al.* estimate (unfortunately without confidence limits) that 27 per cent of seropositive individuals will eventually develop AIDS. Rees's estimate of almost 100 per cent is based not only on his much more pessimistic estimates of the incubation time distribution but also on the assumed prevalence of HIV infection in the US and UK populations, estimates of which are known to be very imprecise.

In conclusion, while a mean incubation period of 15 years (or even more) cannot be ruled out, we believe that the best estimate at the present time remains at about 5 years, although with a very wide confidence interval.

DOMINIQUE COSTAGLIOLA
Unité de Recherches Biomathématiques et Biostatistiques, INSERM,
Université Paris 7, 2 Place Jussieu,
75251 Paris Cedex 05, France
ANGELA M. DOWNS
WHO Collaborating Centre on AIDS,
IMET,
Hôpital Claude Bernard,
10 Avenue Porte d'Auvervilliers,
75019 Paris, France

1. Rees, M. *Nature* **326**, 343-345 (1987).
2. Peterman, T. A. *et al. J. Am. med. Ass.* **254**, 2913-2917 (1985).
3. Rao, C. R. *Linear Statistical Inference and its Applications* 353-366 (Wiley, New York, 1965).
4. Lui, K-J. *et al. Proc. natn. Acad. Sci. U.S.A.* **83**, 3051-3055 (1986).
5. Iversen, O-J. & Engen, S. *J. Epid. Comm. Health* **41**, 55-58 (1986).

Table 1   Alternative presentation of Table 2 of Rees

Year of infection	Length of incubation period (years)							
	0–1	1–2	2–3	3–4	4–5	5–6	6–7	
1978	0.07	0.12	0.20	0.32	0.50	0.73	1.06	
1979	0.45	0.79	1.33	2.15	3.33	4.96		
1980	1.38	2.41	4.05	6.53	10.12			
1981	3.27	5.71	9.58	15.45				
1982	6.70	11.69	19.61					
1983	9.11	15.89						
1984	7.00							
Total	27.98	36.61	34.77	24.45	13.95	5.69	1.06	144.51

Expected cases ($\mu = 15, \sigma = 5$) classified according to length of incubation period
Total absolute deviation (TAD) = 31.8

Table 2   Comparison of calculated total absolute deviations for two pairs of parameter values

Parameter values	Total absolute deviation (TAD)	
	(a)	(b)
$\mu = 5, \quad \sigma = 1.6$	35.6	23.2
$\mu = 15, \quad \sigma = 5$	14.7	31.8

(a) cases classified by year of diagnosis (overlapping 2-year periods), as in Rees[1]. (b) Cases classified by length of incubation period, as in Table 1.