Brief communication

Cancer deaths in India: is the model-based approach valid?

P.C. Gupta,[1] R. Sankaranarayanan,[2] & J. Ferlay[3]

The model-based cancer mortality estimates for India (775 800) are nearly double the data-based estimates (433 000), and are higher than even the incidence estimates (612 300). The model-based approach is therefore, at least in the case of India, incorrect. Established practice is to use real data to validate theoretical models, not to reject the data if apparently the model does not fit.

The problem

In their recent article on global and regional cause-of-death patterns in 1990, Murray & Lopez estimated the total cancer deaths for India as 775 800 (*1*). Based on the World Bank estimates of age/sex-specific total deaths and linear regression modelling procedures, they divided the total number first into three broad categories of causes using adjustment procedures for arriving at corrected totals, and then into specific disease groups, one of which was cancer. For India, they used the Survey of Causes of Death (*8*) to check the validity of their division into the three broad categories, adjusting if the estimates differed by more than 10%. They then used cancer registry data to generate site-specific numbers from total cancer deaths, but preferred indirect estimates of cancer mortality, suggested by the models, rather than estimates based on cancer incidence because "the level of underreporting by registries in some areas (e.g., India) seemed... to be sufficiently high to invalidate this approach" (*1*, p. 456).

In India there exists a network of several well-established population-based cancer registries providing reasonable incidence data for many years. The Cancer Registry in Bombay, for example, has operated for 30 years and the registries in other metropolitan areas for over 10 years. Their data have been included in the well-known volumes on *Cancer incidence in five continents*. Regular reports have been published by the individual registries, with details of registration methods and reliability indices including two research papers (*2, 3*). Combined reports have been published by the Indian Council of Medical Research. Murray & Lopez have not provided any reason for their impression of incompleteness, nor any quantification of it. We have compared the Indian data with that from other countries and have not found any evidence of serious underreporting.

Discussion

Cancer registry data were used for estimating the worldwide cancer incidence (*4*) and mortality (*5*) in 1985. In order to produce a national estimate, the age-sex-site-specific cancer incidence rates from the three longest-functioning cancer registries in metropolitan areas in India (Bombay, Bangalore, Madras) and the only rural registry (Barshi) were used with equal weightage (0.25) for each. This may have resulted in some overestimation, as the rates for Barshi were the lowest and 70% of the Indian population is rural. These estimates can easily be updated to 1990 by taking into account the 5-year trends in incidence for each site and age group from Coleman et al. (*6*) and population changes. The results for all sites combined and some selected sites among men and women are displayed in the Table along with the percentage of "underreporting" if the estimates from the models are taken as correct.

It can be seen that the model-based mortality estimates are about twice as high as the data-based estimates, and are higher than even the incidence estimates. Looking at some specific sites the differences seem even more inexplicable; for example, the incidence and mortality estimates for mouth/pharynx cancers in India are already one of the highest in the world. They are comparatively easily diagnosed cancers and it would be rather ludicrous to suggest that

[1] Visiting Scientist, International Agency for Research on Cancer, Lyon, France.

[2] Scientist, International Agency for Research on Cancer, 150 Cours Albert-Thomas, 69372 Lyon Cedex 08, France. Requests for reprints should be sent to this author.

[3] Assistant (Statistics), International Agency for Research on Cancer, Lyon, France.

Reprint No. **5552**

Table 1: **Incidence and mortality estimates (in thousands) for cancer in India during 1990**

	Incidence (Gupta et al.)	Mortality (Gupta et al.)	Mortality (Murray & Lopez)	Percentage "under-reporting"
All sites	612.3	433.0	775.8	44
Male	290.6	229.5	450.2	49
Female	321.7	203.5	325.6	38
Mouth/pharynx	90.1	66.5	157.1	58
Male	61.3	44.4	108.0	59
Female	28.8	22.1	49.1	55
Oesophagus	42.1	40.6	83.2	51
Male	24.4	23.7	51.5	54
Female	17.7	16.9	31.7	47
Colorectal	27.1	18.5	39.5	53
Male	15.1	10.4	22.7	54
Female	12.0	8.1	16.8	52
Lung	30.8	27.4	55.3	50
Male	26.0	23.2	46.4	50
Female	4.8	4.2	7.9	47
Breast	58.3	30.0	40.7	26
Cervix	95.5	43.5	63.0	31

there is nearly 60% underreporting. The model predicts consistently higher underreporting for men rather than women and the smallest underreporting for female breast and cervix cancers. If there were some underreporting, one would expect it to be the other way around. Using different procedures, the National Cancer Registry Programme of India estimated the total number of cases (incidence) in 1990 as 629 000 (*7*), about 2.7% higher than our estimate (612 300).

It seems clear that the model-based approach is, at least in the case of India, incorrect. The justification for using it is presumably to ensure that the proportionate distribution of deaths by cause is similar to that observed in other countries with a similar total (all causes) mortality. Yet there is no reason to suppose that this is necessarily so; it might be, for example, that mortality from cardiovascular disease is more common in India than in other countries at an equivalent level of socioeconomic development. In any case, established practice is to use real data to validate theoretical models, not to reject the data if apparently the model does not fit.

Acknowledgements

We are grateful to D.M. Parkin and P. Pisani for their help in preparing the estimates.

Résumé

Décès par cancer en Inde: l'approche fondée sur les modèles est-elle valable?

En Inde, les estimations de la mortalité par cancer fondées sur les modèles (775 800) sont près du double des estimations fondées sur les données (433 000), et dépassent même les estimations de l'incidence (612 300). Par conséquent, du moins en ce qui concerne l'Inde, l'approche fondée sur les modèles est incorrecte. La pratique usuelle est d'utiliser les données réelles pour valider les modèles théoriques, et non de rejeter les données si le modèle ne semble pas s'y adapter.

References

1. **Murray CJL, Lopez AD.** Global and regional cause-of-death patterns in 1990. *Bulletin of the World Health Organization*, 1994, **72**: 447–480.
2. **Shanta V et al.** Cancer registration methods in Madras Metropolitan Tumour Registry in India. *European journal of cancer*, 1994, **30A**: 974–978.
3. **Yeole BB, Jussawalla DJ.** An assessment of reliability and completeness of Bombay Cancer Registry data (1963–1985). *Indian journal of cancer*, 1988, **25**: 177–190.
4. **Parkin DM, Pisani P, Ferlay J.** Estimates of the world-wide incidence of eighteen major cancers in 1985. *International journal of cancer*, 1993, **54**: 594–606.
5. **Pisani P, Parkin DM, Ferlay J.** Estimates of the world-wide mortality of eighteen major cancers in 1985. *International journal of cancer*, 1993, **55**: 891–903.
6. **Coleman MP et al.** *Trends in cancer incidence and mortality*. Lyon, International Agency for Research on Cancer, 1993 (IARC Scientific Publication No. 121).
7. **National Cancer Registry Programme.** *Annual report 1987*. New Delhi, Indian Council of Medical Research, 1990.
8. **Registrar-General, Government of India.** *Survey of causes of death (rural). Annual Report 1988*. New Delhi, 1990.