a long term decline or increase in the disability incidence rates. Over the 20th century, French male mortality rates declined substantially, and we would argue that it is quite implausible to assume that the mortality rates of disabled men and women have declined along with those of non-disabled men. In the French male population, life expectancy has not remained unchanged, and it is not reasonable to expect outflows from the disabled state to remain unchanged. We made the arbitrary assumption that the relative risk of death for disabled versus non-disabled people remained constant at 3.66 in the published simulations. We also carried out simulations where the relative risk ranged from 1 to 10 and also varied with age. We found that the results were not highly sensitive to the value assumed and did not affect our conclusions.

In response to criticism of the constant relative risk assumption, we have carried out additional simulations in which we assume that the ratio of the mortality risk for disabled male to the average mortality risk for all males remains constant at its value in the year 1945, and so is independent of changes in disability prevalence. This had very little effect on the estimated health expectancies under any of the scenarios we considered. Our published comment that the results are not highly sensitive to assumptions about the mortality rates of disabled people.

We intended no implication as to the plausibility of the scenario of Barendregt et al by describing ‘hypothetical disease example’; that was the description they themselves used.1 We have no disagreement with their scenario or its results, but repeat that such a simulation is not relevant to the question of how accurate Sullivan’s method is for monitoring at the whole of population level. Thrombolytic therapy may well have been introduced in a three year period in The Netherlands, and may well have had a major effect on survival rates and a detectable effect on cardiovascular mortality at population level. Despite this, in-hospital or post-hospital mortality after myocardial infarction does not account for most cardiovascular deaths (in Australia around 80% of myocardial infarction deaths occur outside hospital).

Figure 1 shows long term trends in mortality rates for The Netherlands for all causes mortality and ischaemic heart disease mortality. Rates are age standardised using the European standard population and five year averages are shown prior to 1985. It is clear that the impact of thrombolytic therapy has not caused any sudden change in all cause mortality rates (the relevant rates for use with Sullivan’s method) at the population level, and that Sullivan’s method would be entirely adequate for monitoring long term trends in Dutch health expectancies. Very few changes in treatment practice would result in such dramatic changes in transition rates as the “hypothetical example” of Barendregt et al, and it is very unlikely that sudden changes in all causes mortality or disability transition rates at the population level will result from new medical interventions.

Finally, to the question of monitoring compression or expansion of morbidity. The example in the letter above is based on the assumption that compression ceases, allowing the disability prevalence in the population to reach its equilibrium value. We have no disagreement with the conclusion that in such a case the Sullivan’s method would give a spurious compression—it is another example of the limitation of Sullivan’s method when there are sudden changes in transition rates. In a more realistic example, where disability incidence and prevalence evolve smoothly, Sullivan’s method will give a reasonably good indication of whether compression or expansion is occurring. In our published scenarios, Sullivan’s method provides quite accurate estimates of the degree to which compression or expansion is occurring.

In conclusion, we do not believe we are in disagreement over the usefulness and limitations of Sullivan’s method or that the example of Barendregt et al contradicts our conclusion that when population health is evolving reasonably smoothly, Sullivan’s method is acceptable.
the study of the paper has revealed that the reference on which the calculation of the concordance between the continuous variables considered in the study is based contains a possible error.

The ICC is the proportion of total variability accounted for by the variability among subjects. A high ICC means that not much of the variability is due to variability in measurement on different occasions (proxy versus true control, in this case) or, what amounts to the same thing, that the agreement between them is high. The ICC for the common situation in which there are two observations for each subject can easily be calculated from an analysis of variance (ANOVA), as shown in the example in table 1 (formula 1). As the ANOVA may be somewhat burdensome to carry out, Deyo et al. [4] offered a simple method for calculating the ICC (see formula 2 in table 1).

The calculation of the ICC by Poulter et al. seems to be based on a commonly cited review article on statistics of concordance, which unfortunately omitted n in the denominator of the original equation for ICC proposed by Bartko. Results based on this mistaken formula underestimate the observed agreement as shown in table 1 (formula 3).

Formulas 1 to 3 assume a random effects model. Another model must nevertheless be distinguished in making inferences about the factors affecting the measurements—the fixed effects model. In the random effects model, the two raters implied in the measurement are assumed to be a sample from a larger population of raters. In the fixed effects model, the two raters are the only raters about which inferences will be made. Although the calculation of ICC assuming a fixed effects model is also based on the ANOVA table, the formula is different (see formula 4 in table 1). Poulter et al. have also proposed a simpler method for calculating the ICC in this case (see formula 5 in table 1).

The calculations of Poulter et al. on the ICC assume a random effects model. Since proxy and "true" controls are the only raters about which inferences were made, we believe that it would be more correct to assume a fixed effects model. Nevertheless, more information concerning the selection of cases and "true" vs. proxy controls will be necessary to strengthen this hypothesis.

Since important differences in results may be obtained, future studies that consider the agreement of continuous variables must be extremely careful in selecting formulas to estimate the appropriate ICC.

**Table 1** Estimation of agreement between two measures (A and B) through the intraclass correlation coefficient (ICC). Hypothetical data considering two measures in 10 patients

<table>
<thead>
<tr>
<th>Subject</th>
<th>Measure A</th>
<th>Measure B</th>
<th>Difference (B-A)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>75</td>
<td>80</td>
<td>5</td>
</tr>
<tr>
<td>2</td>
<td>74</td>
<td>84</td>
<td>10</td>
</tr>
<tr>
<td>3</td>
<td>76</td>
<td>81</td>
<td>5</td>
</tr>
<tr>
<td>4</td>
<td>79</td>
<td>83</td>
<td>4</td>
</tr>
<tr>
<td>5</td>
<td>82</td>
<td>92</td>
<td>10</td>
</tr>
<tr>
<td>6</td>
<td>83</td>
<td>88</td>
<td>5</td>
</tr>
<tr>
<td>7</td>
<td>85</td>
<td>90</td>
<td>5</td>
</tr>
<tr>
<td>8</td>
<td>87</td>
<td>92</td>
<td>5</td>
</tr>
<tr>
<td>9</td>
<td>87</td>
<td>93</td>
<td>6</td>
</tr>
<tr>
<td>Mean</td>
<td>81.6</td>
<td>87.5</td>
<td>5.9</td>
</tr>
<tr>
<td>SD</td>
<td>28.1</td>
<td>25.4</td>
<td>4.8</td>
</tr>
</tbody>
</table>

**ANOVA table on preceding data**

<table>
<thead>
<tr>
<th>Source of variation</th>
<th>DF</th>
<th>SS</th>
<th>MS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between subjects</td>
<td>9</td>
<td>459.45</td>
<td>51.05</td>
</tr>
<tr>
<td>Within patients</td>
<td>1</td>
<td>174.05</td>
<td>174.05</td>
</tr>
<tr>
<td>Residual R</td>
<td>9</td>
<td>21.45</td>
<td>2.38</td>
</tr>
<tr>
<td>Total</td>
<td>19</td>
<td>654.95</td>
<td></td>
</tr>
</tbody>
</table>

**Formulas for the calculation of the ICC**

**Random effects**

**ICC Bartko**

\[
\text{ICC} = \frac{n}{n-1} \left(1 - \frac{\sum(SS_\text{MS})}{\sum(SS_\text{MS} + SS_{\text{MSR}})}\right)
\]

**formula 1**

**ICC Deyo et al.**

\[
\text{ICC} = \frac{\sum(SS_\text{MS}) - \sum(SS_\text{MSR})}{\sum(SS_\text{MS}) - \sum(SS_\text{MSR}) + \sum(SS_\text{MSR})}
\]

**formula 2**

**ICC Kramer and Feinstein**

\[
\text{ICC} = \frac{\sum(SS_\text{MS}) - \sum(SS_\text{MSR})}{\sum(SS_\text{MS}) + \sum(SS_\text{MSR})}
\]

**formula 3**

**Fixed effects**

**ICC Fleiss**

\[
\text{ICC} = \frac{\sum(SS_\text{MS}) - \sum(SS_\text{MSR})}{\sum(SS_\text{MS}) - \sum(SS_\text{MSR}) + \sum(SS_\text{MSR})}
\]

**formula 4**

**ICC Poulter et al.**

\[
\text{ICC} = \frac{\sum(SS_\text{MS}) - \sum(SS_\text{MSR})}{\sum(SS_\text{MS}) - \sum(SS_\text{MSR}) + \sum(SS_\text{MSR})}
\]

**formula 5**

**Reply**

Dr. Prieto et al. are correct that the formula to calculate intraclass correlation coefficients (ICC) included in Kramer and Feinstein’s review is misquoted and understates the observed agreement. However, while we included the reference to Kramer and Feinstein as background information on ICC, in our recent evaluation of the reliability of proxy-derived data in a case-control study, we used the formula included in Bartko’s original paper.

We agree that the random or fixed effects model affects the estimated ICC and our choice of this random effects model is not ideal where raters were not necessarily representative of a larger population of raters. However, the ICC calculated using Bartko’s formula do not differ importantly from those calculated using Fleiss’s formula, assuming a fixed effect model (table 1). The only three exceptions to this finding are for comparisons of true and proxy data when only two pairs are involved (mother/daughter and sister/sister). In such a situation all estimates are unstable.

**Table 2** Comparison of ICC estimates using random and fixed effects models (comparisons relate to data shown in reference 2)

<table>
<thead>
<tr>
<th>Subject</th>
<th>Random effects</th>
<th>Fixed effects</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Weight</td>
<td>Height</td>
</tr>
<tr>
<td></td>
<td>1.00</td>
<td>0.90</td>
</tr>
</tbody>
</table>

**Letters, Notices, Corrigenda, Books reviews**

NOTICES

European Journal of Public Health—volume 7 (2) June 1997

Editorial
Balancing state and market in health system reform R B Saltman

Editorial note
J Palm, P G Svensson

Original articles
Use of provocative emotional appeals in a mass media campaign designed to encourage smoking among adolescents. A Hofstad, B Stray-Pedersen, F Langmark
Managing the gap between demand and publicly affordable health care in an ethical way. J A Drevet
Adolescent motherhood and socioeconomic factors: an ecologic approach. M Nebot, C Borrell, J R Villalba
Impaired cognitive function in elderly exposed to benzodiazepines or other anxiolytics. J Ransamt, J M Merlo, G Blennow, S Hansson, P O Ostergren, A Melander
Cost-effectiveness of vaccination against pneumococcal pneumonia in the Netherlands. R M P M Baltussen, A J H A Ament, R M Leidl, R van Furth
Sick-leave and disability pensions among female assembly workers. B Palsson, V Horn mann, R G Attekwall, K Ohlsson, S Skenfjeld

Cancer
Socioeconomic group, occupation and incidence of breast cancer and genital cancer among women in Denmark. B A Rix, T Skov, E Lynge

AIDS/HIV
Experience with and attitudes towards HIV patients among Italian general practitioners. A C Botti, M Cesa-Bianchi

Methods
Severity measurement using a generic instrument: a feasibility study in ambulatory care involving patients with diabetes or asthma. M Eccles, N Steen, A Hutchinson, C Bradshaw, E McColl

BOOK REVIEWS


Responding to the needs of older people probably remains the biggest challenge facing health and social care, to quote the author himself. This report examines the process of joint commissioning of services by health and local authorities in five contrasting local areas. Much of the work of the project has been recorded in a series of previous King's Fund papers; the present report draws out the key lessons.

Monitoring the process of service development is a slippery task, especially when a multiplicity of agencies is involved. As the report acknowledges, significant changes will take time, often longer than an individual researcher is able to follow them. So here we have an account of work in process, rather than nearly defined historical episodes.

Ponton presents grounds for modest optimism. Achievements differ in the various localities, but rigorous evaluation of planning initiatives is seldom easy, and short term perspectives may be misleading. The qualities of vision and passion, here identified as essential, are particularly difficult to pin down.

One complication to the success of joint commissioning which the report neglects relatively is that of the different occupational settings of the participants. GPs often point out that social services departments have high staff turnover; their own positions are relatively stable, while allowing them consider able scope in deciding the margins of their job. In a primary care led NHS, their contribution to processes like joint commissioning will be increasingly influential.

BERNARD INIECHEN
Lecturer in Epidemiology
Charing Cross and Westminster Medical School, London


Populations are ageing in almost every country, and this book begins to fill a large void in the literature of health in the elderly. There are 45 chapters with an impressive list of contributors. Two thirds are from the UK and the rest from eight different countries. Many are established authorities on their chosen subject. As Margot Jefferys says in her editorial preface, this should become a standard and reference book.

Chapter subjects range well beyond consideration of individual diseases, to encompass, for example, health economics, community care, migration and ethnicity, and istrogenesis. Its scope aims (and generally succeeds) in being truly international. The editors have succeeded in keeping contributors to a uniform chapter length of about nine pages. They have been less successful in achieving a universal standard of up to date ness. However, given the explosion of knowledge in the past decade, and the ease with which it can be retrieved, this is a serious point. There are dangers ahead too—volumes like this age quickly in matters of detail.

Nevertheless, Epidemiology in Old Age is a splendid achievement. The price is too steep for individuals, but medical libraries will find it of great value. It is extremely well produced with a vivid purple cover. Specialists will find their particular concerns examined in more detail elsewhere, but for those who want to venture into new aspects of geriatric epidemiology, their search begins here.

BERNARD INIECHEN
Lecturer in Epidemiology
Charing Cross and Westminster Medical School, London