A-PDF MERGER DEMO

Università degli Studi di Pisa
Dipartimento di Statistica e Matematica
Applicata all'Economia

Report n.217

Heterogeneity in regional notification patterns and its impact on aggregate national case notification data: the example of measles in Italy

J. R.Williams, P.Manfredi, S.Salmaso, M.Ciofi

Pisa, Novembre 2001

Via Cosimo Ridolfi, 10 – 56124 PISA – Tel. Segr. Amm. 050 945231 Segr. Stud. 050 945317 Fax 050 945375
Cod. Fisc. 80003670504 - P. IVA 00286820501 - Web http://statmat.ec.unipi.it/ - E-mail: dipstat@ec.unipi.it

Heterogeneity in regional notification patterns and its impact on aggregate national case notification data: the example of measles in Italy

John R Williams[1,2], Piero Manfredi[1], Stefania Salmaso[3], Marta Ciofi[3]

[1]Dipartimento di Statistica e Matematica Applicata all'Economia
Università degli Studi di Pisa
Via C Ridolfi, 10
56124 Pisa, Italy

[2]Department of Infectious Disease Epidemiology
Division of Primary Care & Population Health Sciences
Faculty of Medicine
Imperial College of Science, Technology and Medicine
Norfolk Place
London W2 1PG, U.K.

[3]Istituto Superiore di Sanità
Viale Regina Elena, 98
00161 Rome, Italy

Keywords: *measles, case notification, Italy*

Word count: *3283*

Abstract

A monthly time series of measles case notifications exists for Italy from 1949 onwards. The usefulness of this series is seriously undermined by extensive under-reporting. This varies strikingly between regions. Such heterogeneity suggests a possibility of significant distortions in epidemic patterns seen in the aggregated national data set. Estimates of under-reporting are presented here for each Italian region using a simple procedure based upon the approximate equality between births and measles cases. The estimates confirm the impression of universally poor levels of reporting in Southern Italy. These estimates are used to provide a corrected national data set eliminating much potential distortion. The corrected national time series is an order of magnitude greater but, despite great heterogeneity between regions in under-reporting and in epidemic patterns, the shape of the corrected national time series remains close to that of the aggregated uncorrected data. This suggests such aggregate data is quite robust to great heterogeneity in reporting and epidemic patterns at the regional level, and as a result the corrected Italian data set maintains an epidemic pattern which is quite distinct from that of England and Wales.

Introduction

Highly infectious diseases such as measles necessitate vaccination of a very high proportion of the population in a sustained and effective manner in order to achieve control. Therefore national policy makers must possess data which reflect as closely as possible the true scale and pattern of infection, thus allowing vaccination policy to be tuned to achieve maximum impact.

Optimisation of vaccination programme design can be greatly facilitated through insights gained from the mathematical modelling of measles transmission dynamics. Such modelling can shed light on reasons for observed patterns of infection in both pre-vaccination and post-vaccination eras and upon likely effects of the continuation of existing policies. However, modelling work must be based upon sound data if full confidence is to be placed in results and, since epidemiology of specific infectious diseases may differ according to geographic area, the availability of good local data is of great importance. Seroepidemiological surveys provide reliable information about patterns of experience of infection (e.g. Salmaso *et al*, 2000) but, in the absence of continuing large scale representative programmes of serological surveillance, reliance must be placed upon case notification data for information about patterns of infection over time.

Much important analysis has been based on a relatively small number of long data series. These provide a valuable epidemiological resource, but their value relies heavily upon the consistency and the reliability of the relevant systems of case reporting. One European example is measles case notification data from England & Wales, providing raw material for much work in this field (e.g. Fine & Clarkson, 1982; Anderson, *et* al, 1984; Grenfell & Anderson, 1985; Finkenstadt & Grenfell, 2000), and widely considered to be of good quality although, even here, concerns have been raised about under-reporting, and mis-reporting of measles cases remains a danger (Ramsay *et* al, 1997). Elsewhere in Europe, availability and quality of case notification data varies widely between countries and, in some instances, within countries (ref).

In this paper we take Italy as an example of a country possessing a long and potentially valuable time series of measles case notifications, but whose immediate usefulness may be undermined by substantial inconsistencies in the reporting process. In Italy measles has been a notifiable disease since the end of the 19th century (Fig 0) although most regretably data prior to 1949 were apparently destroyed in the 1970's. However, from 1949 onwards data are available by age-group and geographic area. It is widely recognised that throughout this period in Italy there has been considerable under-reporting of measles cases (Salmaso *et al*, 2000). Santoro *et al* (1984), for example, observe that reported incidence may be an order of magnitude less than true incidence. The degree of under-reporting also varies widely between regions; indeed in a few regions notifications are consistently so few as to appear far below the bounds of credibility. With such a heterogeneous pattern of under-reporting occurring side by side with the existence of significant regional heterogeneity in epidemic patterns, there has been concern that there may be distortion of the epidemiological patterns observed in aggregate national data with a consequent effect upon the estimates of epidemiological parameters made from such data.

We have here attempted to overcome these inconsistencies by making use of the observed fact that, in the pre-vaccination era, nearly all were infected by measles well before reaching adulthood (ref). Thus, for each region, there should be an approximate correspondence over time between regional numbers of births and numbers of measles cases. A simple procedure is employed making use of this correspondence to revise regional notifications data in order to provide a corrected aggregate national data set; results are compared with estimates of national case notifications from measles mortality data (for the purposes of further comparison a similar procedure was employed for the national data set of measles cases for England & Wales).

Although in essence this approach to correction of data is not novel, being an extension of that of Clarkson & Fine (1985) (see also Finkenstadt & Grenfell , 2000; Santoro et al, 1984), we believe that what is new is its systematic application to regional datasets to provide corrected aggregate national data and that this approach does provide a much sounder basis for the epidemiological analysis of national data sets. Subsequent publications will discuss the time series analysis of this data set, an investigation of the corrected national age distribution of cases and its variation over time and a theoretical consideration of the potential impact on observed epidemic patterns of the substantial levels of internal migration in Italy that have taken place in the latter part of its pre-vaccination era.

Materials and methods

Source Data

Monthly measles case notification data and annual births data were provided by the Istituto Nazionale di Statistica (ISTAT) for the period from the first available year, 1949, to 1996. Figure 1 shows a standardised regional time series of monthly case notification data for the whole period 1949-1996. Data on measles deaths in Italy and England & Wales were also obtained from published sources.

In Italy vaccination for measles is simply recommended rather than being compulsory (as is also the case for mumps and rubella). Nationally vaccination against measles was first made available towards the end of 1976 and became officially recommended in 1979. There is little data on vaccine uptake but initially it is believed to have been low (Salmaso *et al*, 2000) and, for the first 10 years at least, to have remained at disappointingly low levels (ICONA working group, 1998).

Estimation of degree of under-reporting

Prior to widespread commencement of measles vaccination, most individuals were infected by measles before adulthood **(ref)**. Clearly because of the cyclic epidemic pattern, variation in the age at which individuals are infected, and fluctuations and trends in birth numbers, there could not be an exact correspondence between births and cases. Nevertheless, ignoring migrational flows and assuming endemic equilibrium, during the pre-vaccination period numbers of measles cases over time would have been approximately equivalent to numbers of births, providing a basis for the estimation of the degree of under-reporting (Clarkson & Fine, 1985).

One approach to such estimation would be to select a suitable value for the force of infection (FOI: the per capita yearly incidence of infection in the susceptible proportion of the population) and, by using a simple catalytic model (Anderson & May, 1991), to estimate the equilibrium age distribution of cases and hence distribution of the birth years of those infected in an average year.

Although perhaps theoretically sound, this has two main drawbacks. First, it is quite demanding in terms of births data. Second, it presumes availability of some "suitable" FOI (a possible choice could be the EURO FOI estimated by Edmunds et al. (2000) for a wide range of European countries). This assumption is clearly not neutral (**nor** even tautological), as it implies imposing the structure of the hypothesised FOI upon the very data from which subsequently an estimation of the true FOI would be made. Instead we seek here a correction tool based on minimal assumptions, to minimise possible biases introduced by the correcting algorithm.

In pre-vaccination Europe around half of all measles cases would be expected to occur in the first 4-7 years of life (Edmunds et al, 2000), with numbers of cases at older ages becoming more and more widely distributed with increasing age. In these circumstances it was felt that estimation of under reporting using a moving average to smooth yearly birth data would be more than adequate (Appendix 1 provides a theoretical justification). In the event, values obtained by this means proved sufficiently close to those derived using the catalytic modelling approach above to aditionally justify its use in this instance on grounds of simplicity, greater transparency and ease of application as a general method. A births curve smoothed by averaging over some appropriately delayed period of years (Grenfell *et al*,1995), should encapsulate much of the magnitude and trend of pre-vaccination measles cases. Therefore the ratio of the moving averages of births to that of notifications should provide a suitably robust estimate of the degree of under-notification. Here moving averages of both births and cases for each region were calculated as:

$$y_t = \frac{0.5x_{t-n/2} + x_{t-n/2+1} + \ldots + x_{t-n/2+(n-1)} + 0.5x_{t-n/2+n}}{n}, \quad n = 8.$$

where x_t denotes the original time series of births or case notifications, as appropriate, $n+1$ the number of terms included in the moving average and y_t the moving average series. The estimated level of under-reporting, U_t, was therefore simply:

$$U_t = \frac{C_t}{B_{t-m}}, \quad t = 1957, 1958, \ldots .1982$$

where $C_t = y_{t(births)}$ and $B_t = y_{t(cases)}$. The lag, m, between averages for births and those of cases is intended to take into account the cohorts supplying the greater proportion of cases occurring in any given year (Edmunds *et al*, 1999; Grenfell *et al*, 1995). As an example, in the present instance where a lag of 4 years was chosen (the same as used by Clarkson & Fine (1985)), U_{1957} is a function of birth numbers for the years 1949-57 and case notifications for the years 1953-1961.

Results

For each region, Table 1 shows mean numbers of notifications per thousand live births for the pre-vaccination period 1949-1976, for the first 10 years post-vaccination 1977-1986 (when levels of vaccination cover were believed to be quite low), and for the period 1987-96. To highlight the low numbers of cases shown for some regions, Table 1 also includes national data from England & Wales which demonstrates a 25-fold difference between the pre-vaccination period 1949-1966 and the recent post-vaccination period 1989-96. The latter figure of 22.5 notifications 1000 live births^{-1}, covers a period when vaccine coverage of the order of 80% had been achieved (Ramsey et al, 1994); in contrast the almost identical figure for the Italian region of Campania during the pre-vaccination period 1949-76 is plainly not credible.

Table 1 makes clear the wide variation in numbers of case notifications between regions in both pre- and post-vaccination eras. At its greatest, during the pre-vaccination period, this reaches an order of magnitude difference between Emilia Romagna in the north (209.3 notifications 1000 live births^{-1}) and Campania in the south (22.6 notifications 1000 live births^{-1}). If data for the 4 largest cities (population > 1.000.000) is taken into account, the difference is even greater with Campania's regional capital of Naples providing only 16.0 notifications 1000 live births^{-1}. For the majority of regions there is also clear evidence of a decline in the number of notifications per 1000 births during the pre-vaccination period. These contrasts are maintained on a wider geographic scale with, for example, figures of 107.9 and 34.3 notifications 1000 live births^{-1} in North and South Italy for 1949-76 and, for 1987-96, 76.9 and 34.0 respectively. In contrast, the national notification rate of 572.6 1000 live births^{-1} for England & Wales for the pre-vaccination period 1949-66 is some 8 times greater than that for Italy (73.1) in the comparable period 1949-76.

There is a further strong contrast between Italy and England & Wales in terms of the relationship between recorded measles deaths and notified cases (Figure 8) where the Italian data displays a substantially higher relative number of deaths

Smoothing of births and case notifications

The trends of smoothed case notifications data for the Italian regions (not shown) were more varied than those of births, although generally slightly declining towards the start of vaccination in the mid-1970's, in some cases from a broad peak or plateau and in others with occasional sharper peaks or other fluctuations (in contrast smoothing of the national pre-vaccination notifications data for England & Wales showed a more or less constant number of cases). After the start of vaccination in the mid-1970's, some Italian regions saw a decrease in notifications, but others, strangely, an increase; yet others showed a decrease followed by increase.

Under-reporting factors

For England & Wales, the magnitude of U_t, suggests about 60% of cases are notified, although with a slowly decreasing trend over time. In Italy regional patterns of U_t and its mean values are quite variable for the period 1957-76 prior to vaccination (Table 2). When considering the pattern of *Ut* values by 5-years intervals, most of the regions show a decline over time; this reaches its maximum in Val d'Aosta, where *Ut* decreased from 15.4% in 1952-1956, to 3.8 in 1972-76, but it is particularly evident also in some regions of Southern Italy, such as Campania and Sicilia. Over the whole of the pre-vaccination period the data for Emilia suggests the highest level of reporting, estimated at 21.2% of cases, whereas Campania, the lowest, is estimated to report only 2.0% of cases. Figure 2 makes clear the north-south contrast between regions by providing a summary of these data in map form.

For purposes of further comparison Table 3 shows U_t for Abruzzo, a region in the middle range of estimates of under-reporting, compared with estimates of under-reporting where weights of the moving average for births are derived using a catalytic model together with various estimates of the force of infection as discussed above(Edmunds et al, 2000); a further comparison was made using weighting based upon the 1971-76 age distribution of actual reported cases. The simple moving average, U_t, compares well with the estimates using the more complex weighting systems estimates. Figure 3 compares the actual national age distribution of case notifications for the years 1971-76 with birth year distributions (i.e. age distributions) suggested by FOI-based estimates and with that of the simple approach used here. It is clear therefore that FOI-based estimates do not necessarily provide a closer correspondence to the age distribution of cases (but see caveat in Discussion concerning potential age bias in reporting of cases).

Corrected case notification data for Italy

Figure 4 shows corrected aggregate national notification data where, prior to aggregation, the appropriate mean under-reporting correction is applied to the data for each region. Although there is great variability in magnitude of the correction for the different regions, and the magnitude of the overall correction is very large (Figure 4a), the effect of the correction on yearly epidemic patterns is remarkably small with very close correspondence between the shapes of uncorrected and corrected data, and the timing of epidemic peaks is changed by, at the most, one month or so (Figure 4b).

Although this adjusted Italian case data is now of similar magnitude to that of similarly corrected data for England & Wales, the corrected Italian data continues to show a strongly distinct pattern of roughly 3-year periodicity compared with that of 2-years seen for England & Wales (Figure 5). The clear pattern of a single seasonal peak occurring in March or April is also maintained (Figure 6); this occurs a month or so later than that seen in weekly pre-vaccination data for England & Wales (Anderson, Grenfell & May, 1984) which also shows small troughs in January and April, absent, or not clearly apparent, from the monthly Italian data. The results of the present work suggests that such national shapes and trends in notifications data can be quite robust to very substantial variation in under-reporting between regions. It may however be the case that detailed time series analysis might reveal differences between corrected and uncorrected data series that are not immediately apparent from visual inspection; this issue will be addressed in a forthcoming paper.

Discussion

The degree of under-reporting seen in **Italian** regional notification data of the pre-vaccination era is in some instances, strikingly large, so that corrected aggregate national data for Italy is an order of magnitude greater than for the uncorrected data. Nevertheless, although they vary widely between regions. shapes, location and trends of adjusted and unadjusted national data are very similar. Although the proportion of cases notified may be poor, it seems likely that deaths arising from measles cases would be more reliably reported, so that a relatively high ratio of deaths to cases might be expected in the circumstances described. Figure 8 confirms this expectation by comparing notified deaths and cases for the period 1958-77 with those observed in England & Wales over similar and earlier timescales (Ramsey *et al*, 1994). It may be argued that the higher Italian 'deaths to cases' ratios may to some extent be a result of poor living conditions or standards of health in certain regions, and an indication of the ratio from one developing country (Peru) is also included for further comparison (Sniadack *et al*, 1999). A comparison of infant mortality rates for Italy (52.7 and 35.6 1000^{-1} live births in 1951-60 and 1961-70 respectively (Livi-Bacci, 1990)) with those of England & Wales (29.1 and 17.7 1000^{-1} live births in 1950-2 and 1970-2 respectively(US Census Bureau, 1993) and Peru (47.48 1000^{-1} live births in 1995 (US Census Bureau website)) suggests that these factors may have played a part but it is undoubtedly the case that the high Italian 'deaths to cases' ratio also reflect the very poor levels of case reporting illustrated above (Salmaso *et al*, 2000) (the term 'case fatality rate' is not used here because of the unreliability of case reporting).

Although the relationship between births and cases clearly begins to break down with increasing levels of vaccination, in absence of evidence to the contrary, it is reasonable to assume that levels of under-reporting seen in the pre-vaccination era might continue after the start of vaccination. Thus correction of post-vaccination data using mean adjustment factors calculated for the whole or part of the data from pre-vaccination years should prove more satisfactory than reliance upon uncorrected data which there is every reason to suppose substantially under represents true numbers of cases and distorts the contribution of individual regions. Whether it is more appropriate to extrapolate the trend seen in pre-vaccination era correction factors or simply to employ a correction factor calculated for the immediately pre-vaccination years is a matter for debate. The former might appear more soundly based but, until underlying social and administrative factors determining the degree of under-reporting are better understood, it would be more prudent to adopt corrections derived from the latter part of the pre-vaccination era. However a number of factors in the post-vaccination period may potentially disturb the magnitude of or trend in under-reporting, as a result of changing perceptions arising directly from institution of a vaccination programme. For example, such a programme, by underlining the public health impact of measles, might encourage increased reporting; alternatively, reporting might decrease as a result of complacency arising from the very existence of a vaccination programme. Also it has been argued that, as measles becomes less common with increasing vaccination, there is an increased tendency to misdiagnose simply as a result of lack of experience in identifying measles cases; as this might lead to both an increase in under-reporting and an increase in the number of cases of other infections misreported as measles, its overall impact is not clear (Blackburn *et al*, 2000; Ramsey *et al*, 1997)

A possible additional factor in Italy affecting the apparent pattern of under-reporting during the pre-vaccination period is the extremely high volume of systematic internal migration during the period 1950 to 1975, combined with the age distribution of susceptible individuals amongst migrant families. As a result of depletion of the pool of susceptibles in the exporting regions, use of the birth rates to correct notifications for those regions could have given rise to an over-estimation of the true number of cases. The extent of this effect is, of course, modulated by the age distribution of infection in those regions; theoretical aspects of this issue will be addressed in a further publication which will suggest that, although migrations significantly affect the true number of cases, the validity of under reporting estimates is very little affected.

Overall, though corrected age-aggregated notification data for Italy are certainly much closer to true numbers of cases, the issue raised by Edmunds *et al* (2000) of age bias in reporting of measles cases in Italy remains to be considered. If there is indeed a substantial increase in under-reporting with decreasing age, there would be an overestimation of the average age of infection and underestimation of the force of infection in the youngest age groups. This may in turn impact on health policy both directly and, indirectly, through the effect on parameter estimation for modelling work informing health policy. It is therefore of paramount importance that investigations be undertaken to address this issue and the authors are aware of a new study under the auspices of the Istituto Superiore di Sanità which will cast valuable light upon this.

Clearly there is a need to improve overall levels of case notification within the regions of Italy. However under-reporting appears to be of such dramatic extent, in some regions at least, that exhortations to medical practioners to notify all cases seem unlikely to achieve the desired result; further research will need to be undertaken if we are to understand fully the reasons for such poor levels of reporting and if sound strategies for optimising the impact of vaccination are to be developed.

Overall, despite qualifications outlined above, it is believed that the adoption of the procedure described here provides, by eliminating potential distortions arising from wide variation in regional reporting levels, both i) a national time series of measles data, and ii) a national age distribution of cases, of improved reliability which will prove a useful basic resource to aid further understanding of the epidemiology of measles infection, both within Italy and in the wider European context.

References

Anderson RM, May RM *Infectious diseases of humans: dynamics and control.* Oxford: OUP; 1991.

Anderson RM, Grenfell BT, May RM. Oscillatory fluctuations in the incidence of infectious disease and the impact of vaccination: Time series analysis. *Journal of Hygiene.* 1984; **93**(3):587-608.

Blackburn N. Schoub B. O'Connell K. Reliability of the clinical surveillance criteria for measles diagnosis. *Bulletin of the World Health Organization.* 2000; **78**(6):861

Clarkson JA, Fine PEM The efficiency of measles and pertussis notification in England & Wales. *International Journal Epidemiology.* 1985; **14**(1):153-68.

Edmunds WJ. Gay NJ. Kretzchmar M. Pebody RG. Wachmann H The pre-vaccination epidemiology of measles, mumps and rubella in Europe: Implications for modelling studies. *Epidemiology & Infection.* 2000; **125**(3):635-650

Fine PEM, Clarkson JA Measles in England and Wales: I, an analysis of factors underlying seasonal patterns. *International Journal Epidemiology.* 1982; **11**:5-14.

Finkenstadt BF, Grenfell BT Time series modelling of childhood diseases: a dynamical systems approach. *Appl Statist.* 2000; **449**(2):187-205.

Finkenstadt B, Grenfell BT Empirical determinants of measles metapopulation dynamics in England and Wales. *Proc. R. Soc. London* B. 1998; **265**,211-220.

Gomes MC. Gomes JJ. Paulo AC Diphtheria, pertussis, and measles in Portugal before and after mass vaccination: A time series analysis. *European Journal of Epidemiology.* 1999; **15**(9): 791-798.

Grenfell B, Bolker B, Kleczkowski A Seasonality, demography and the dynamics of measles in developed countries. In: Mollison D, editor. *Epidemic models: their structure and relation to data.* Cambridge: CUP; 1995. p.248-68

Grenfell BT, Anderson RM The estimation of age-related rates of infection from case notifications and serological data. *Journal of Hygiene (Cambridge)* 1985; **95**:419-36.

ICONA working group (1998). ICONA: national infant vaccination coverage. Rapporti ISTISAN 98/33 Istituto Superiore di Sanità (in Italian)

Ramsey M, Gay N, Miller E, Rush M, White J, Morgan-Capner P *et al.* The epidemiology of measles in England and Wales: rationale for the 1994 national vaccination campaign *CDR Review* 1994; **4**(12):R141-R146

Ramsay M. Brugha R. Brown D. Surveillance of measles in England and Wales: Implications of a national saliva testing programme. *Bulletin of the World Health Organization.* 1997; **75**(6): 515-521.

Salmaso S, Gabutti G, Rota MC, Giordano C, Penna C, Mandolini D *et al* Pattern of susceptibility to measles in Italy. *Bulletin of the World Health Organization.* 2000; **78**(8): 950-5.

Santoro R, Ruggeri FM, Battaglia M, Rapicetta M, Grandolfo ME, Annesi I, *et al* Measles Epidemiology in Italy. *International Journal Epidemiology.* 1984; **13**(2):201-9.

Sniadack DH. Moscoso B. Aguilar R. Heath J. Bellini W. Chiu MC Measles epidemiology and outbreak response immunization in a rural community in Peru. *Bulletin of the World Health Organization.* 1999; **77**(7):545-552

US Census Bureau, 1993 Statistical Abstract of the United States: 1993 (113th edition) Washigton DC, US Government Printing Office

APPENDIX: using births as estimators for cases

We consider the simplest SIR compartmental model describing infection in a stationary homogeneously mixed population in absence of vaccination:

$$\begin{aligned}\dot{X}(t) &= B(t) - (\mu + \lambda(t))X(t)\\ \dot{Y}(t) &= \lambda(t)X(t) - (\mu + v)Y(t) \qquad (1)\\ \dot{Z}(t) &= vY(t) - \mu Z(t)\end{aligned}$$

where $X(t)$, $Y(t)$, $Z(t)$ represent numbers of susceptible, infective and recovered immune individuals at time t, $B(t)$ births per unit time, $\lambda(t)=\beta Y(t)$ the force of infection (FOI), μ the mortality rate, and v the recovery rate. The total population $N=X+Y+Z$ is assumed stationary with $B=\mu N$. At equilibrium the numbers of births (B) and cases (C) respectively are $B=(\mu+\lambda)X$ and $C=\lambda X$, so that: $C = B - \mu X$. A prediction of (1) is that the susceptible fraction at equilibrium, X/N, is the reciprocal of R_0 (the basic reproduction ratio of the infection (Anderson & May, 1991). Thus: $C = B - \mu N / R_0 = B(1-(1/R_0))$. So for very high values of R_0 such as for pre-vaccination measles, and a sufficiently low mortality rate, it approximately holds C/B≈1. This relation is basically the consequence of the fact that the age window over which the forces of mortality and infection operate are separated. It is worth noting that the basic relation between births and cases holds in more general circumstances. Suppose that there is a steady oscillation of (1) around its long term equilibrium, as is typical of periodically forced SIR and SEIR models. In this case the relation C/B≈1 remains correct provided we consider average values of cases and births over the appropriate time period. More generally if the model is non-stationary, provided state variables are bounded in the long term, the equilibrium condition (time derivatives of the state variables are zero) may be replaced by the weaker condition that their long term averages are zero: $E(\dot{X}(t)) = E(\dot{Y}(t)) = E(\dot{Z}(t)) = 0$, so that $E(C)=E(B)-\mu E(X)$ indicating that the basic relation still holds in a broader sense, i.e. if yearly numbers of cases and births are replaced by their averages over a sufficiently long period.

The previous relations are preserved when age-structure is explicitly introduced. Consider a SIR model with stationary population and no additional mortality arising from disease

$$\begin{aligned}&\frac{\partial X}{\partial a} + \frac{\partial X}{\partial t} = -(\mu(a) + \lambda(a,t))X(a,t) \quad X(0,t) = B(t)\\ &\frac{\partial Y}{\partial a} + \frac{\partial Y}{\partial t} = \lambda(a,t)X(a,t) - (\mu(a) + v)Y(a,t) \quad Y(0,t) = 0 \qquad (2)\\ &\frac{\partial Z}{\partial a} + \frac{\partial Z}{\partial t} = vY(a,t) - \mu(a)Z(a,t) \quad Z(0,t)\end{aligned}$$

where X, Y, Z, λ, μ are as in equation (1) with the added dimension of age, a.

With a stationary population $B(t)=B$, and assuming equilibrium, the total number of cases of infection per unit time is:

$$C=\int_0^\infty \lambda(a)X(a)da = B\int_0^\infty \lambda(a)p(a)\Lambda(a)da \qquad (3)$$

where $p(a)$, $\Lambda(a)$ denote respectively survival to death and infection functions. Under Type I mortality ($\mu(a)=0$, $a<L$, $\mu(a)=\infty$, $a\geq L$), commonly used to approximate mortality in industrialised developed countries:

$$C=B\int_0^L \Lambda(a)\lambda(a)da = B(1-\Lambda(L)) \qquad (4)$$

where $F(L)=1-\Lambda$ is the fraction who have experienced the disease by age L. Since for measles in the pre-vaccination era $F(L)$ is usually very close to 100%, again the relation $C\approx B$ strongly. In the special case of homogeneous mixing by age one finds in particular $C=\lambda X = B\left(1-e^{-\lambda L}\right)$.

Using the approximate relation $A=1/\lambda$ (holding under Type I mortality (Anderson and May, 1991) one gets $C\approx B\left(1-e^{-L/A}\right)=B\left(1-e^{-R_0}\right)$ showing that yearly numbers of cases at equilibrium are quite close to yearly numbers of births. Under Type II mortality ($\mu(a)=\mu$) one recovers (2). A possibility offered by (3) for estimating the degree of under notification of measles cases under more general circumstances appears if one explicitly introduces time, showing that the total number of cases is a weighted average of past births

$$C(t)=(1-F(L))\int_0^L B(t-a)G(a)da \qquad (5)$$

the weights being given by the (normalised) density of infection: $G(a)=\Lambda(a)\lambda(a)/(1-\Lambda(L))$; this is the catalytic approach mentioned in the text. As long as births do not fluctuate wildly, cases can be reconstructed from births provided a suitable force of infection is assumed. The simplified approach used in this paper follows by considering time averages of (6) over a generic interval (t_0,t_1):

$$\overline{C}(t_0,t_1)=\frac{1}{t_1-t_0}\int_{t_0}^{t_1} C(t)dt=\frac{1}{t_1-t_0}\int_{t_0}^{t_1}\left(\int_0^L B(t-a)G(a)da\right)dt \qquad (6)$$

By the mean value theorem a value ζ exists such that

$$\overline{C}(t_0,t_1)=(t_1-t_0)^{-1}\int_{t_0}^{t_1} B(t-\zeta)\left(\int_0^L G(a)da\right)dt=(t_1-t_0)^{-1}\int_{t_0}^{t_1} B(t-\zeta)dt \qquad (7)$$
$$=\overline{B}(t_0-\zeta,t_1-\zeta)$$

(7) suggests that a moving average of births, delayed by a suitable choice of ζ, may be used to reconstruct the time series of present (moving averaged) cases. The literature (Edmunds *et al*, 2000, and references therein) suggests that possible shapes of the age densities of infection for measles in pre-vaccination regimes are sufficiently well behaved so that the choice of ζ is easy. Compared to (6), (8) has the advantage that assumptions about the full shape of the force of infection function are unnecessary.

We finally point out that the relationship between cases and births is robust to changes in basic model assumptions, extending to more general models such as SEIR's. Moreover, if the population is exponentially increasing rather than stationary (typical of developing countries) it holds in a more general form involving discounted births rather than actual births.

Table 1. Mean number of notifications of measles cases 1000 live births^{-1} for the specified period for Italy by region (north to south), broad geographic region (North, Central, South (see Fig 2)) and for England & Wales Specific data is not available for the major cities: Milan in Lombardia, Turin in Piemonte, Rome in Lazio, and Naples in Campania; these are represented by their provinces (the major part of whose populations are found in these cities), with additionally mean figures for the 4 regions concerned excluding these provinces.

		1949-76	**1977-86**	**1987-96**
Piemonte		89,7	93,0	93,6
	Turin	*63,5*	*81,5*	-
	Excl. Turin	*116,9*	*107,0*	-
Valle D'aosta		94,2	45,8	30,1
Lombardia		86,3	103,2	89,4
	Milan	*63,5*	*94,5*	-
	Excl. Milan	*102,0*	*109,8*	-
Trentino A.A.		70,9	94,6	114,7
Veneto		72,3	86,0	89,4
Friuli Venezia G.		168,0	229,0	147,7
Liguria		134,4	134,5	67,7
Emilia Romana		209,3	221,3	92,3
Toscana		116,8	147,1	99,6
Umbria		129,7	99,4	58,9
Marche		113,6	127,9	87,7
Lazio		77,2	56,4	36,0
	Rome	*82,4*	*61,9*	-
	Excl. Rome	*65,2*	*43,6*	-
Abruzzo		80,1	61,0	53,3
Molise		83,0	76,4	85,6
Campania		22,6	11,6	22,8
	Naples	*16,0*	*7,1*	-
	Excl. Naples	*30,4*	*17,5*	-
Puglia		28,3	33,7	46,3
Basilicata		48,5	45,3	51,3
Calabria		33,7	15,6	14,8
Sicilia		28,3	13,1	27,2
Sardegna		57,3	39,9	23,5
	NORTH	108,9	123,1	76,9
	CENTRE	88,2	71,9	41,7
	SOUTH	30,4	20,2	34,0
	ITALY	72,6	69,7	56,0
England & Wales		572,6		22,5
		(1949-66)		(1989-96)

Table 2. Estimates by region of percentage of measles cases notified in Italy for the specified periods (i.e. 100 U_t with a 4 year lag between moving averages of births and cases)

	1957-61	**1962-66**	**1967-71**	**1972-76**	**Pre-vaccination mean**
Abruzzo	8,6	8,8	7,4	5,8	7,9
Basilicata	4,9	4,8	4,2	3,2	4,4
Calabria	3,9	3,4	2,7	2,1	3,1
Campania	2,9	2,3	1,5	1,0	2,0
Naples	2,2	1,7	1,1	0,6	1,5
Excl. Naples	3,5	2,9	2,0	1,6	2,6
Emilia Romana	22,8	22,3	19,9	18,4	21,2
Friuli Venezia G.	20,7	20,0	15,8	12,7	17,9
Lazio	9,8	8,3	6,9	4,7	7,8
Rome	11,4	8,7	7,2	4,7	8,4
Excl. Rome	7,1	7,2	6,0	4,7	6,5
Liguria	16,1	14,3	11,1	10,4	13,2
Lombardia	11,0	10,0	7,6	6,3	9,0
Milan	8,9	7,6	5,7	4,7	7,0
Excl. Milan	12,2	11,5	9,1	7,6	10,4
Marche	12,4	11,4	10,6	9,5	11,2
Molise	9,2	11,0	7,7	5,2	8,6
Piemonte	12,8	9,8	7,5	5,7	9,4
Turin	10,2	7,2	5,3	4,3	7,0
Excl. Turin	14,6	12,4	10,0	7,7	11,6
Puglia	3,0	3,4	2,7	2,1	2,9
Sardegna	6,4	5,1	4,3	3,3	5,0
Sicilia	3,4	3,2	2,4	1,6	2,8
Toscana	15,2	14,1	12,4	11,8	13,5
Trentino A.A.	7,1	7,8	6,7	6,2	7,0
Umbria	15,5	13,5	12,0	7,9	12,8
Valle D'aosta	12,3	8,4	6,3	3,6	8,2
Veneto	8,4	8,1	6,5	5,3	7,3

Table 3: Estimated levels of measles case notification 1960-1974 arising from the use of different weighting systems for the moving average of numbers of births using as an example the region of Abruzzo

Weighting system	**% cases reported**
Catalytic model using force of infection (FOI) from Italian seroprevalence data	8.9
Catalytic model using FOI from case notifications data	8.7
Catalytic model using EURO FOI (see text)	8.4
Age distribution of cases 1971-76	8.7
Simple moving average	8.3

Figure Legends

Figure 1. Monthly measles case notifications for 1949-1996 standardised by regional mean. Regions are shown in customary geographic order and data for each region are standardised by their mean. (NB as the city of Trieste was not incorporated into the region of Friuli Venezia Giulia until 1954, when the city's postwar status within Italy was finally resolved, 2 time series are shown for Friuli V. G.: up to 1954, and 1954 onwards,).

Figure 2. Map showing estimated mean percentage of measles cases reported by region for the pre-vaccination period 1953-1976

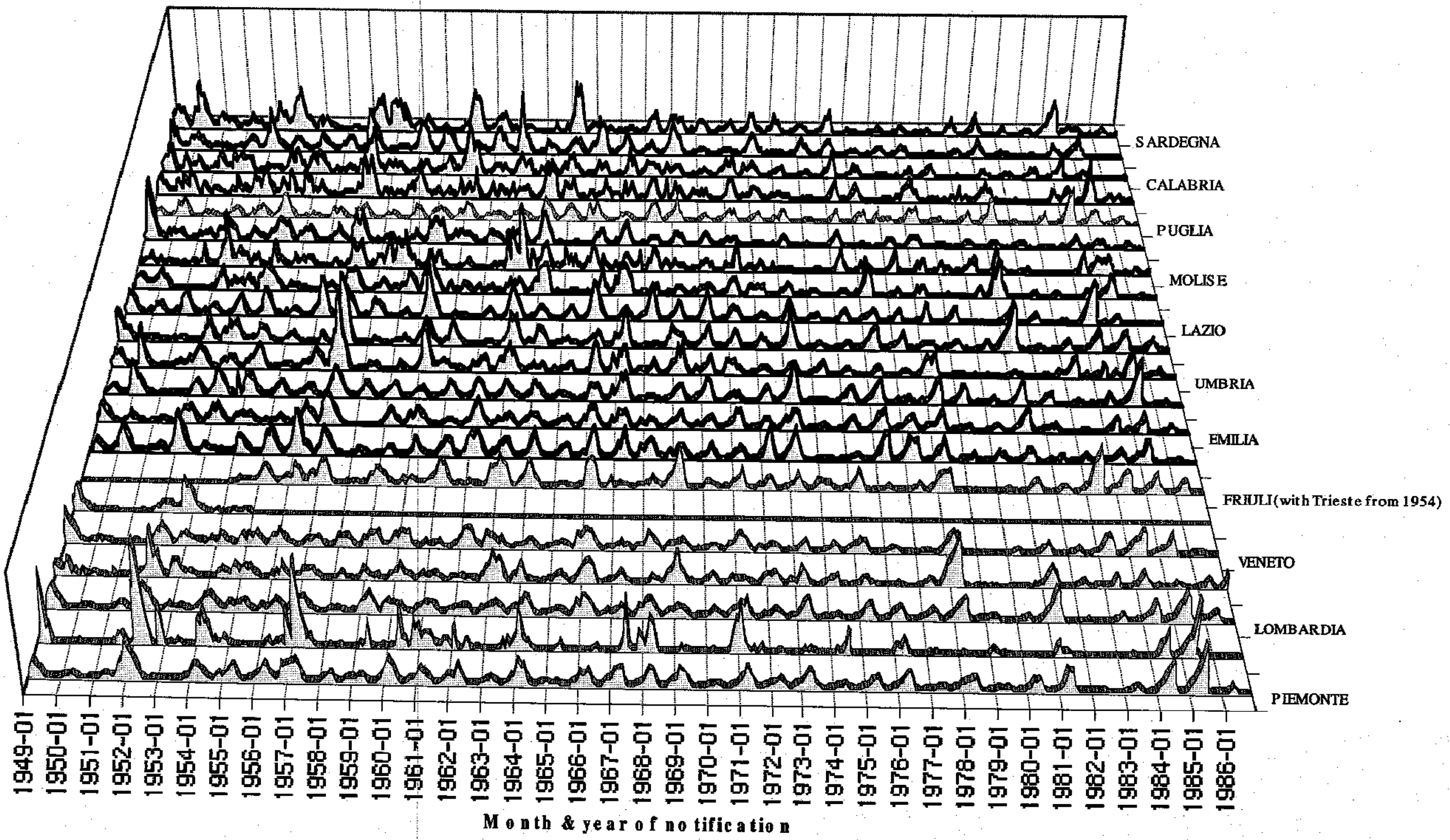
Figure 3. Age distribution of measles case reports for Italy for 1971-76 (thick solid line) compared with age distribution of weights of 9-year moving average (thick dashed line), and those of the age distributions suggested by force of infection (FOI) estimates of Edmunds *et al* (2000) made, respectively, from seroprevalence data (Δ), case notifications (O) and that suggested by the composite 'EURO FOI' of those authors corresponding to a number of European countries (■)

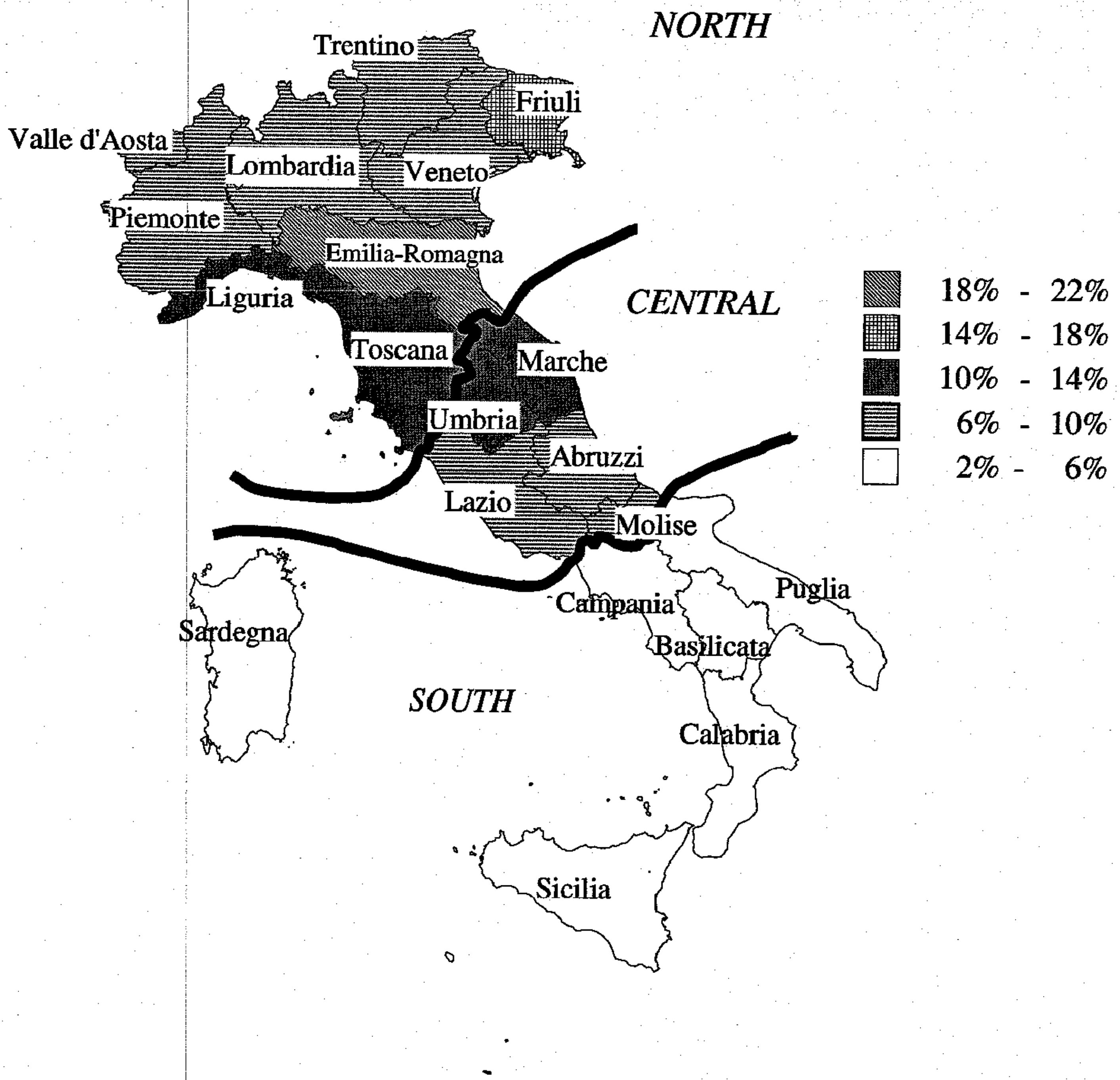
Figure 4. Comparison of (a) unadjusted national Italian measles notification data (solid line) and nationally aggregated regional notification data for the years 1953-82 adjusted using the pre-vaccination mean value of U_t from Table 2 (dashed line); (b) the same data but with the uncorrected national data (solid line) scaled to be of the same magnitude as that of the aggregate adjusted regional data (dashed line).

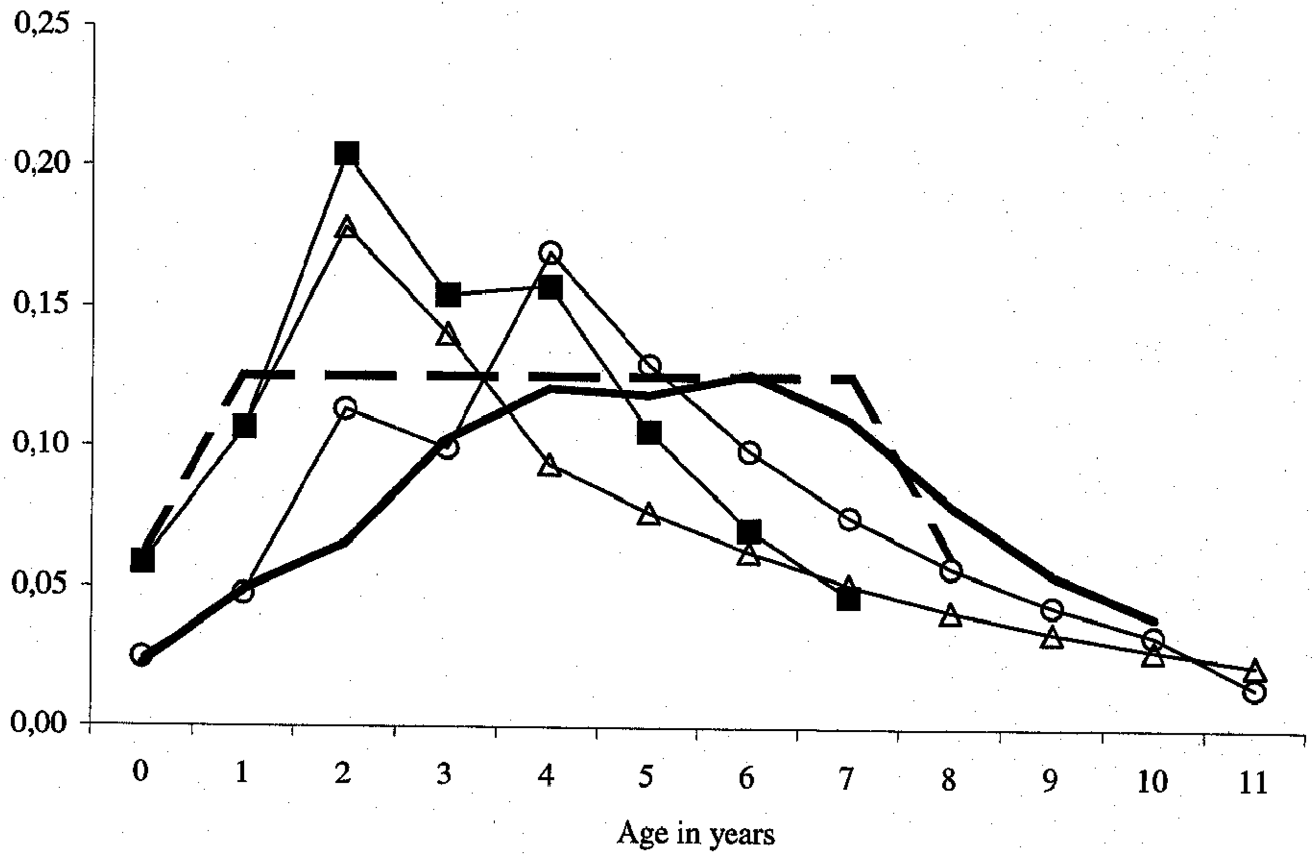
Figure 5. Comparison of Italian nationally aggregated regional yearly notification data (open and solid triangles) adjusted for under reporting with similarly adjusted data for England & Wales (open circles). Adjustments are by moving average of births as described in the text (solid triangles) and by the mean of the moving average of births for the period 1953-82 (open triangles and circles)

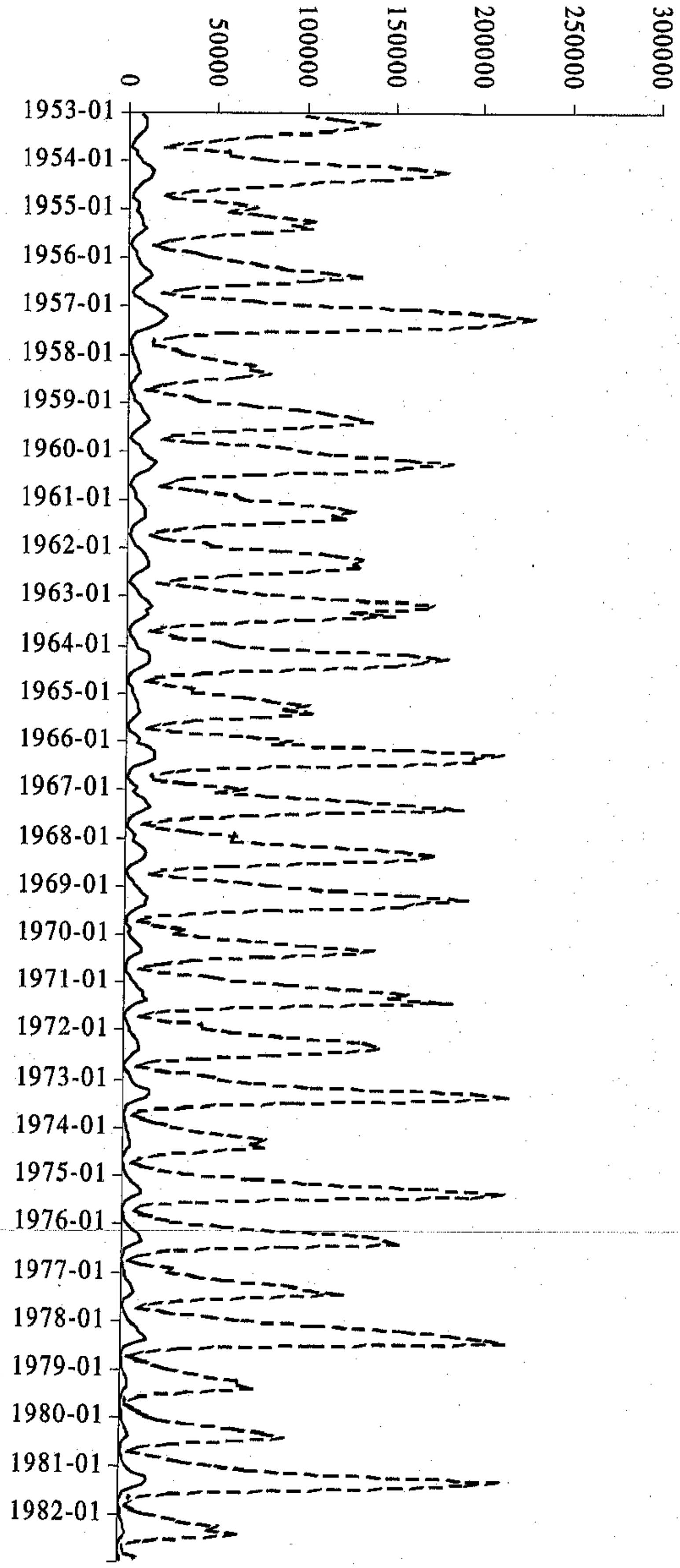
Figure 6. Adjusted nationally aggregated regional notification data by month and year of notification for the period 1953-82.

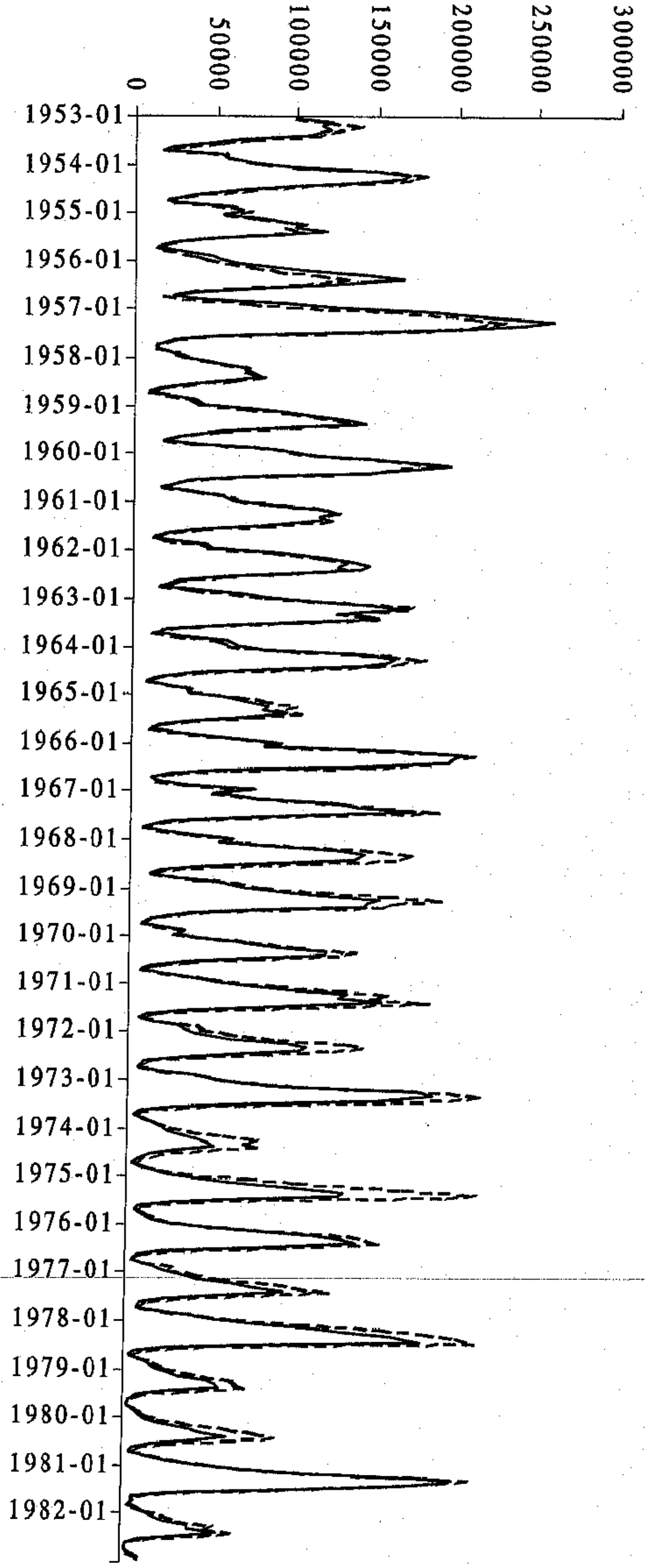
Figure 7. Relationship between notified measles deaths and notified cases (unadjusted) for the Italian regions for the period 1958 -1977 compared with the approximate range in this relationship for England & Wales for the late 1940's and early 1950's (solid line) and for the late 1950's to early 1960's (dotted line). Also shown (dashed and dotted line) is an example of this relationship from a recent outbreak in a developing country (Peru)(Sniadack *et al*, 1999).

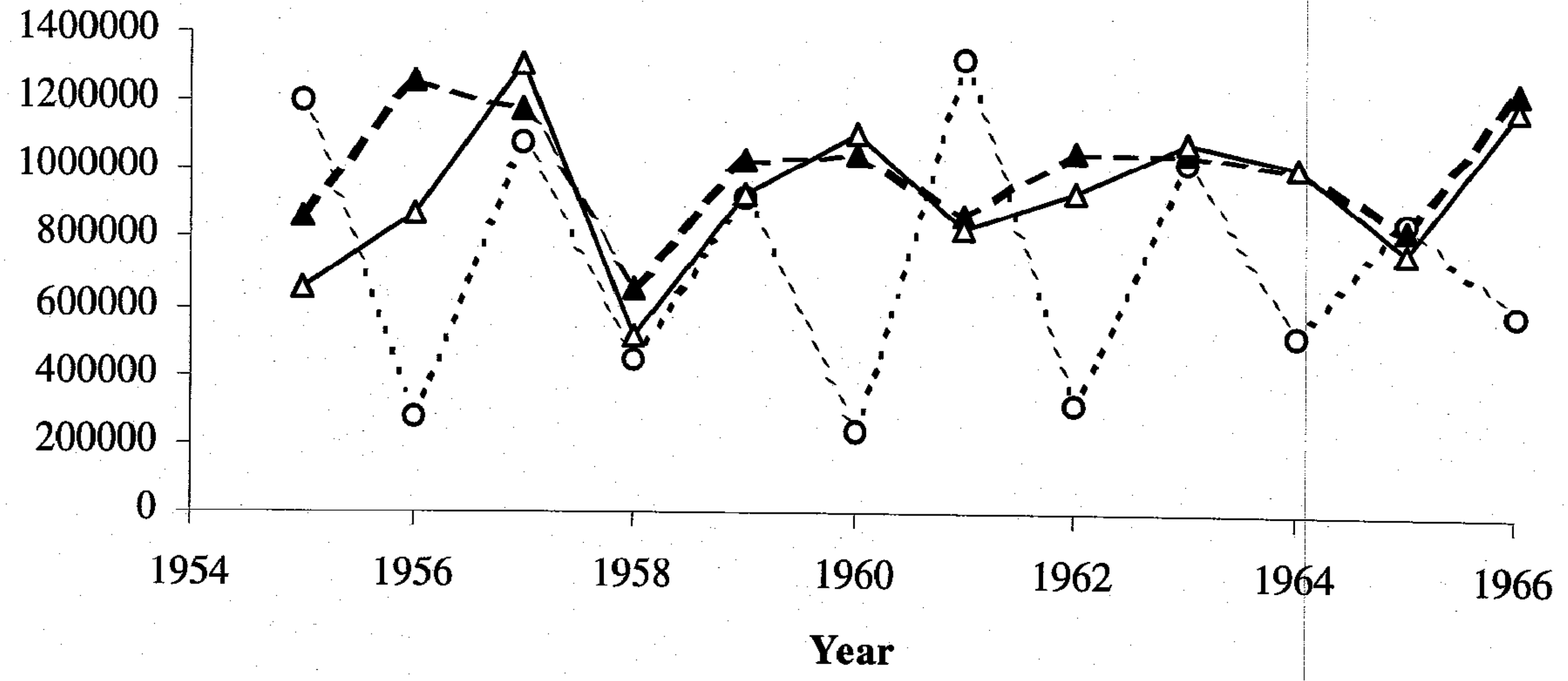

SARDEGNA
CALABRIA
PUGLIA
MOLISE
LAZIO
UMBRIA
EMILIA
FRIULI (with Trieste from 1954)
VENETO
LOMBARDIA
PIEMONTE
1949-01
1950-01
1951-01
1952-01
1953-01
1954-01
1955-01
1956-01
1957-01
1958-01
1959-01
1960-01
1961-01
1962-01
1963-01
1964-01
1965-01
1966-01
1967-01
1968-01
1969-01
1970-01
1971-01
1972-01
1973-01
1974-01
1975-01
1976-01
1977-01
1978-01
1979-01
1980-01
1981-01
1982-01
1983-01
1984-01
1985-01
1986-01
Month & year of notification

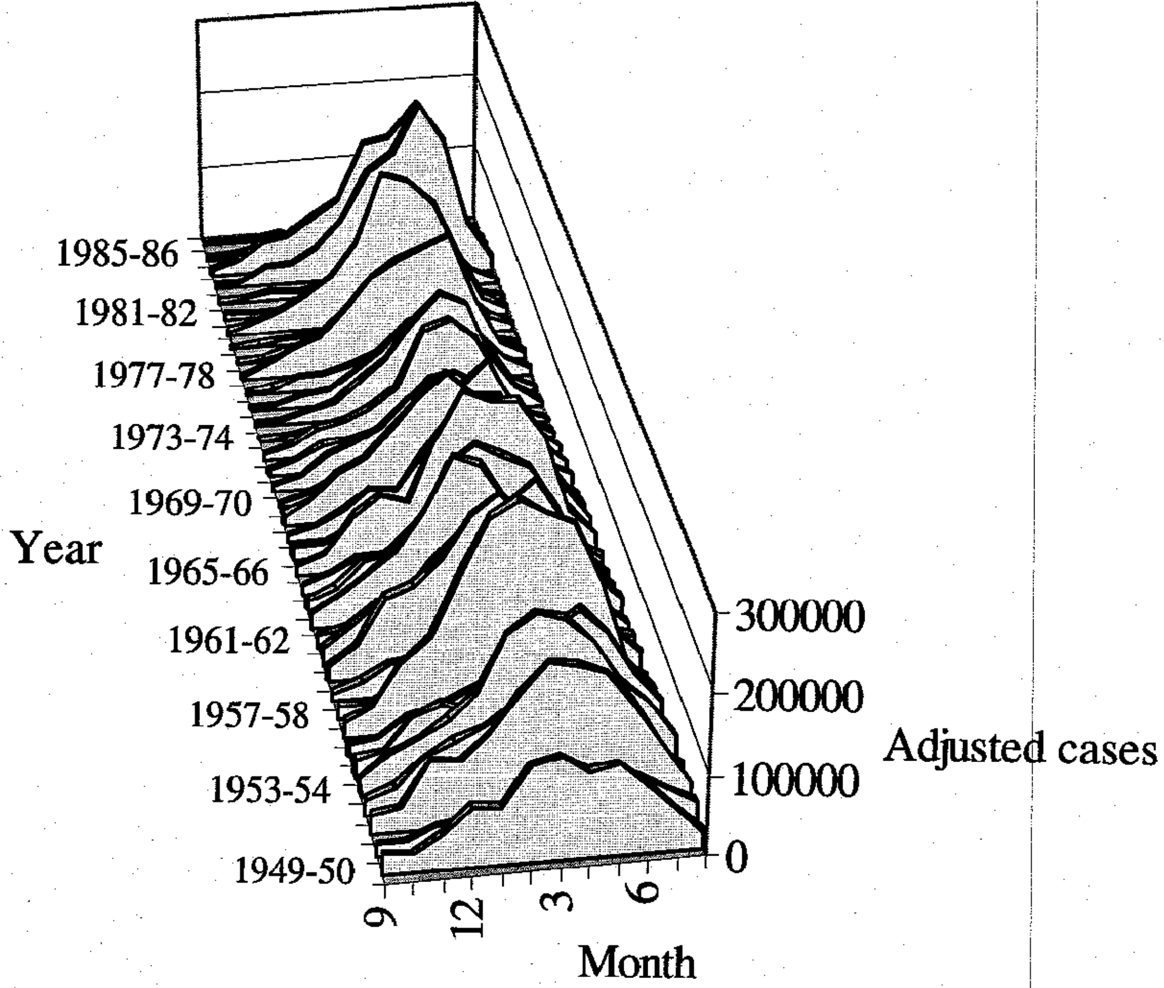

NORTH
Trentino
Friuli
Valle d'Aosta
Lombardia
Veneto
Piemonte
Emilia-Romagna
Liguria
CENTRAL
Toscana
Marche
Umbria
Abruzzi
Lazio
Molise
Puglia
Campania
Sardegna
Basilicata
SOUTH
Calabria
Sicilia
▨ 18% - 22%
▦ 14% - 18%
■ 10% - 14%
▤ 6% - 10%
□ 2% - 6%

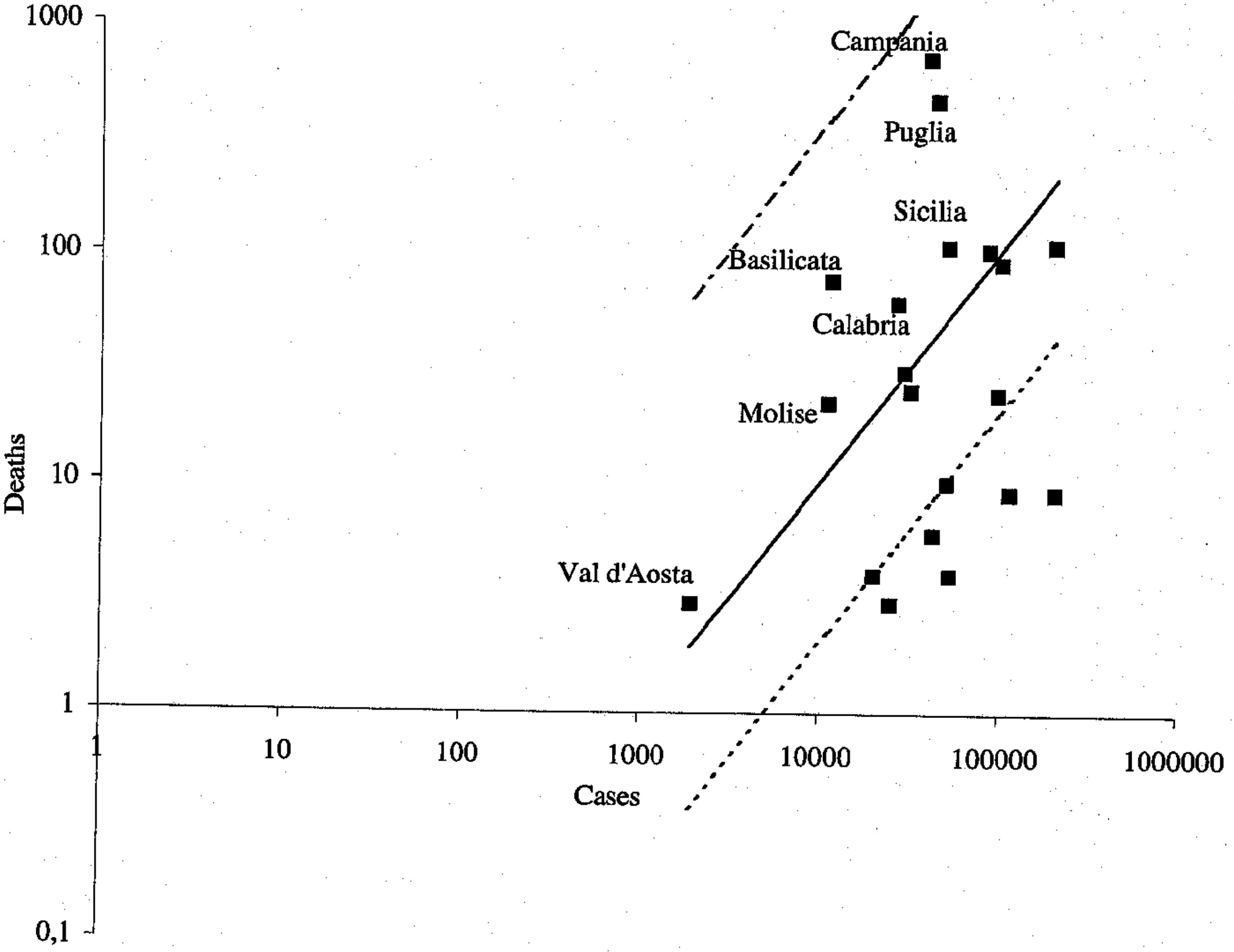

0,25
0,20
0,15
0,10
0,05
0,00
0
1
2
3
4
5
6
7
8
9
10
11
Age in years


0
50000
100000
150000
200000
250000
300000
1953-01
1954-01
1955-01
1956-01
1957-01
1958-01
1959-01
1960-01
1961-01
1962-01
1963-01
1964-01
1965-01
1966-01
1967-01
1968-01
1969-01
1970-01
1971-01
1972-01
1973-01
1974-01
1975-01
1976-01
1977-01
1978-01
1979-01
1980-01
1981-01
1982-01


0
50000
100000
150000
200000
250000
300000
1953-01
1954-01
1955-01
1956-01
1957-01
1958-01
1959-01
1960-01
1961-01
1962-01
1963-01
1964-01
1965-01
1966-01
1967-01
1968-01
1969-01
1970-01
1971-01
1972-01
1973-01
1974-01
1975-01
1976-01
1977-01
1978-01
1979-01
1980-01
1981-01
1982-01


1400000
1200000
1000000
800000
600000
400000
200000
0
1954
1956
1958
1960
1962
1964
1966
Year


1985-86
1981-82
1977-78
1973-74
1969-70
Year
1965-66
1961-62
1957-58
1953-54
1949-50
9
12
3
6
Month
300000
200000
100000
0
Adjusted cases


1000
100
10
1
0,1
Deaths
1
10
100
1000
10000
100000
1000000
Cases
Campania
Puglia
Sicilia
Basilicata
Calabria
Molise
Val d'Aosta