JIA 113 (1986) 103-149

A STATISTICAL STUDY OF THE VARIABILITY OF SICKNESS DATA

BY R. H. DAW, B.Sc., F.I.A.

The aim of science is to seek the simplest explanation of complex facts . . . seek simplicity and distrust it.

A. N. WHITEHEAD (Bibby, 1983)

1. INTRODUCTION

1.1. IN 1979 the first set of graduated sickness rates derived from the data supplied by a number of insurance companies was published by the PHI Sub-Committee (*C.M.I.R.* **4,** 1979).

Sickness is a much more complex phenomenon than mortality. I get the impression from working on the 1972–75 sickness experience of, as it were, a certain looseness and lack of coherence in the data which I have not found in mortality data. Some of this could be due to the much smaller size of the data and to the heterogeneity of the experience. Much of the latter has been removed in the Standard experience of 1975–78 but, not having studied these data, I can say nothing about the effect of this. It is therefore not surprising that the PHI Sub-Committee had considerable doubts "about the possibility of judging the success of the graduation by applying standard statistical tests" such as those normally used for mortality data. They rejected the validity of the χ^2 test "in view of the lack of knowledge regarding the distribution of sickness rates by age, and the lack of independent events within duration of sickness claims" (*C.M.I.R.* **4,** 1979, pp. 17 and 19).

1.2. It is the purpose of this paper to study the variations in the Sickness Experience of 1972–75 for individual PHI policies (*C.M.I.R.* **4,** 1979) for male lives. The data for female lives are too small for this. An attempt is made to get some idea of the magnitude and distribution of the variability of the rates of claim inceptions and of sickness. The Report gives the crude data for age-groups only but the PHI Sub-Committee kindly supplied me with tabulations of the figures for individual ages.

The initial study of the sickness data will be made by applying to it the r_x test of Redington and Michaelson (1940) which was originally devised for investigating mortality experiences (Daw, 1945, 1974 and 1982, Beard, 1951 and Jager, 1953). As possible confirmation of the results a similar study was made of the Manchester Unity 1893–97 (Whole Society) sickness experience. (Watson, 1903).

1.3. In applying the r_x test to sickness rates a number of new, mainly technical, matters need to be considered. Some of these are dealt with in Appendices A and B; this is not because they are unimportant, but in order not to interrupt the flow of the general description of the work. It is hoped that the text of the paper will

enable a reader, who is prepared to accept the technical matters on trust, to get a general understanding of the work and the results without studying the Appendices. Other readers may study the Appendices either before or after reading the paper—or both.

1.4. The numerical tables in this paper relate to the 1972–75 sickness experience of individual policies on male lives unless the heading states otherwise. Reference to a particular deferred period and sickness period will often be in the form, e.g. D4, S13/13, which denotes deferred period 4 weeks, sickness period 13 weeks deferred 13 weeks. The numbering of paragraphs, formulae and tables commences with the number of the section of the paper in which it occurs or, in the case of the Appendices, with A or B.

2. FEATURES OF THE DATA

2.1. Before proceeding further there are some features of the 1972–75 sickness experience which need to be described.

Quantity of data available

2.2. In considering the results of this study, it should be borne in mind that the 1972–75 sickness experience virtually ceases at age 64 and that 30 is the youngest age covered by the official graduations of the sickness rates; for certain sickness periods the commencing age is later. Thus the largest age range to which the r_x test could be applied was 35 years. This is about half the range usually available in a mortality experience.

Further, in mortality experiences to which the r_x test has been applied the exposed to risk at individual ages is usually a 5- or 6-figure number, whereas the average figure per age in the 1972–75 experience is in most cases below 4,000. (See also Table 6.1). These limitations on the size of the data available mean that variations in the indices calculated can be quite large.

2.3. The tabulations supplied to me show the number of claim inceptions for deferred periods 1, 4, 13 and 26 weeks in the age range 30–64 for the male 1972–75 experience as being 10,371, 2,462, 576 and 269 respectively. In the case of sickness rates these numbers of inceptions will apply only to the first period of sickness after the end of the deferred period. For later sickness periods the number of claims on which the sickness rates are based will progressively decrease, so that some of the later sickness periods must be based on a very small number of claimants. Table 3.2.3(a) of *C.M.I.R.* **4** (1979) gives an indication of these figures but the text of § 3.2.3 must be studied carefully.

Duplicate policies

2.4. The 1972–75 experience contains duplicate policies and one of the effects of duplicates is to increase the variance of mortality rates, sickness claim inception rates and sickness rates. *C.M.I.R.* **4** (1979) gives no indication of the extent to which duplicates are present but *C.M.I.R.* **7** (1984) Appendix F, gives

the results of an analysis of the duplicate policies included in the sickness inceptions of the 1975–78 experience. I understand that there is no evidence to suggest that the proportion and distribution of duplicates in the 1975–78 experience would not also apply approximately to that of 1972–75. However in the discussion of the 1975–78 experience, it was suggested that the high inflation during that period might have resulted in an increase in the number of duplicate policies taken out, in order to keep the total sickness benefit in line with inflation.

2.5. Appendix F, of *C.M.I.R.* 7 (1984) indicates that the data for deferred period 1 week contains an average of 1·6 policies per life, and that deferred periods 4, 13 and 26 each contain about 1·1 policies per life. From this information, Appendix F determined factors for the increase in the standard deviation of claim inception rates due to the presence of duplicates.

2.6. In Daw (1984), it is shown that the increase factors for duplicates in inception rates should also apply to the standard deviation of the various sickness rates. This assumes that both the average number of policies per life and the distribution at inception apply also to policies on which no claims were made, and that these are not modified in later sickness periods by say selective lapsing. The factors for the increase in the standard deviation of claim inception rates and sickness rates were:

1·5 for deferred period 1 week
1·1 for deferred periods 4, 13 and 26 weeks

However, as suggested in § 2.4, it must not be forgotten that these factors could be on the high side.

Correlations between deviations at successive ages

2.7. The term 'deviations' is used to describe the deviations of an ungraduated sickness rate (z_x) from the corresponding true underlying rate, usually taken as the graduated rate (z'_x). In statistical literature the term 'error' is more often used for this purpose.

2.8. The method used in the C.M.I. sickness experiences is to define the year of age as the calendar year, the age x being taken as the nearest age on the 1 January. If a sickness claim continues past the 31 December, then the period of sickness after that date is allocated to the age $x+1$. This will introduce positive correlation between the deviations at age x and age $x+1$. No correlation will be introduced between deviations at age x and age $x+2$ or later ages, except in the case of sickness period 104/all where the duration of a claim can exceed one year. The longer the sickness period the more likely is a claim to run past the year end, so that the correlation introduced might be expected to increase with length of sickness period.

2.9. Some contributing offices define age in a different way from that described in § 2.8, which necessitates their claims being allocated half to one age and half to the next. This can introduce positive correlation of a similar nature to that described in the previous paragraph.

Heterogeneity

2.10. In §3.1 of *C.M.I.R.* **4** (1979), mention is made of "the evident heterogeneity of the data". By heterogeneity is meant that the data contain groups of lives subject to differing rates of sickness, for example the differences between one office and another (see Table Se 1.1.4 of *C.M.I.R.* **4**). It is sometimes argued that heterogeneity within ages reduces the variance of sickness and mortality rates. Beyond mentioning that, if the sampling is random, such heterogeneity will not affect the variance, the question will not be discussed further, because I feel sure that any possible effect which heterogeneity *within* individual ages could have on the variance is far too small to be detected in the sickness data dealt with in this paper. (Those interested in the matter might read, (i) Kendall, p. 197 in the discussion of Daw (1945), (ii) Coward (1949) p. 21(d) and Perks, p. 31, in the discussion, and (iii) Daw (1974) section 2). Significant heterogeneity *between* ages would show up as an increase in the value of σ_r obtained from the r_x test.

2.11. A feature of the data which might also be regarded as heterogeneity is the rapid increase in the volume of PHI business which has taken place before and during the period of the investigation. Table Se 1.1.1 of *C.M.I.R.* **4** (1979) shows that the number of policies in the male sickness data increased by over 40% from 1 January 1972 to 31 December 1975. While this is of importance as regards the uses to which the sickness rates are put, it seems unlikely to have much effect on the results of the present study. This may, perhaps, not be so for sickness period 104/all where, for example, claims exceeding 4 years in length must all have come from the smaller numbers of policies effected before 1972.

3. APPLICATION OF THE r_x TEST

The r_x *test*

3.1. The r_x test of Redington and Michaelson (1940) is described in Appendix A, where the various formulae are given. Briefly, as applied to mortality, sickness claim inception and sickness rates, it involves calculating the third differences of the ungraduated rates at individual ages. Each third difference is then divided by its theoretical standard deviation to obtain a series of values of r_x. The reason for taking third differences is that, for underlying rates of this type, differences of the third order are small, so that the third differences are composed almost entirely of the deviations (defined in §2.7) in the ungraduated rates.

For mortality and claim inception rates there is a formula for the standard deviation of the rate based on the binomial distribution, but for sickness rates there is no such simple formula. Appendix B discusses how V_x, the variance of the sickness (in weeks) of one individual aged x, may be calculated. (The corresponding standard deviation is $\sqrt{V_x}$.)

3.2. Having calculated the values of r_x for each individual age the standard deviation of these values, σ_r, is then calculated. If all the various assumptions on

which the r_x test is based are correct, then the values of σ_r should be close to unity. However a number of factors can affect the values of σ_r actually obtained:

(i) As stated in §2.6 the presence of duplicates will increase the standard deviation of sickness rates and the values of σ_r will be increased by the same factor.
(ii) If there is correlation between the deviations at successive ages, as described in §§2.8 and 2.9, then σ_r will no longer be an estimate of the standard deviation of the deviations in the original data (see §A9).
(iii) Errors or approximations in the data or method of constructing the table of rates can also cause an increase in σ_r (Daw, 1982).

3.3. Separate values of σ_r have been calculated in respect of the 1972–75 sickness experience for:

(i) claim inception rates for each of the deferred periods 1, 4, 13 and 26 weeks, and
(ii) sickness rates for the same deferred periods dealing separately with sickness periods 1/3, 4/9, 13/13, 26/26 and 52/52 weeks.

No calculations have been made for deferred period 52 weeks because the data are too sparce. Consideration of sickness period 104/all is deferred until Section 7.

Tables 3.1, 3.3 and 3.4 give the values of σ_r obtained; these are in respect of ages 30–64 (i.e. 32 values of r_x) with a few exceptions indicated in Tables 3.3 and 3.4 where the graduated sickness rates do not commence until an age later than 30.

Claim inception rates

3.4. Allowing for duplicates by the factors given in §2.6, the values expected for σ_r would be about 1·5 for deferred period 1 week and 1·1 for the other three deferred periods. These agree well with those shown in Table 3.1 except for deferred period 26 weeks for which σ_r is 1·54. This might perhaps be accounted for by the small number of sickness claims in this deferred period for which the figures given in §2.3 indicate an average number of claims per age of less than 8.

3.5. As explained in §A8 the process of taking third differences results in

Table 3.1. *Values of σ_r for claim inception rates*

Deferred period (weeks)	σ_r
1	1·54
4	1·18
13	1·08
26	1·54

Table 3.2. *Serial correlation coefficients between values of* r_x *for claim inception rates*

Deferred period (weeks)	ρ_1	ρ_2	ρ_3
1	−·89	+·66	−·47
4	−·59	−·05	+·29
13	−·81	+·50	−·37
26	−·81	+·45	−·19
Average of deferred periods 4, 13 and 26 weeks	−·74	+·30	−·09
Theoretical values by formula (A10)	−·75	+·30	−·05

Note: The average coefficients have been calculated by means of Fisher's z transformation (Fisher, 1941, §35) which is also used to obtain any other such averages given in this paper.

successive values of r_x not being independent. There will be correlation between r_x and r_{x+1}, r_x and r_{x+2}, r_x and r_{x+3}; the coefficients measuring this are called the serial correlation coefficients of lag 1, lag 2 and lag 3 respectively and have been denoted by ρ_1, ρ_2 and ρ_3. Table 3.2 gives the values of these coefficients calculated from the values of r_x on which σ_r is based.

The averages shown in Table 3.2 of the serial correlation coefficients for deferred periods 4, 13 and 26 weeks compare very well with the theoretical values, bearing in mind the large variations between the individual coefficients and that the average is over only three values. The figures for deferred period 1 week have been omitted from the average shown because they look rather larger than the others and, as is shown later in the paper, the data for this deferred period are suspect (§§4.3–4.5).

3.6. The results of the r_x test give little reason to think that the variation of the inception rates of the 1972–75 experience, differs from that on the assumption of binomial variation, provided allowance is made for duplicates. Thus the usual statistical tests applied to mortality experiences would seem valid for inception rates.

Sickness rates

3.7. Tables 3.3 and 3.4 set out the values of σ_r in respect of sickness rates. In most cases two values of σ_r are shown based on different methods of calculating the variance of sickness; these are described in Appendix B. Method I is that of Coward (1949) and assumes that no duplicates are present. Method II is based on

Table 3.3. *Values of σ_r for sickness rates*

Deferred period (weeks)	Sickness period (weeks) 1/3 I	1/3 II	4/9 I	4/9 II	13/13 I	13/13 II	26/26 I	26/26 II	52/52 I	52/52 II	Average for all sickness periods except S1/3 method I	
1	1·74	1·57	·91	·95	·58	·60	·93	·95	·93*	N.A.	·84	
4			·92	·95	·84	·86	·83	·87	1·04	N.A.	·91	
13					1·10	1·15	·93	·97	·75	N.A.	·93	·93
26							1·11	1·14	·87†	·88†	·99	
Average	1·74	1·57	·92	·95	·84	·87	·95	·98	·90			

Notes: I and II denote the method used to calculate the variance of sickness (see Appendix B).

* Based on ages 35–64, i.e., 27 values of r_x.

† Based on ages 40–64, i.e., 22 values of r_x.

the moments of the duration of sickness claims given in *C.M.I.R.* **4** (1979). As the experience included duplicates I expected that the resulting variances of sickness would be those taking account of duplicates. On this reasoning the value of σ_r by method I would include the increase due to duplicates and would therefore be higher than the corresponding value of σ_r by method II.

Looking now at Tables 3.3 and 3.4 it will be seen that σ_r (I) is in fact less than σ_r (II) in all but three cases. The only case where the numerical difference is greater than ·05 is D1, S1/3 where σ_r (I) is substantially greater than σ_r (II), i.e., 1·74 compared with 1·57, which might be attributed to the effect of duplicates, although according to the figures in § 2.6 the expected values would be 1·5 and 1·0. The safest course seems to be to say that, for some unknown reason method II has failed in its intention of taking account of the effect of duplicates on the sickness rates.

Table 3.4. *Values of σ_r for sickness rates using data for ages 40–64, i.e., 22 values of r_x*

Deferred period (weeks)	Sickness period (weeks) 52/52 I	52/52 II	104/all I	104/all II
1	·97	·96	N.A.	·47
4	1·16	1·15	N.A.	·58
13	·67	·70	N.A.	·56
26	·87	·88	N.A.	·47
Average	·92	·92	—	·52

Note: For meaning of I and II see note to Table 3.3.

Combined effect of duplicates and correlation

3.8. The values of σ_r by method I would be expected to be somewhere near unity if all the assumptions were correct. However § 3.2 describes some features which tend to produce a value of σ_r different from unity. Two of these will now be considered:

(i) The presence of duplicates will increase the value of σ_r by a factor, suggested in § 2.6 as 1·5 for deferred period 1 week and 1·1 for the other three deferred periods, and

(ii) §§ 2.8 and 2.9 describe the serial correlation between the deviations of sickness rates and suggests that this will be positive and confined to that between ages x and $x+1$, so far as the sickness periods shown in Table 3.3 are concerned. § A9 shows that, if there is correlation of this nature, the value of σ_r will no longer be an estimate of the variance of sickness rates. Table A2 shows how various levels of correlation will change the value of σ_r and enables an estimate of the serial correlation coefficient of lag 1, R_1, to be made from the value of σ_r.

Thus σ_r in Table 3.3 has been increased by duplicates and reduced by the positive correlation. It will be assumed as a starting point that the duplicates factors of § 2.6 will apply and the value of σ_r in Table 3.3, after reduction to remove the effect of duplicates, can then be used to enter Table A2 to get an estimate of R_1. The calculations are set out in Table 3.5, where column (5) gives the resulting estimate of R_1, made from the adjusted values of σ_r in column (4).

3.9. § A10 describes the effect of serial correlations R_1, R_2, . . . in the original data (but see § A11) on the correlations ρ_1, ρ_2, ρ_3 between the successive values of

Table 3.5. *Estimates of serial correlation lag 1 (R_1) in the original data*

Deferred period (weeks) (1)	Average value of σ_r (I) (Table 3.3) (2)	Duplicates factor (3)	σ_r Adjusted for duplicates (2)÷(3) (4)	Estimates of R_1 etc. By Table A2 (formula (A13)) (5)	By formulae (A14) R_1 (6)	R_2 (7)	R_3 (8)
1, S1/3 only	1·74	*1·5	1·16	−·23	−·07	+·39	−·18
		1·1	1·58	−·998	−·87	+·25	−·51
1, excluding S1/3	·84	*1·5	·56	+·46	+·57	+·32	+·16
		1·1	·76	+·28	+·32	+·12	+·13
		1·0	·84	+·20	+·20	+·03	+·11
4, 13 and 26	·93	*1·1	·85	+·18	+·16	−·05	+·02
		1·0	·93	+·09	+·02	−·18	−·02
4, 13, 26 and 1 (excluding S1/3)	·90	1·1	·82	+·22	+·22	+·02	+·06

* The duplicates factors of § 2.6.

Table 3.6. *Serial correlations of* r_x.

Deferred period (weeks)	Sickness period (weeks)	ρ_1	ρ_2	ρ_3
1	1/3	−·90	+·65	−·41
	4/9	−·80	+·45	−·24
	13/13	−·64	+·11	+·10
	26/26	−·59	−·12	+·44
	52/52	−·62	−·02	+·25
4	4/9	−·52	−·26	+·52
	13/13	−·65	+·03	+·23
	26/26	−·77	+·46	−·42
	52/52	−·72	+·21	+·13
13	13/13	−·76	+·33	−·16
	26/26	−·63	+·06	+·01
	52/52	−·54	−·13	+·38
26	26/26	−·72	+·20	+·17
	52/52	−·78	+·47	−·32
Average coefficients:				
(i) D1 (excluding S1/3)		−·68	+·12	+·14
(ii) D4, 13 and 26		−·68	+·14	+·08

r_x which will no longer take the theoretical values of formula (A10). Formulae (A14) give the relationship between the R's and the ρ's.

Table 3.6 sets out the serial correlation coefficients calculated from the values of r_x, from which the σ_r's were determined. These coefficients show considerable variation but the consistently high values for D1, S1/3 should be noted.

By substituting in formulae (A14) the values of σ_r (adjusted for duplicates) and the corresponding average values of ρ_1, ρ_2, ρ_3 from Table 3.6, estimates of R_1, R_2, R_3 can be made which take account of both σ_r and the ρ's. These estimates are given in columns (6), (7) and (8) of Table 3.5. The calculations in Table 3.5 have been made not only for the appropriate duplicates factors of § 2.6, but for other factors as well. This is because the figures in Table 3.5 and others to be discussed later (e.g. Section 6) indicated that the effect of duplicates might be less than that expected.

Discussion of Table 3.5

3.10. The expected form of the correlation in the original series is of a small positive correlation R_1, and zero for R_2 and R_3, at least so far as the sickness periods in Table 3.3 are concerned. The estimate of R_1 in column (5) is based on σ_r alone and assumes that R_2 and R_3 are zero. Looking at the lines (marked *) of the table where the duplicates factors of § 2.6 are used, it will be seen that for D4, 13 and 26 the estimate of R_1 in column (5) of +·18, agrees well with that in column (6) of +·16 which takes account of both σ_r and the ρ's. Also the corresponding estimates of R_2 and R_3 in columns (7) and (8) are small. Thus for these deferred

periods the duplicates factor of 1·1 gives values for the R's which reproduce the expected pattern; a lower duplicates factor of 1·0 (i.e., no effect) does not reproduce this pattern so well. Duplicate factor 1·1 seems suitable for D4, 13 and 26.

The results for D1, S1/3 only, for duplicates factor of 1·5 do not fit in with the expected pattern and the two estimates of R_1 differ appreciably. In view of the short period of sickness, only three weeks, there would probably not be much carry over of claims from one age to the next and a low value of R_1 might be reasonable, but higher correlations R_2 and R_3 do not seem appropriate. Reduction of the duplicates factor to 1·1 gives an absurd value of R_1 in column (5) and a high negative value in column (6). In view of the suspect nature of the data for D1, S1/3 (§§ 4.3–4.5) the results for this sickness period will not be considered further here.

For D1, excluding S1/3 the expected duplicates factor of 1·5 gives large values for R_1, the two estimates not agreeing well, and also quite large values for R_2 and R_3. If the duplicates factor is reduced to 1·1 the correlation pattern is fairly satisfactory, but the best result is for factor 1·0. As these sickness periods are thought to contain a much higher proportion of duplicates than D4, 13 and 26 it would be surprising to find that duplicates had *less* effect than in D4, 13 and 26. If it be thought that a duplicates factor of 1·1 could apply to D1, excluding S1/3, then it might be legitimate to combine these sickness periods with D4, 13 and 26. The effect of doing this is shown in the last line of Table 3.5 which agrees very satisfactorily with the expected pattern.

3.11. On the basis of Table 3.5 and the above discussion an estimated value for R_1 of $+\cdot 2$ seems appropriate for all sickness periods in Table 3·3 except S1/3. This estimate will be used in the remainder of this paper.

3.12. It was suggested in § 2.8 that the correlation between consecutive ages in the original data might increase with length of sickness period; this would be shown by a falling trend in the values of σ_r from left to right in Table 3.3 but the table shows no clear evidence of such a trend. It is perhaps worth mentioning (§ A13) that by appropriate substitutions in formula (A16) the 95% confidence limits for σ_r for deferred periods 4, 13 and 26 weeks combined are 1·26 to ·60, and for deferred period 1 week (excluding S1/3) they are 1·14 to ·54. All the values of σ_r (I) in Table 3.3, except that for S1/3, lie within the respective limits. Thus, so far as it goes, this tends to confirm the absence of trend but then a small trend could well be masked by the variations of σ_r and the ρ's.

Looking now at Table 3.4 it will be seen that for S104/all, each value of σ_r (II) is lower than any in Table 3.3, which tends to indicate a higher degree of correlation. Also as mentioned in § 2.8 sickness claims in this period can exceed one year in length and so may be allocated partly to each of 3 or 4 consecutive ages, thus introducing serial correlation at lag 2 or lags 2 and 3 into the original data. More than four consecutive ages cannot be involved because the experience covers only the four years 1972–75.

Manchester Unity 1893–97 (Whole Society)

3.13. The figures given by Coward (1949) for the variance of Manchester Unity sickness, enable calculations similar to those described above, to be made for three of the sickness periods with comparatively little labour. The results are given in Table 3.7.

Table 3.7. *Manchester Unity sickness rates (using data for ages 20–70, i.e., 48 values of* r_x*)*

Sickness period	σ_r (method I)
First 3 months	·85
Second 3 months	·68
Second 6 months	1·15
Average	·89

Average serial correlations of r_x

ρ_1	ρ_2	ρ_3
−·70	+·18	+·12

3.14. The Manchester Unity experience does not contain duplicates. The average of the three values of σ_r is ·89 and for this average Table A2 gives an estimate for R_1 of +·14. Substitution in formulae (A14) of the values of σ_r and the ρ's of Table 3.7 gives estimates for the R's of $R_1 = +·16$, $R_2 = +·09$ and $R_3 = +·16$. By formula (A16) the 95% confidence limits for σ_r are 1·15 to·63 and none of the three values in Table 3.7 fall outside these limits. All the figures for the Manchester Unity experience are very similar to those derived above for the 1972–75 sickness experience, and thus tend to confirm the type of pattern deduced in respect of that experience. The fact that both experiences lead to an estimate in the region of +·2 for R_1 is perhaps a little surprising.

The meaning of R_1

3.15. In this section of the paper it has been assumed that the estimates of R_1 relate to the serial correlation between the deviations in the original data. However § A11 shows that they in fact relate to a very complicated function of these deviations. The matter is discussed further in §§ 6.14–6.16 where it is suggested that it may not be too unreasonable to assume that R_1 does not differ greatly from the correlation between the deviations so far as the 1972–75 experience is concerned.

4. TESTING THE GRADUATION OF THE 1972 75 SICKNESS RATES

4.1. Section 3 has indicated that the variance of the deviations (defined in § 2.7) of sickness, appears to correspond with the values arrived at by method I of

Appendix B, when allowance is made for the effect of duplicates and the serial correlation between consecutive deviations. The data for deferred period 1, S1/3 has been found suspect (§§ 4.3–4.5).

4.2. Although the various statistical tests involving the assumption of normal distribution of the deviations were not applied to the graduations of the 1972–75 sickness rates, the results were carefully studied and a comparison made of the net premiums calculated by the graduated and ungraduated sickness rates (*C.M.I.R.* **4**, 1979, Table Seg 3.3.5). Also a statistical test of the runs of signs of the deviations was applied (§§ 3.3.3 and 3.3.4, Table Seg 3.3.4). This showed that of the sickness periods for which σ_r is given in Table 3.3 those for D4, S52/52, D13, S13/13 and 52/52 gave values of χ_1^2 slightly above the 5% level, in all cases because the number of runs were a little too few. Positive serial correlation between successive ages might be expected to result in fewer runs than when the ages are independent. However the result for D1, S1/3 was quite different and is discussed in the next paragraph. Sickness periods for which the runs of signs test gave a result significant at the 5% level are marked with an asterisk in Table 4.1. In general the results of the tests which were made, showed the graduations of the 1972–75 sickness rates to be reasonably satisfactory.

Deferred period 1 week, Sickness 1/3

4.3. The value of χ_1^2 for the runs of signs test for D1, S1/3 was 15·1, which corresponds with a chance of about ·0001 that the observed number of runs or more would occur in random sampling. Unfortunately the test used produces large values of χ_1^2 for both too few and too many runs. § 3.3.4 of *C.M.I.R.* **4** (1979), states that in this particular case the high value of χ_1^2 is due to too many runs, and that it is the only deferred and sickness period for which this feature was found.

For D1, S1/3 the graduation gave deviations between actual (A) and expected (E) weeks of sickness such that $A-E$ was positive for 18 ages and negative for 17 ages. Thus the maximum possible number of runs of signs is 35, when positive and negative signs occur alternately. Actually the graduated rates gave 30 runs. If the observed series of + and − signs is regarded as an ordered sequence, it needs the position of only three of the negative signs to be changed in order to convert the series to one having the maximum number of runs. In *C.M.I.R.* **4** (1979), § 3.3.4 the only comment made by the PHI Sub-Committee was that too many runs of signs "would often be taken to indicate overgraduation", and this in spite of the use of a mathematical formula with only 5 constants. The Sub-Committee apparently did not realize or did not feel it worthy of comment, that this feature must result from some peculiarity of the original data. However the pattern of signs seemed to me very extraordinary, and made me wonder whether it could perhaps be a real life example, similar to the peculiar mortality experience hypothecated by Redington in the discussion of Seal (1941), where at even ages all the lives exposed to risk were mine workers, and at odd ages all were insurance clerks. On graphing the ungraduated rates (Figure 1), it became quite

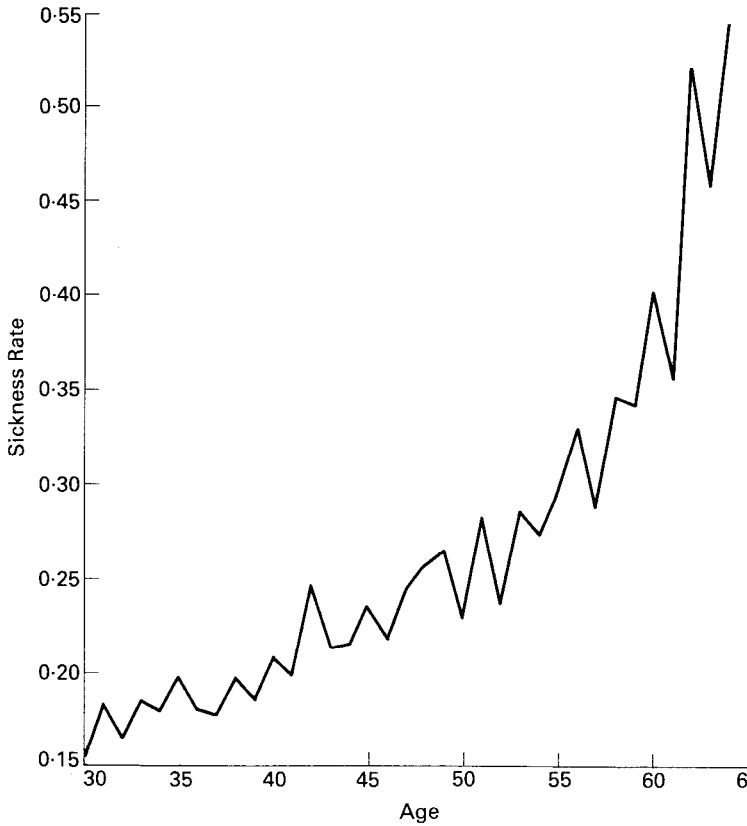


Figure 1. Ungraduated sickness rates for 1972–75 experience D1, S1/3

obvious where the graduated curve *had* to run; it just had to go between each successive pair of points with few exceptions, and this was where the graduation did, in fact, run. The extraordinary zig-zag pattern of variation in the data could be the reason for:

(i) the high negative correlation (lag 1) between the successive values of r_x of $-\cdot 90$ shown in Table 3.6, and,
(ii) the high values of σ_r (methods I and II) in Table 3.3.

4.4. In an attempt to explain matters the corresponding data of the Aggregate experience of individual policies in 1975–78 for D1, S1/3 (*C.M.I.R.* **7**, 1984) was examined. It was found that the first differences of the ungraduated sickness rates contained 21 runs of signs in 1975–78 as compared with 29 in 1972–75. Also the third differences of the ungraduated rates were much smaller and less variable

than those for 1972–75. The total of these third differences (disregarding signs) over the age range 30–64 was 2·16 for 1975–78, as compared with 3·72 for 1972–75. Thus, in spite of the two periods having the year 1975 in common, the later period does not give evidence of the peculiar zig-zag pattern found in the 1972–75 experience.

4.5. I have been quite unable to think of anything which might account for the regular zig-zag pattern of the 1972–75 data. I cannot see how so regular a pattern could be produced by duplicates, by differences between offices, by selective lapsing, or by heterogeneity in the data. In theory it could be a freak random result (chance ·0001). The only suggestion I can make is that some extraordinary error has crept into the scheduling or processing of the data. Until some real explanation is found I consider that the data of D1, S1/3 must be regarded as defective.

C.M.I.R. **7,** (1984), mentions in § 1.2 that "suspicion arose over the reliability of some data included in the deferred 1 and 4 weeks tables for 1972 and 1973, and it was felt unsafe to continue to rely on that data". No details are given. Does this refer to D1, S1/3?

Application of statistical tests

4.6. The next step is to test the graduations given in *C.M.I.R.* **4,** (1979) using the statistical tests normally applied to mortality rates, but with appropriate modifications, and to see how the tests perform. This requires the calculation for each individual age of the difference between the actual (A) and expected (E) weeks of sickness, and to divide it by the standard deviation of the weeks of sickness; that is to calculate

$$\frac{A-E}{\sqrt{E_x V_x}} = \delta_x \text{ (say)} \tag{4.1}$$

where V_x is the variance of sickness by method I of Appendix B which has already been calculated. It should be noted that while δ_x is the standardized deviate of $A-E$, it is also the standardized deviate of $z_x - z_x'$ or

$$\frac{A-E}{E_x}.$$

χ^2 test

4.7. The first test which will be applied is the χ^2 test. χ_f^2 is calculated for each period of sickness by the formula

$$\chi_f^2 = \sum_x \delta_x^2 \tag{4.2}$$

where f is the number of degrees of freedom and equals the number of values of δ_x^2 less the number of constants in the graduation formula, i.e., 5 for the 1972–75

Table 4.1. *Application of χ^2 test to graduated 1972–75 sickness rates*

Deferred period (1)	Sickness period (2)	f (3)	χ_f^2 (4)	Adjusted χ_f^2 (5)	Significance of adjusted χ_f^2 (6)	Runs of signs test —5% significance indicated by* (see § 4.2) (7)
1	1/3	30	47·8	22·1	—	*
	4/9	30	32·3	15·0	s	
	13/13	30	29·8	13·8	s	
	26/26	30	34·5	16·0	s	
	52/52	25	37·2	17·3	—	
4	4/9	30	46·8	40·3	—	
	13/13	30	32·0	27·5	—	
	26/26	30	36·9	31·8	—	
	52/52	30	39·8	34·3	—	*
13	13/13	30	47·0	40·5	—	*
	26/26	30	51·6	44·4	S	
	52/52	30	48·8	42·0	—	*
26	26/26	30	39·7	34·2	—	
	52/52	20	17·7	15·5	—	

Note: Upper and lower 5% limits of:
Adj. χ^2_{30}, 43·8 and 18·5,
Adj. χ^2_{25}, 37·7 and 14·6,
Adj. χ^2_{20}, 31·4 and 10·9.
S = greater than upper limit; s = less than lower limit.

graduations (see Benjamin and Pollard, 1980, §§ 11.8–11.14 and § 14.56). The resulting values of χ_f^2 are given in Table 4.1, column (4).

The values of χ_f^2 in column (4) of Table 4.1. need two adjustments:

(i) *For duplicates*

In spite of the curious results obtained in Section 3 for deferred period 1 week and in § 4.3 for S1/3, the duplicate factors of § 2.6 will be used for all deferred periods. Thus the adjustment to χ_f^2 is, for deferred period 1 week, division by $(1.5)^2$ and for the other deferred periods, division by $(1{\cdot}1)^2$.

(ii) *For serial correlation*

Rhodes (1927) p. 138 gives the average value of the variance of a sample of n drawn from a normal population, with mean of zero and variance of unity, where the consecutive items in the sample are correlated with a serial correlation coefficient (lag 1) of R_1. The formula for the average variance of a sample is

$$\left(1-\frac{1}{n}\right)\left(1-\frac{2R_1}{n}\right) = F\,(\text{say}) \tag{4.3}$$

For positive values of R_1 formula (4.3) gives a variance of the sample which is less than that of the population sampled (i.e., unity) and thus will in effect be based on a variance less than unity. A reasonable adjustment would seem to be to divide χ_f^2 by F, where F is calculated by formula (4.3) putting $n = f + 5$ and $R_1 = +\cdot 2$, the estimate of § 3.11. The values of F for the various values of n needed are:

n	F
25	·945
30	·954
35	·960

Adjustment (i) decreases χ_f^2 and adjustment (ii) increases it; the net effect is given in column (5) of Table 4.1.

Results of the χ^2 test

4.8. (i) For deferred periods 4, 13 and 26 weeks, column (6) of Table 4.1 shows that only for D13, S26/26 does Adjusted χ_f^2 exceed the upper 5% limit and then only slightly. Thus on the basis of the findings in § 3.10 it seems reasonable to say that the χ^2 test shows the graduations of these deferred periods to be satisfactory.

(ii) As regards D1 (excluding S1/3), all but one of the adjusted values of χ_f^2 are below the lower 5% limit and the remaining one is but little above. This would usually be regarded as showing that the graduated curve is too close to the ungraduated. However, in view of the findings in § 3.10, it would seem more sensible to say that the duplicates have had nothing like the expected effect in increasing the variance of the deviations of these sickness periods. Table 4.2 shows the values of Adjusted χ_f^2 when the duplicates factor is taken (i) as $(1\cdot5)^2$, as in Table 4.1, (ii) as $(1\cdot1)^2$ and (iii) as 1·0, i.e., no effect.

On the basis of the χ_f^2 test alone the significance columns of Table 4.2 indicate a slight preference for duplicates factor of $(1\cdot1)^2$ over 1·0. This is in line with the

Table 4.2. *Application of χ_f^2 test to Deferred Period 1 week (excluding S1/3) on three assumptions regarding the effect of duplicates*

Sickness period	f	Values and significance of χ_f^2 for duplicates factors of: $(1\cdot5)^2$		$(1\cdot1)^2$		1·0	
(1)	(2)	(3)		(4)		(5)	
4/9	30	15·0	s	27·8	—	33·6	—
13/13	30	13·8	s	25·7	—	31·0	—
26/26	30	16·0	s	29·7	—	35·9	—
52/52	25	17·3	—	32·2	—	39·0	S

See footnote to Table 4.1.

finding in § 3.10, that there is reasonable agreement between the two estimates of R_1 for duplicates factors of $(1{\cdot}1)^2$ and of 1·0.

(iii) In the case of D1, S1/3 the value of Adjusted χ_f^2 in Table 4.1 is above the lower 5% limit but not greatly so. In view of the conclusion in § 4.5 that the data are defective, there is no point in considering further the possible effect of duplicates or serial correlation in the original data.

4.9. For the χ_f^2 test to be valid the frequency distribution of the deviates δ_x should be normal. This question is considered in §§ 6.9–6.11.

Other statistical tests

4.10. Benjamin and Pollard (1980) describe in §§ 11.15–11.23 certain other statistical tests, involving the assumption of normality, which can be applied to mortality graduations, but which do not take account of the effect of duplicates, although appropriate adjustments can easily be made if the necessary information is available. Apart from stating in § 11.46 that duplicates increase the variance, the official text book does not consider at all how this should be taken into account; I regard this as a considerable defect.

These other tests have been applied to the 1972–75 sickness graduations and in nearly every case gave non-significant results, even when applied without adjustment to data containing duplicates. It seemed that, applied to the short span of ages available in the 1972–75 sickness data, these tests were too insensitive to be of much use. No further consideration of them will be given in this paper.

Manchester Unity experience

4.11. The Manchester Unity (Whole Society) sickness data do not contain duplicates and were graduated by a summation formula. The χ^2 test has been applied to the graduated sickness rates for ages 20–70 of each of the three sickness periods included in Table 3.7. The resulting values of

$$\chi_f^2 = \sum_x \delta_x^2$$

need adjustment only for the serial correlation, R_1, in the original data (§ 4.7(ii)). The appropriate value of F is obtained by putting $n = 51$ and $R_1 = +{\cdot}2$ (see § 3.14) in formula (4.3); this gives $F = {\cdot}973$, and each value of χ_f^2 was divided by this factor to get Adjusted χ_f^2 . The value of f was obtained by the method of Benjamin and Pollard (1980) § 13.92 as 37, i.e., a deduction of 14 from the number of values of δ_x. The figures are set out in Table 4.3.

Table 4.3 shows that the only value of Adj. χ_f^2 which lies outside the 5% limits is that for Second 3 months sickness; this would appear to indicate that the graduated rates follow the data a little too closely, a feature which could arise from the use of a summation formula. However the value of σ_r (Table 3.7) for Second 3 months sickness is quite close to the lower 95% confidence limit which denotes rather small variations in the original data, and the low value of Adj. χ_f^2

Table 4.3. *Application of χ^2_f test to graduation of Manchester Unity sickness rates*

Sickness period	f	χ^2_f	Adjusted χ^2_f	Significance of Adj. χ^2_f
(1)	(2)	(3)	(4)	(5)
First 3 months	37	34·0	34·9	—
Second 3 months	37	22·4	23·0	s
Second 6 months	37	47·7	49·0	—

Note: Upper and lower 5% limits of Adj. χ^2_f: 52·2 and 24.1.
S = above upper limit, s = below lower limit.

for this sickness period may simply follow from this. The Manchester Unity graduation is therefore found to be reasonably satisfactory as judged by the χ^2 test.

5. PRACTICAL GRADUATION TESTS

5.1. The χ^2 graduation tests used in Section 4, and the other tests mentioned in that section, made use of values of the variance of sickness calculated very laboriously by Coward's (1949) method (Appendix B, method I). In normal circumstances the labour involved would be quite impractical and some simpler way would have to be found. Following Coward's (1949) Table 1 in respect of the Manchester Unity experience, Table 5.1 has now been calculated showing, for quinquennial ages, the ratio of the variance of sickness, V_x, to the graduated 1972–75 sickness rates, z'_x. This ratio is denoted by k_x, i.e., $k_x = V_x/z'_x$.

5.2. It will be seen from Table 5.1 that the value of k_x varies comparatively little with age or deferred period, but substantially with length of sickness period. Some of the irregularities in the columns of the table are accounted for by the small size of some of the graduated sickness rates, where a change of ·001 in the rate has a substantial effect on the ratio in the table. A study of the table indicates that it can be represented reasonably closely by a set of integers which vary only with sickness period.

Table 5.2 sets out these integers and compares them with the corresponding integers used in *C.M.I.R.* 7, (1984), § 6.5 for testing the graduations of the 1975–78 experience. (The method and the two sets of integers were arrived at quite independently by the Sub-Committee and myself).

5.3. All the values of δ_x (formula (4.1)) were recalculated using the approximate expression

$$\frac{A-E}{\sqrt{kE}}$$

(Note that both the E's in this expression represent the expected weeks of

Table 5.1. *Values of* $k_x = V_x/z'_x$

Sickness period	Deferred period	Age 30	35	40	45	50	55	60	64	Average
1/3	1	1·6	1·7	1·9	2·0	2·0	2·1	2·1	2·0	1·9
4/9	1	7·1	7·1	7·2	7·3	7·3	7·3	7·2	7·1	7·2
	4	7·1	7·3	7·5	7·5	7·6	7·6	7·7	7·7	7·5
13/13	1	10·6	10·8	10·8	10·8	10·8	10·8	10·8	10·7	10·8
	4	10·9	10·9	10·9	10·9	11·0	11·0	11·1	11·0	11·0
	13	11·0	10·4	10·6	11·0	11·3	11·6	11·4	10·9	11·0
26/26	1	19·5	19·6	19·7	19·7	19·8	19·8	19·8	19·7	19·7
	4	19·7	19·7	19·7	19·8	19·9	20·0	20·1	20·1	19·9
	13	20·0	20·0	20·0	20·2	20·3	20·5	20·6	20·7	20·3
	26	21·6	19·5	18·9	20·0	20·9	21·4	21·3	21·0	20·6
52/52	1		31·6	31·8	32·0	32·2	32·3	32·2	32·0	32·0
	4	31·9	31·9	32·0	32·2	32·3	32·4	32·5	32·5	32·2
	13	32·1	32·3	32·1	32·1	32·3	32·8	33·4	33·7	32·6
	26			33·7	33·5	33·3	33·4	33·7	33·9	33·6

Table 5.2. *Values of* k *for all ages and deferred periods*

Sickness period	1/3	4/9	13/13	26/26	52/52
(i) 1972–75 experience (as used in this paper)	2	7	11	20	32
(ii) 1975–78 experience (as used in *C.M.I.R.* 7)	2	6	10	19	30

Table 5.3. *Changes in Adj.* χ^2_f *of Table 4.1 from use of the* k *approximation*

Deferred period (1)	Sickness period (2)	Change in Adj. χ^2_f (3)	Significance (4)
1	1/3	−0·1	—
	4/9	+0·5	s
	13/13	−0·3	s
	26/26	—	s
	52/52	+0·1	—
4	4/9	+3·1	—
	13/13	—	—
	26/26	−0·2	—
	52/52	+0·4	—
13	13/13	+0·4	—
	26/26	+1·4	S
	52/52	+1·0	—
26	26/26	+1·5	—
	52/52	+1·0	—

Note: S = above upper 5% limit of χ^2_f.
s = below lower 5% limit of χ^2_f.

sickness). The approximate values of δ_x on the basis of k show only small differences from the more accurate values and have an almost negligible effect on the χ_f^2test of §4.7. The changes in the values of Adj. χ_f^2 in Table 4.1 are shown in Table 5.3 and there is no change in the significance in column (6) of Table 4.1.

Thus the use of k, provided it can be determined with reasonable accuracy, would seem to be a practical method of testing the graduation of sickness rates by the usual tests. Satisfactory estimates of k could probably be obtained by calculating V_x for one or two ages of each sickness period, using method I.

6. DISCUSSION

6.1. Now that the results of the various tests have been set out and certain tentative conclusions drawn, it is necessary to draw the various threads together and try to arrive at a comprehensive statement of the overall position, so far as it is possible to do this. In particular three matters need to be discussed; these are the effect of duplicates, the extent to which the χ^2 test is valid for testing graduations of sickness rates and the meaning of R_1 in relation to the original sickness data.

Further discussion of the effect of duplicates

6.2. In mortality experiences the effect of duplicate policies has usually shown up clearly in the values of σ_r (Daw, 1945, 1974). Their effect on the 1972–75 sickness experience is not so clear, and in the case of sickness rates, there is the further complication of serial correlation between successive ages. Broadly the effect of duplicates seems to be that a duplicates factor of 1·1 applies to deferred periods 4, 13 and 26. The same factor probably also applies to deferred period 1 week, excluding S1/3 (§3.10). Because of the defective nature of the data no similar statement can be made for D1, S1/3 (§4.5). The results in §4.8 of the application of the χ_f^2 test to the graduations in *C.M.I.R.* **4** (1979) appear to substantiate these statements.

6.3. As mentioned in §2.2 the exposed to risk of the 1972–75 sickness experience is considerably smaller than that for the mortality tables examined by the r_x test. For the 1972–75 sickness experience the average exposed to risk per individual age for deferred periods 1, 4, 13 and 26 weeks is about 2,100, 2,800, 3,100 and 4,000 respectively, with the maximum at any age in the region of $1\frac{1}{2}$ times the average. These figures do not vary greatly for the individual sickness periods making up each deferred period. Thus the smaller numbers involved would result in more irregularity in the distribution of no-claim duplicates, and duplicate claims among the exposed to risk and sickness at individual ages.

6.4. Table 6.1 gives a *very* rough estimate of the number of duplicate policies and duplicate claims for each sickness period corresponding with the average numbers exposed to risk of §6.3. These estimates were made by assuming 1·6 policies per life for deferred period 1 week and 1·1 policies per life for the other

Table 6.1. *Rough estimates of the number of duplicate policies and duplicate claims per age*

Deferred period (weeks) (1)	Average exposed to risk per age (2)	Average number of duplicates in the exposed to risk per age (3)	Average number of duplicate claims per age in claims for sickness period: 1/3 (4)	4/9 (5)	13/13 (6)	26/26 (7)
1	2,100	790	107	24	6	2
4	2,800	250		7	2	0·6
13	3,100	280			2	0·6
26	4,000	360				0·8
Inception rates for age 50			·13527	·02983	·00702	·00225

three deferred periods. The numbers of duplicate claims were obtained by use of the graduated inception rates for age 50, making the crude assumption that the inception rates for, say, deferred period 13 weeks applied to sickness period 13/13 for all deferred periods.

Table 6.1 shows that only two out of the 10 cells have an average of more than 10 duplicate claims per age, and for three cells the average is less than one per age. Even allowing for the crudeness of these estimates, it is perhaps not surprising that the two longest sickness periods of deferred period 1 week do not show a greater effect of duplicates than for deferred periods 4, 13 and 26 weeks. It must also be remembered that the proportion of duplicates will not be the same for each age but will show considerable irregularities although, broadly, an increase with age would be expected.

6.5. The next step is to examine the frequency distribution of δ_x (formula (4.1)). Table 6.2 gives this distribution of δ_x for, (a) the 1972–75 graduations for all the deferred and sickness periods included in Table 3.3 except D1, S1/3 and, (b) the Manchester Unity (Whole Society) sickness graduations. The corresponding figures for the normal distribution are also included in the table.

In the case of the 1972–75 sickness, the variance of the normal distribution in the table is 1·16. This is the duplicates factor $(1{\cdot}1)^2$ (which is assumed to apply to all the deferred and sickness periods for which values of δ_x are included in the table), reduced by the application of F of ·96 (§4.7(ii)). As the correlation R_1 occurs only within, and not between each sickness period it seemed appropriate to use the value of F for samples of 35 rather than to regard the actual figures in Table 6.2 as one sample of 440 items for this purpose. For the Manchester Unity experience, which does not contain duplicates, the normal distribution shown in Table 6.2(b) has a variance of ·97, i.e., the value of F given in §4.11.

An interesting feature of Table 6.2 is the excess of the number of negative over positive values of δ_x for the 1972–75 experience. This feature is set out in more

Table 6.2. *Frequency distributions of values of δ_x for (a) 1972–75 (excluding S1/3) and (b) Manchester Unity Sickness Rates*

	Values of δ_x														
	Below −3·0	−3·0 to −2·5	−2·5 to −2·0	−2·0 to −1·5	−1·5 to −1·0	−1·0 to −·5	−·5 to 0	0 to ·5	·5 to 1·0	1·0 to 1·5	1·5 to 2·0	2·0 to 2·5	2·5 to 3·0	Above 3·0	Total	Goodness of fit test
(a) *1972–75 Sickness (excluding S1/3)*																
(i) Actual number (A)	—	1	1	23	52	78	88	68	53	34	19	15	5	3	440	$\chi^2_{10} = 28{\cdot}18$
(ii) Expected by normal curve (E)	1·2	3·3	9·5	22·0	41·6	63·8	78·6	78·6	63·8	41·6	22·0	9·5	3·3	1·2	440·0	$·005 > P > ·001$
(iii) $A-E$	−1·2	−2·3	−8·5	+1·0	+10·4	+14·2	+9·4	−10·6	−10·8	−7·6	−3·0	+5·5	+1·7	+1·8	—	
(b) *Manchester Unity Sickness*																
(i) Actual number (A)	—	2	—	3	10	18	37	40	28	9	4	2	—	—	153	$\chi^2_9 = 14{\cdot}22$
(ii) Expected by normal curve (E)	·2	·7	2·4	6·5	13·9	23·1	29·7	29·7	23·1	13·9	6·5	2·4	·7	·2	153·0	$P = ·12$
(iii) $A-E$	− ·2	+1·3	−2·4	−3·5	−3·9	−5·1	+7·3	+10·3	+4·9	−4·9	−2·5	− ·4	− ·7	− ·2	—	

Note: (i) The normal curve figures are those having zero mean, with variance of 1·16 for 1972–75 sickness and ·97 for Manchester Unity.
(ii) The brackets over certain columns at each end of the table indicate the columns combined in calculating χ^2 shown on the right.

Table 6.3. *Comparison of numbers of positive and negative values of δ_x in Table 6.2 for the 1972–75 experience*

Numerical value of δ_x	Negative values		Positive values	
	Actual numbers	Expected by normal curve	Actual numbers	Expected by normal curve
(1)	(2)	(3)	(4)	(5)
Over 2·0	2	14·0	23	14·0
2·0 and less	241	206·0	174	206·0
Total	243	220·0	197	220·0

detail in Table 6.3, which shows that there is an excess of negative over positive values of δ_x of 2·0 or less and that the reverse applies for values over 2·0.

Unlike the 1972–75 experience, the Manchester Unity (Table 6.2) shows a deficiency of negative values (i.e., 70 negative and 83 positive).

6.6. The effect of duplicates in a small sickness experience needs to be considered in some detail. The introduction of a no-claim duplicate increases the exposed to risk at the particular age and thus reduces the sickness rate. On the contrary the introduction of a duplicate claim increases both the exposed to risk and the weeks of sickness, the net effect being usually an increase in the sickness rate. If therefore the number of no-claim duplicates is comparatively small, and the number of duplicate claims much smaller, then the age distribution of duplicates and their claims can be very irregular. In the light of Table 6.1 this would seem likely to be the case in the 1972–75 experience. Also those ages where no-claim duplicates predominate would tend to show negative values of δ_x and this could explain the large number of negative values of smaller sizes. Correspondingly the ages which include duplicate claims would tend to give increased sickness rates and hence positive deviations, some of which might be quite large.

If the exposed to risk had been very much larger than in the 1972–75 experience the numbers of no-claim duplicates and duplicate claims would also be much larger and the two effects just explained would be more likely to affect all, or most, ages and to produce the averaging effect on which the duplicates factors given in *C.M.I.R.* **7** (1984) depend so that the distribution of δ_x would not be distorted in the ways described.

6.7. It is also of interest to look at individual large values of δ_x. Table 6.4 gives the detailed distribution of values of δ_x numerically greater than 2·0. For the sake of completeness the figures for S1/3 are included in this table, but there is nothing exceptional about them.

The three negative deviations in Table 6.4 are unremarkable except that in total they are substantially fewer in number than would be expected by the normal distribution. This can be explained by the argument of § 6.6 regarding no-claim duplicates.

The positive deviations are however much more numerous and there is a

Table 6.4. *Distribution of large values of δ_x by age for 1972–75 graduated sickness rates*

Deferred period (weeks)	Sickness period	Values of δ_x									
		Negative values of δ_x		Positive values of δ_x							
		Age: 61	Number	Age: 42	46	47	53	62			Number
1	1/3	−2·08	1	2·48				2·72			2
	4/9										—
	13/13				3·01						1
	26/26					3·06	2·73				2
	52/52			2·67			2·07				2
		Age: 57		Age: 31	34	42	43	45			
4	4/9	−2·70	1				2·36				1
	13/13						2·25				1
	26/26				2·05	2·02	2·06				3
	52/52			2·23				3·78			2
				Age: 33	36	37	45	58	59	60	
13	13/13				2·15		2·35		2·57		3
	26/26				2·80	2·25		2·09			3
	52/52			2·52					2·23	2·30	3
		Age: 49		Age: 50	59						
26	26/26	−2·01	1	2·15	2·41						2
	52/52										
		Total	3								25

tendency for several large deviations to cluster at one age or at one age and the following age, e.g.,

D1. Ages 46 and 47 (1 value each), age 53 (2 values)
D4. Ages 43 (3 values) and/or 42 and 43 (1 value each)
D13. Ages 36 and 37 (2 values and 1 value respectively)
Age 58, 59, 60 (1 value each)

Each of these could be accounted for by the presence at the age(s) given of several duplicates, subject to long claims which run from one sickness period on into the next, and in some cases pass to the next age as well. I had hoped that some, at least, of these examples could be checked against the original data to see whether duplicate claims were involved as suggested. Unfortunately I am informed that this is not feasible at this stage. However the suggestion is possibly supported by the Manchester Unity experience, which does not contain duplicates. Table 6.2 (b) shows that this experience has only 2 positive and 2 negative values of δ_x which are numerically greater than 2·0, as compared with 3·3 of each expected by the normal curve, although the summation graduation could have something to do with this. There is no evidence of the distortions which it is suggested are due to duplicates.

6.8. Table 6.4 gives no evidence of a greater proportion of duplicates in deferred period 1 week (excluding S1/3) as compared with deferred period 4, 13 and 26 weeks. It is not evident whether this is

(i) because the proportion of duplicates for deferred period 1 week in the 1972–75 experience is much less than that found in 1975–78 in Appendix F of *C.M.I.R.* 7, (1984),
(ii) because some or all of the defects in the data of S1/3 also extend into the later sickness periods, or
(iii) because the duplicates, even if in greater proportion in D1 (excl. S1/3), are still not numerous enough to produce the expected average effect.

Validity of graduation tests based on the normal curve

6.9. The χ^2 test as applied to sickness rates in Section 4 and the other statistical tests mentioned in § 4.10 assume that the deviations in the original data are distributed as the normal curve. Although the numerator of r_x is assumed to be composed almost entirely of these deviations, it became clear in the discussion of Daw (1945) that, apart from giving the scale of the distribution (i.e., the standard deviation, σ_r), the frequency distribution of r_x gives no information about the shape of the distribution of the deviations in mortality data. The same applies to the distribution of r_x from sickness data.

In testing a graduation the deviations between the actual and expected sickness are regarded as the theoretical deviations and, if the graduation is 'correct', it is their distribution which is involved in the statistical tests.

6.10. Table 6.2 (a) gives the combined frequency distribution of the values of δ_x in respect of the sickness graduations of the 1972–75 experience and also shows the corresponding figures for the normal curve, taking account of the effect of duplicates and the serial correlation (§ 6.5). The table also gives the result of a χ^2 test, to test the goodness of fit of the actual numbers with those of the normal curve shown in the table. The value of P in the table (i.e., less than ·005) indicates that it is rather unlikely that the observed differences can have arisen from random effects. The discussion of §§ 6.5–6.7 leads to the conclusion that the no-claim duplicates and the duplicate claims are not numerous enough for their effects to average out over individual ages. The different effects of no-claim duplicates (increased numbers of small negative values of δ_x) and of duplicate claims (increased numbers of large positive values of δ_x) are shown in the comparison with the normal curve in line (iii) of Table 6.2 (a).

The mean and variance of the distribution of δ_x for the 1972–75 rates are respectively $-·002$ and 1·14. The mean is effectively zero, as it should be for satisfactory graduations, and the variance of 1·14 is quite close to that of 1·16 determined independently in § 6.5 without any reference to the distribution of δ_x.

6.11. In determining χ^2_f by formula (4.2) the individual values of δ_x are squared, so that the value of χ^2_f is not affected by the signs of δ_x (which is why a test of signs is also needed). Thus in considering the validity of the χ^2 test, it is the normality of

Table 6.5. *Frequency distribution of* $|\delta_x|$ *for 1972–75 (excluding S1/3) sickness graduations. (Figures derived from Table 6.2)*

	Value of $\|\delta_x\|$					
	Above 2·0	1·5 to 2·0	1·0 to 1·5	0·5 to 1·0	0 to 0·5	Total
(a) 1972–75 Sickness (excluding S1/3)						
(i) Actual number (A)	25	42	86	131	156	440
(ii) Expected by normal curve (E)	28·0	44·0	83·2	127·6	157·2	440·0
(iii) $A-E$	−3·0	−2·0	+2·8	+3·4	−1·2	—
(b) Manchester Unity Sickness						
(i) Actual number (A)	4	7	19	46	77	153
(ii) Expected by normal curve (E)	6·6	13·0	27·8	46·2	59·4	153·0
(iii) $A-E$	−2·6	−6·0	−8·8	−0·2	+17·6	—
Goodness of fit test	(a) 1972–75 $\chi^2_4 = \cdot 60$; $\cdot 975 > P > \cdot 95$					
	(b) M.U. $\chi^2_4 = 11 \cdot 80$; $P \doteq \cdot 02$					

the distribution of $|\delta_x|$ which is relevant. Table 6.5 (a) gives the distribution of $|\delta_x|$ obtained from Table 6.2 (a). A χ^2 test of the goodness of fit with the normal curve gives a somewhat high value for P (i.e., greater than ·95) which, if anything, indicates an unexpectedly good agreement with the normal curve. Perhaps the degrees of freedom of the χ^2 should be reduced from 4 to 3 because the variance used for the normal curve has been deduced, to a limited extent from the inceptions of the deferred periods involved. This would give a value of P of about ·9, i.e., within the upper and lower 5% limits. However there is no ground for thinking that the distribution of $|\delta_x|$ differs appreciably from normality. The χ^2 test would therefore seem valid for use in testing the 1972–75 sickness graduations, provided appropriate adjustments are made on account of duplicates and serial correlation. Other statistical tests involving the assumption of normality, but which are affected by the signs of the values of δ_x, require the normality of the distribution of δ_x, rather than $|\delta_x|$, and so may be less valid than the χ^2 test.

6.12. The distribution of δ_x for the Manchester Unity experience in Table 6.2 (b) does not differ significantly from the normal curve there shown, with zero mean and variance of ·97 ($P = \cdot 12$). However, in spite of this result, it is perhaps relevant to mention that the mean and variance of δ_x are respectively +·05 and ·69 (thus the variance is well below that of ·97 on which the normal figures are calculated). As regards the χ^2 test of the graduations, Table 6.5 (b)

gives the distribution of $|\delta_x|$ for the Manchester Unity. This brings out strongly the tendency towards small values of δ_x and the value of P of ·02 gives less assurance that the distribution may be regarded as normal. These features and the fact that the variance of δ_x (=·69) is less than ·97 could be explained as a tendency of a summation graduation to follow too closely the ungraduated rates.

6.13. As regards the assumption of normality, the remarks of Professor M. Greenwood in the discussion of Daw (1945) are worthy of note. He found that certain samples of 1,000 showed an excellent fit to the normal curve but for one large sample of 20,000 the fit 'was execrable'. He suggested that in practice statistics nearly always contained a certain amount of extraneous matter, and that in a small experience this heterogeneity was concealed by the roughness of the data. An increased number of cases removed the roughness but the heterogeneity remained. I certainly feel that the 1972–75 sickness experience contains quite a lot of extraneous matter. In view of the results just described I am uncertain whether Prof. Greenwood's remarks are comforting or the reverse.

How is R_1 *to be regarded?*

6.14. In § A11 it is shown that R_1 is the serial correlation coefficient between successive values of some very complicated function of the deviations of the sickness rates. It was suggested that R_1 might be somewhere in between the serial correlations (lag 1) of the values of ε_x and of $\varepsilon_x/SD\ (\varepsilon_x)$. In the context of testing graduations these two functions would be $(A-E)/E_x$ and δ_x respectively, provided the graduation had been found satisfactory, as is the case with those of the 1972–75 experiences, excluding S1/3.

The values of R_1 have been calculated by formulae (A14) for individual deferred and sickness periods using a duplicates factor of 1·1. These are set out in Table 6.6 and compared with the corresponding serial correlation coefficients (lag 1) between successive values of δ_x and of $(A-E)/E_x$.

The averages in the table show R_1 to be greater than both the other two coefficients. (It is purely chance that the average of R_1 of +·2 is exactly equal to the estimate made in § 3.11). However consideration of the body of the table

Table 6.6. *Values of* R_1 *by formulae (A14) (duplicates factor 1·1) compared with serial correlation coefficients (lag 1) between successive values of* δ_x *and of* $(A-E)/E_x$ *for the 1972–75 sickness experience. S1/3 is omitted*

Deferred period (weeks)	S4/9			S13/13			S26/26			S52/52		
	R_1	δ_x	$\frac{A-E}{E_x}$	R_1	δ_x	$\frac{A-E}{E_x}$	R_1	δ_x	$\frac{A-E}{E_x}$	R_1	δ_x	$\frac{A-E}{E_x}$
1	+·30	+·17	−·09	+·59	+·41	+·22	+·14	−·03	−·11	+·12	+·15	+·02
4	+·06	+·10	−·11	+·29	−·01	−·47	+·30	+·22	−·03	+·13	−·06	+·15
13				−·06	+·10	+·21	+·03	+·24	+·14	+·35	+·34	+·53
26							+·05	−·11	+·14	+·31	−·11	−·18
Average correlation coefficients for all deferred and sickness periods.										+·20	+·12	+·04

Table 6.7. *Manchester Unity experience. Values of* R_1 *by formulae (A14) compared with serial correlation coefficients (lag 1) between successive values of* δ_x *and of* $(A-E)/E_x$.

Sickness period	R_1	δ_x	$\frac{A-E}{E_x}$
First 3 months	+·22	−·16	−·14
Second 3 months	+·46	−·002	+·06
Second 6 months	−·30	−·33	−·21
Average	+·14	−·17	−·10

shows substantial variations in all three coefficients. Out of the 13 sets of three coefficients, 6 show a decreasing series in the order R_1, δ_x, $(A-E)/E_x$, and of the remaining 7 sets, R_1 is in the second place 5 times and in the third place twice. Only one of the 6 possible arrangements of the three coefficients is not represented. It could therefore be argued that, taking account of the general variability and the relative smallness of the data, R_1, or perhaps a somewhat smaller value, could be taken as an estimate of the serial correlation (lag 1) between the standardized deviations in the original series of sickness rates.

The more complicated correlation pattern of S104/all is discussed in § 7.4.

6.15. It is interesting to note that for the Manchester Unity experience the estimate of R_1 by the r_x test is also in the region of +·2. Table 6.7 corresponds with Table 6.6, but, in spite of the much greater range of ages available in each sickness period, shows a more confused pattern than that of the 1972–75 experience. The averages of the coefficients in Table 6.7 show R_1 to be positive as would be expected, but those for δ_x and $(A-E)/E_x$, which are calculated from the deviations between graduated and ungraduated sickness rates, are negative. Thus this table gives little reason to think that R_1 is related in some way to the serial correlation in δ_x or $(A-E)/E_x$. Possibly this has something to do with the use of a summation formula to graduate this experience. Perhaps the deviations δ_x and $(A-E)/E_x$ for a summation graduation differ considerably from the deviations which are taken into account in the r_x test (from which R_1 is calculated) and which are largely independent of the graduation.

6.16. The tentative suggestion is made that the estimate of R_1 obtained from the r_x test might be looked upon as being of the same sign and similar magnitude to, or perhaps a little greater than, the correlation between successive values of δ_x (the standardized deviation between graduated and ungraduated sickness rates) derived from a satisfactory graduation by a mathematical formula.

7. SICKNESS PERIOD 104/all OF 1972–75 SICKNESS

7.1. Apart from brief comments in §§ 2.11 and 3.12, sickness period 104/all has not yet been considered. This is principally because, for various reasons, it was

thought that any conclusions drawn from these might be less reliable than those already described.

The number of weeks of sickness was considered quite inadequate to justify calculating r_x for any ages below 40 and even for higher ages the number of policies on which claims were made must be very small and the number of duplicate claims even smaller. As explained in § B9 it seemed impracticable to use method I to calculate the variance of sickness, while the numerical values of the moments of sickness duration (*C.M.I.R.* **4** (1979), Table Se 1.2.3) needed in order to use method II did not start until age 40. The graduated sickness rates given in *C.M.I.R.* **4** (1979) also did not start until age 40 (age 42 for deferred period 4 weeks). Thus only 22 values, at the most, of r_x were available to calculate each value of σ_r and this small amount of data, combined with the greater degree of correlation between successive items of the data (§ 3.12), seemed likely to reduce the reliance which could be placed on any results. Also *C.M.I.R.* **7** (1984), §§ 4.3–4.6, explains the dubious nature of the exposed to risk used for S104/all.

The r_x *test*

7.2. This paragraph sets out the results of calculations, similar to those already described, in respect of S104/all. Table 3.4 gives the values of σ_r and Table 7.1 the serial correlations, ρ_1, ρ_2, ρ_3 between successive values of r_x for S104/all.

The values of σ_r for S104/all in Table 3.4 show much less variation than those in Table 3.3, and give no indication that duplicates have a greater effect in D1 than in the other deferred periods. Accordingly all the four periods will be considered together in estimating the values of R_1, R_2, R_3 by means of formulae (A14). Separate calculations will be made for duplicates factors of 1·1 and 1·0 (no effect). The number of duplicate claims in S104/all will tend to be reduced because they come from the smaller volume of sickness business effected in earlier years (see § 2.11). It may well be that a duplicates factor of 1·0 might not be unreasonable for S104/all.

Table 7.2 gives the results of the application of formulae (A14), substituting

Table 7.1. *Serial correlations of* r_x *for sickness rates of S104/all*

Deferred period (weeks)	ρ_1	ρ_2	ρ_3
1	−·54	+·23	−·58
4	−·58	−·23	+·66
13	−·69	+·03	+·43
26	−·72	+·17	+·22
Average coefficients	−·64	+·05	+·20

Table 7.2. *Estimates of* R_1, R_2, R_3 *for S104/all by formulae (A14)*

Duplicates factor	σ_r	Adjusted σ_r (2)÷(1)	Estimated serial corrrelations		
			R_1	R_2	R_3
(1)	(2)	(3)	(4)	(5)	(6)
1·1	·52	·473	+·65	+·36	+·17
1·0		·52	+·60	+·32	+·16

the average value of σ_r for all four deferred periods and the corresponding averages of ρ_1, ρ_2, ρ_3 of Table 7.1.

As suggested in § 3.12, the correlations between the successive deviations of the original data (but see §§ A11 and 6.14) are considerably greater than those found for earlier sickness periods and extend to the larger lags. (For this reason R_1 cannot be estimated from Table A2). Also the three values are all positive and decrease with increasing lag, thus fitting in with the reasoning which suggested the correlation in the first place. In what follows the estimates of the three coefficients will be taken as

$$R_1 = +\cdot 6,\ R_2 = +\cdot 3,\ R_3 = +\cdot 2 \tag{7.1}$$

Testing the graduations

7.3. The values of δ_x were calculated in respect of the graduated sickness rates for S104/all given in *C.M.I.R.* **4,** (1979) and the values of χ_f^2 calculated. The two adjustments described in § 4.7 were applied to the values of χ_f^2. For adjustment (ii) formula (4.3) must be replaced by

$$F = \left(1-\frac{1}{n}\right) - \frac{2}{n}\left\{R_1\left(1-\frac{1}{n}\right) + R_2\left(1-\frac{2}{n}\right) + R_3\left(1-\frac{3}{n}\right)\right\} \tag{7.2}$$

(Substituting $R_2 = R_3 = 0$ in formula (7.2) gives formula (4.3) as it should). Table 7.3 gives the results of applying these adjustments to χ_f^2.

Table 7.3. *Application of* χ^2 *test to sickness rates S104/all*

Deferred period (weeks)	n	f	F	χ_f^2	Adjusted χ_f^2 for duplicates factor of:	
					$(1\cdot1)^2$	1·0
(1)	(2)	(3)	(4)	(5)	(6)	(7)
1	25	20	·878	20·5	19·3	23·3
4	23	18	·868	18·3	17·5	20·1
13	25	20	·878	19·4	18·3	22·1
26	25	20	·878	19·7	18·5	22·4

Note: Upper and lower 5% limits of Adj. χ^2_{20}; 31·4 and 10·9.

The adjusted values of χ_f^2 all lie well within the upper and lower 5% limits, and therefore indicate a satisfactory graduation but give little indication of which duplicates factor is the more appropriate. In fact, in general, the figures for S104/all give the impression of smaller variation as compared with those already studied; perhaps this is due to the large correlations indicated by the R values. The test of runs of signs of $A-E$ given in *C.M.I.R.* **4** (1979) § 3.3.4 shows that for S104/all, deferred periods 13 and 26 weeks have slightly too few runs of signs at the 5% level, (perhaps due to the correlation) the other two deferred periods being satisfactory. The distribution of signs of δ_x is rather different from that given in Table 6.3. Out of the 98 values of δ_x, 54 have positive signs and 44 negative signs. Only 3 are numerically greater than 2·0, one being positive and two negative. Both these comparisons of positive and negative signs are in the opposite direction to those of Table 6.3, but are very similar to the corresponding figures for the Manchester Unity experience. This might perhaps indicate that there are few duplicates in the data for S104/all.

Meaning of R_1, R_2, R_3

7.4. Table 7.4 gives the serial correlation coefficients of lags 1, 2 and 3 calculated from the successive values of δ_x for S104/all.

The agreement between R_1 and the corresponding (lag 1) coefficient for the δ_x's is proportionately better than that considered in § 6.14 and the figure for δ_x is again the smaller of the two. The values of R_2 and R_3 do not agree well with those for δ_x but, in view of the large variability of these correlation coefficients, good agreement is perhaps too much to expect. The fact that δ_x gives a correlation lower than R_1 on two occasions might be thought to give a little support to the view that the correlation in the original data is less than R_1 calculated from the r_x test, which relates to the correlation between some complicated function of the deviations in the original data (see § A11). Maybe this is straining the interpretation of the figures too far.

Table 7.4. *Serial correlation coefficients of lags 1, 2 and 3 between values of* δ_x *for S104/all*

Deferred period (weeks) (1)	Number of values of δ_x (2)	Serial correlations lag 1 (3)	lag 2 (4)	lag 3 (5)
1	25	+·39	−·34	−·54
4	23	+·45	+·14	+·08
13	25	+·43	−·003	−·41
26	25	+·58	+·35	+·04
Average coefficient		+·47	+·04	−·23
Estimated values of R_1, R_2, R_3 (for comparison)		+·6	+·3	+·2

Table 7.5. *Values of* $k_x = V_x/z'_x$ *for S104/all*

Deferred period (weeks)	Age 40	45	50	55	60	64	Average
1	73	42	46	51	39	40	48
4	..	35	43	53	49	34	43
13	36	49	54	48	48	41	46
26	78	70	61	49	45	46	58

Practical graduation tests

7.5. The values of k_x, the ratio of the variance of sickness to the graduated sickness rate have been calculated and are shown in Table 7.5.

The values of k_x in Table 7.5 are much more variable than those of Table 5.1 for the earlier sickness periods and do not give nearly the same impression of constancy. This is probably because they are based on variances of sickness calculated from ungraduated moments of the duration of sickness (i.e. method II of Appendix B), and because of the small amount of data available, particularly at the younger ages. The large values of k_x of 70 and over all occur at the younger ages and involve small graduated sickness rates of ·1 or less. If these values are omitted the remainder average 45 and it is suggested that this value might be suitable for use in the 'Practical graduation tests' discussed in Section 5. After writing this, reference was made to *C.M.I.R.* **7**, (1984) § 6.5 and it was found that here also a figure of 45 was adopted.

7.6. The 'practical method' of Section 5 has been applied to the graduations of S104/all, using $k = 45$. In spite of the variability of k_x, the resulting values of χ^2_f for deferred period 1, 4 and 13 weeks differ from those in Table 7.3, column (5) by, at the most, ·4. But for deferred period 26 weeks the 'practical' χ^2_f is 24·3 compared with 19·7 in Table 7.3. While the small number of claims at the younger ages may be the cause, it gives a warning that care is needed in using the approximate method.

8. ASSESSMENT OF RESULTS

8.1. This study must be regarded as only a start. It is confined to the 1972–75 sickness experience of individual PHI policies, supplemented by the application of the same methods to the Manchester Unity experience. Any conclusions apply only to these experiences. Study of more sickness data by the same methods, or, perhaps better, by different methods, is needed to find whether the results are particular to the data studied or are of more general application. It would have been better if the more homogeneous 1975–78 Standard experience had been used, but this was not published until most of the extensive calculations on the 1972–75 figures had been completed.

8.2. The work was greatly helped by the results of the analysis of duplicates in

the 1975–78 experience set out in *C.M.I.R.* **7,** (1984), Appendix F. However it seems a pity that, having gone to the length of identifying first, second, third and subsequent policies, the last group was not broken down into . . . 3rd, 4th, 5th . . . etc. policies. The length of the tail of the distribution has an appreciable effect on the resulting increase in the variance of sickness (Daw, 1951). As yet there is no information whatever regarding the shape of the distribution of duplicates in sickness data. This could be quite different from the distribution in mortality data, about which there is a little information; the inflation of the seventies could have had an appreciable effect on the shape.

8.3. The r_x test has been applied in a largely pragmatic manner, involving taking averages of various sets of values, sometimes showing wide variations, and searching for patterns. Apart from when studying graduation tests, little use has been made of statistical significance tests, largely because there was so much internal correlation in the figures involved.

The methods used have, I think, produced a reasonably consistent overall picture of the sickness data studied, although there have been some confused patches. For example, (i) the defective data of D1 S1/3, (ii) the resulting uncertainty regarding the other sickness periods of D1, (iii) the apparent small effect of the greater number of duplicates in D1 (excl. S1/3) as compared with the other three deferred periods and (iv) the failure of σ_r by methods I and II to show up the effect of duplicates which still puzzles me.

The suggestion regarding the meaning of R_1 made in §§ 6.16 and 7.4 can only be resolved by applying the same methods to other sets of sickness data graduated by a mathematical formula—a long and laborious undertaking.

8.4. Some success can, I think, be claimed in showing up the effect of duplicates in sickness data, when the numbers are not large enough, for the different effects of no-claim duplicates and duplicate claims to average out in individual ages. These effects could be an inherent difficulty because sickness rates for sickness of longer durations become very low until the last residual period of 104/all is reached. Certain features of S104/all, discussed in § 7.3, tend to indicate that there may be few duplicates in the data of this sickness period.

Some success has also been achieved in revealing the nature of the correlation between sickness data at successive ages. The results seem to confirm the expected nature of this correlation and also to provide some numerical measure of it.

Surprisingly the estimates of R_1 made from the r_x test turned out to be of about $+\cdot 2$ for both the 1972–75 and the Manchester Unity experiences, although these are separated in time by some 80 years. One wonders whether other sets of sickness data allocated to ages in the same way would also give rise to estimates of about $+\cdot 2$. As expected S104/all showed much higher and more extensive correlation in the original data.

8.5. The r_x test has shown up the necessity of taking account of the effect of duplicates and the serial correlation in the original data when applying statistical tests to the graduation of sickness rates. Methods are proposed of adjusting for

these effects the values of χ_f^2 obtained when testing sickness graduations. It is argued that, if this is done, the χ_f^2 test, being based on the squares of the deviations, is a valid test of sickness graduations by a mathematical formula, but that there may be some doubts about similar tests which are affected by the signs of the deviations (§ 6.11). The need for a test of runs of signs to be combined with the χ_f^2 test is also brought out.

8.6. The paper is perhaps open to the criticism that reasons (e.g. duplicates) have been assigned in too definite a fashion to account for certain features of the various calculations. It is possible that some of these features may be due to the smallness of the data or to peculiarities in them. However the reasons proposed are often supported by other evidence in addition to that primarily being considered.

8.7. While this paper still leaves some loose ends, I like to think that it contributes something to our understanding of the nature of sickness experiences and perhaps also to the methods of studying them.

REFERENCES

ANDERSON, J. F. & DOW, J. B. (1948) *Actuarial Statistics, Vol. II. Construction of Mortality and Other Tables.* C.U.P., Cambridge.

ANDERSON, O. (1926) Ueber die Anwendung der Differenzenmethode ("Variate Difference Method") bei Reihenausgleichungen, Stabilitätsuntersuchungen, und Korrelationmessungen. *Biometrika* **18,** 293–320.

BEARD, R. E. (1947) The standard deviation of the distribution of sickness. *J.S.S.* **7,** 23.

BEARD, R. E. (1951) Some notes on graduation. *J.I.A.* **77,** 382.

BENJAMIN, B. & POLLARD, J. H. (1980) *The Analysis of Mortality and Other Actuarial Statistics.* Heinemann, London.

BIBBY, J. (1983) *Quotes, damned quotes and . . . .* Damart Books: Halifax.

COMMAND 6907. (1913) Memorandum on Rates of Sickness and Disablement. pp. 590–594. (See *Reprints from the course of reading for the examinations of the Institute of Actuaries*, 1946. p. 99–103. Institute of Actuaries Students' Society: London.)

CONTINUOUS MORTALITY INVESTIGATION REPORTS:

(1979) Sickness Experience 1972–75 for Individual Policies. *C.M.I.R.* **4,** 1 and discussion in *J.I.A.* (1979) **106,** 289.

(1984) Sickness Experience 1975–78 for Individual P.H.I. Policies. C.M.I.R. **7,** 1 and discussion in *J.I.A.* (1984) **111,** 503.

COWARD, L. E. (1949) The distribution of sickness. *J.I.A.* **75,** 12.

DAW, R. H. (1945) On the validity of statistical tests of the graduation of a mortality table. *J.I.A.* **72,** 174.

DAW, R. H. (1951) Duplicate policies in Mortality Data. *J.I.A.* **77,** 261.

DAW, R. H. (1974) A study of the variance of mortality rates. *J.I.A.* **101,** 415.

DAW, R. H. (1982) The improving accuracy of the basic data and the exposed to risk of the English Life Tables. *J.I.A.* **109,** 117 and 459.

DAW, R. H. (1984) The effect of duplicate policies upon the variance of sickness data. *J.I.A.* **111,** 567.

FISHER, R. A. (1941) *Statistical Methods for Research Workers.* Oliver & Boyd; London (Eighth Edition).

JAGER, J. DE. (1953) Stochastical investigations on mortality tables. *Het. Verzekerings—Archief, Actuarieel Bijvoegsel.* **30,** 76.

KENDALL, M. G. & STUART, A. (1966) *The Advanced Theory of Statistics.* Vol. III (First Edition). Charles Griffin: London.

QUENOUILLE, M. H. (1953) Modifications to the variate difference method. *Biometrika* **40,** 383.

REDINGTON, F. M. & MICHAELSON, R. L. (1940) An aspect of the *a priori* probability theory of mortality. *Trans. Twelfth Int. Cong. Act., Lucerne.* **I,** 225.

RHODES, E. C. (1927) The precision of means and standard deviations when the individual errors are correlated. *J. Roy. Stat. Soc.* **90,** 135.

RHODES, F. (1946) Percentage table for the estimation of sickness rates for special periods of sickness. *J.I.A.* **72,** 455.

SEAL, H. L. (1941) Tests of a mortality table graduation. *J.I.A.* **71,** 5.

WATSON, A. W. (1903) *An Account of an Investigation of the Sickness and Mortality Experience of the Independent Order of Oddfellows, Manchester Unity during the five years 1893–1897.* I.O.O.F, M.U: Manchester.

YULE, G. U. (1921) On the time-correlation problem, with special reference to the variate difference correlation method. *J. Roy. Stat. Soc.* **84,** 497.

APPENDIX A

THE r_x TEST

A1. The r_x test was put forward by Redington and Michaelson (1940) for application to mortality tables to study the variability of the data. The basic assumption is that the third differences of what may be called the underlying mortality rates are small; this will usually be found to be so for graduated rates of mortality. Thus if the initial data and the method of calculating q_x, the ungraduated rates, are accurate, then the third differences calculated from the ungraduated set of mortality rates should be made up almost entirely of the random errors. Further if each value of $\Delta^3 q_x$ is divided by its statistical standard deviation, the resulting values, denoted by r_x, should have a mean close to zero and a variance of unity.

A2. The variance of q_x according to the binomial distribution is $q''_x p''_x / E_x$ where q''_x is the true underlying rate of mortality. Thus the variance of $\Delta^3 q_x$ (i.e., of $q_{x+3} - 3q_{x+2} + 3q_{x+1} - q_x$) is

$$\frac{p''_{x+3}\, q''_{x+3}}{E_{x+3}} + 9\frac{p''_{x+2}\, q''_{x+2}}{E_{x+2}} + 9\frac{p''_{x+1}\, q''_{x+1}}{E_{x+1}} + \frac{p''_x\, q''_x}{E_x} = \sigma^2(\Delta^3 q_x) \qquad \text{(A1)}$$

and

$$r_x = \frac{\Delta^3 q_x}{\sigma(\Delta^3 q_x)} \qquad \text{(A2)}$$

In practice the underlying mortality rates are not known and q''_x is taken as being either the graduated rate q'_x, or the ungraduated rate q_x, the latter choice making the r_x test independent of any graduation; the choice has been found to have negligible effect on the result of the test.

If therefore the mortality rates show variations which follow the binomial distribution, then the standard deviation of the calculated values of r_x, denoted by σ_r, should be near to unity. If the mortality data contain duplicate policies on the same life then σ_r will become greater than unity, while defects in the original data or in the methods of calculating the exposed to risk or the ungraduated rates of mortality may also increase σ_r (Daw, 1974, 1982).

A3. In this paper the r_x test is applied to sickness claim inception rates and to sickness rates. In the case of inception rates the formulae given above for mortality rates are used, but q_x is replaced by the inception rate at age x.

For the ungraduated sickness rates of a particular deferred period and sickness period, denoted by z_x, the formula for r_x is

$$r_x = \frac{\Delta^3 z_x}{\sqrt{\sigma^2(\Delta^3 z_x)}} \qquad \text{(A3)}$$

where
$$\sigma^2(\Delta^3 z_x) = \frac{V_{x+3}}{E_{x+3}} + 9\frac{V_{x+2}}{E_{x+2}} + 9\frac{V_{x+1}}{E_{x+1}} + \frac{V_x}{E_x} \tag{A4}$$

V_x denoting the variance of sickness (in weeks) of one individual at age x and E_x the exposed to risk. Thus the variance of weeks of sickness among E_x lives is $E_x V_x$ and the variance of the sickness rate is $E_x V_x / E_x^2$ or V_x/E_x.

For sickness rates, unlike mortality rates, there is no simple formula for V_x and considerable calculation is required in order to determine the values. In this paper V_x has been calculated for the 1972–75 sickness experience by two methods which are described in Appendix B.

A4. In dealing with the 1972–75 sickness data it has again been assumed, as for mortality, that third differences of z_x are small. The graduated sickness rates are given to only three places of decimals and by third differences, as might be expected, the values have become somewhat irregular but they are nevertheless comparatively small.

Relation of r_x *test to variate difference method*

A5. In the discussion of Daw (1945) Seal explained that the r_x test was closely related to what is known as the variate difference method, which has been extensively studied mathematically over many years. These results have comparatively little relevance to the application of the r_x test to mortality data, but, when it is applied to the more complex situation of sickness rates, the work on the variate difference method is needed.

A6. The variate difference method is dealt with in Kendall and Stuart (1966) §46.24 onwards. It was originally devised to study time series and the initial concept is of an observed series U_t which consists of a polynomial trend $F(t)$ plus a random error element ε_t with zero mean and variance v which is the same for all t. By successive differencing the polynomial element $F(t)$ is eliminated and at that stage the differences are composed entirely of the random errors ε_t. Thus if the polynomial trend has been eliminated by third differences then

$$E(\Delta^3 \varepsilon_t) = 0 \tag{A5}$$
$$\Delta^3 \varepsilon_t = \varepsilon_{t+3} - \binom{3}{1}\varepsilon_{t+2} + \binom{3}{2}\varepsilon_{t+1} - \varepsilon_t$$
$$\text{Var}\,(\Delta^3 \varepsilon_t) = v \sum_{j=0}^{3} \binom{3}{j}^2 = 20v \tag{A6}$$

Rearranging (A.6) gives

$$\frac{\text{Var}\,(\Delta^3 \varepsilon_t)}{20\,v} = \frac{\text{Var}\,(\Delta^3 U_t)}{20\,v} = 1 \tag{A7}$$

If therefore v is known, a test of its accuracy or of the assumptions on which it has been determined would be to calculate

$$\frac{1}{20v(n-3)} \sum_{t=1}^{n-3} (\Delta^3 U_t)^2 \tag{A8}$$

(where n is the number of values of U_t available) and to see whether the resulting value was near to unity.

A7. The essential difference between the variate difference method and the r_x test is that in the latter the variance of each rate (of mortality or sickness) differs at each age, whereas the variance of each U_t is assumed to be the same.

Now (A8) can be written in the form

$$\frac{1}{n-3}\sum_{t=1}^{n-3}\frac{(U_{t+3}-3U_{t+2}+3U_{t+1}-U_t)^2}{\operatorname{Var}\varepsilon_{t+3}+9\operatorname{Var}\varepsilon_{t+2}+9\operatorname{Var}\varepsilon_{t+1}+\operatorname{Var}\varepsilon_t} \tag{A9}$$

where, for the variate difference method, $\operatorname{Var}\varepsilon_t = v$ for all values of t. In the r_x test the expression which is summed corresponds to r_x^2 and

$$\sigma_r^2 = \frac{1}{n-3}\sum_x r_x^2.$$

Thus the basic individual item in the variate difference method is $\Delta^3 U_t$ which has a mean of zero and variance of v. In the r_x test the individual item is r_x which has a mean of zero and a variance of unity, if all assumptions hold. Thus formulae calculated in respect of the variate difference method will apply to the r_x test if v is put equal to unity.

A8. If the errors ε_t are independent of each other (i.e., there is no serial correlation between successive values), the effect of differencing the series of errors is to introduce serial correlations, because successive terms of the difference series are no longer independent of each other. For third differences the first three serial correlations, say ρ_1, ρ_2, ρ_3, of lags 1, 2 and 3 respectively (i.e., between terms t and $t+1$, t and $t+2$, t and $t+3$) have theoretical values:

$$\rho_1 = -\cdot 75,\ \rho_2 = +\cdot 30,\ \rho_3 = -\cdot 05, \qquad \rho_4 = \rho_5 = \rho_6 = \ldots = 0 \tag{A10}$$

These values were given by Yule (1921) and Anderson (1926) in his formula (19) gave the general formula for the serial correlations for differences of order k and lag j ($k \geqslant j$) as

$$\rho_{j\cdot k} = (-1)^j \frac{k!\,k!}{(k-j)!\,(k+j)!} \tag{A11}$$

For $k=3, j=1, 2, 3$ formula (A11) gives the values of formula (A10).

These coefficients are the theoretical correlations between the successive values of $\Delta^3 U_t$ (variate difference method) and also between the successive values of r_x. Table A1 gives the average values of the three serial correlations calculated from the values of r_x in respect of various mortality tables. Bearing in mind that the individual values of ρ_1, ρ_2 and ρ_3 show quite large variations (e.g., the range of ρ_1 for the tables averaged in Table A1 is from $-\cdot 61$ to $-\cdot 90$) the agreement between the averages in that table and the theoretical values would seem quite good.

Table A1. *Average of the serial correlations of* r_x *for various mortality experiences*

Mortality experience	ρ_1	ρ_2	ρ_3
8 mortality experiences given in Daw (1945)	−·77	+·20	−·002
6 C.M.I. Assured lives mortality experiences for each of the years 1961–1966 and 6 population mortality experiences of the Netherlands for each of the years 1961–1966 from Daw (1974)	−·73	+·23	+·03
Theoretical values of formula (A10)	−·75	+·30	+·05

Effect of serial correlation between the errors of the original series

A9. In §§ 2.8 and 2.9 it is explained that the methods of constructing the 1972–75 sickness experience introduce a positive serial correlation between the deviations or errors (§ 2.7) in the sickness rates at adjacent ages. One of the effects of such correlation is that σ_r^2 is no longer an estimate of the variance of the deviations in the original series. For the variate difference method, if the error terms of the original series are not independent but have serial correlations R_1, R_2, R_3 etc., then formula (A7) no longer holds. For third differences Kendall and Stuart (1966) formula (46.91) gives

$$\frac{\text{Var}\,(\Delta^3\varepsilon_t)}{20v} = 1 - 1{\cdot}5R_1 + 0{\cdot}6R_2 - 0{\cdot}1R_3 \qquad \text{(A12)}$$

As shown in § A7 the variate difference formulae apply to the r_x test if v is put equal to unity. Thus assuming that the serial correlation in the sickness data is confined to that of lag 1 (i.e., $R_2 = R_3 = \ldots = 0$) then formula (A12) becomes

$$\sigma_r^2 = 1 - 1{\cdot}5R_1 \qquad \text{(A13)}$$

The value of σ_r^2 is therefore reduced by positive serial correlation and increased by negative correlation. Table A2 gives the values of σ_r for a number of values of

Table A2. *Values of* σ_r *in respect of given values of* R_1 *by formula* (A13)

R_1	σ_r	R_1	σ_r
+·5	·509	+·10	·922
+·4	·632	+·05	·962
+·3	·742	zero	1·000
+·25	·791		
+·2	·837	−·2	1·140
+·15	·880	−·25	1·173

R_1. Thus an estimate of R_1, the serial correlation between successive deviations in the original data, can be made from the value of σ_r calculated in respect of the sickness experience.

A10. In the case of the variate-difference method the effect of serial correlations R_1, R_2 . . . etc., between successive errors in the original series is to make the serial correlations ρ_1, ρ_2, ρ_3 differ for each successive order of differences and also to differ from the values given by formula (A11). Quenouille (1953) p. 502 gives formulae for the values of ρ_1, ρ_2, ρ_3 for the variate difference method. These formulae can be adapted for use in the r_x test to calculate values for the first three serial correlations R_1, R_2, R_3 by substituting in Quenouille's formulae the value of σ_r^2 (adjusted for duplicates) and the values of ρ_1, ρ_2, ρ_3 calculated from the successive values of r_x. For third differences, as used in the r_x test, the formulae eventually obtained are

$$\begin{aligned} 1{\cdot}75-\sigma_r^2(1-\rho_1) &= 2{\cdot}80R_1-1{\cdot}40R_2+\ {\cdot}40R_3 \\ {\cdot}70-\sigma_r^2(1-\rho_2) &= \ {\cdot}70R_1+\ {\cdot}40R_2-\ {\cdot}65R_3 \\ 1{\cdot}05-\sigma_r^2(1-\rho_3) &= 1{\cdot}80R_1-1{\cdot}35R_2+1{\cdot}10R_3 \end{aligned} \qquad \text{(A14)}$$

These equations can then be solved to obtain estimates of R_1, R_2, R_3. The value obtained for R_1 will not necessarily be the same as that given by formula (A13) because this assumes that R_2 and R_3 are zero and takes account only of the value of σ_r^2 and not at all of the actual values of ρ_1, ρ_2, ρ_3 calculated from the r_x's. If all the R's in formula (A14) are put equal to zero and σ_r^2 equal to unity, the resulting values of the ρ's are those of formula (A10), as of course they should be, since the terms of the original series are then independent.

Meaning of R_1

A11. In the case of the variate difference method R_1 is quite clearly the serial correlation between the successive errors in the original series, since the correlation coefficient will be the same, whether it is calculated from (i) the successive errors or (ii) the successive errors standardized by division by $\sqrt{v}$. But with the r_x test it is not immediately clear to what function of the original series R_1 relates.

The model which is used is that

$$z_x = F(x) + \varepsilon_x \qquad \text{(A15)}$$

where $F(x)$ is eliminated (or nearly so) by taking third differences of z_x. Thus $\Delta^3 z_x \doteqdot \varepsilon_{x+3} - 3\,\varepsilon_{x+2} + 3\,\varepsilon_{x+1} - \varepsilon_x$. But $r_x = \Delta^3 z_x / \sigma(\Delta^3 \varepsilon_x)$. Now r_x must be the third difference of some function of the errors in the original series of sickness rates but that function is neither the errors of the original series, ε_x, nor the standardized errors, $\varepsilon_x / SD(\varepsilon_x)$ of the original series. It would seem to be some complicated function, perhaps one for which the serial correlation coefficient lies somewhere between that for ε_x and $\varepsilon_x / SD(\varepsilon_x)$.

The earlier part of this Appendix and Section 3 of the paper is written on the basis that the serial correlation (lag 1) between successive values of ε_t is the same

as the R_1 estimated by the methods of §§ A9 and A10, although this would seem to be subject to some degree of doubt. The matter is considered further in §§ 6.14–6.16.

Accuracy of the formulae

A12. It should be pointed out that many of the variate difference formulae given in this Appendix are applicable to long series, since what may be called end-effects have been ignored. For example the mean of $\Delta^3 U_t$ in formula (A8) is assumed to be zero. The end-effects will be small in long series but in short series, such as those for sickness rates of PHI policies considered in this paper covering only 35 terms or less, they could be expected to cause some departures from the formulae. However in view of the complexity of sickness rates this would seem to be unimportant. In calculating the values of σ_r given in Tables 3.1, 3.3 and 3.4 account has been taken of the mean of the values of r_x, although in fact this has rarely made more than a minute difference in the resulting value of σ_r.

Sampling standard deviation of σ_r

A13. Rhodes (1927) gives formulae for the sampling variance of the variance of samples of n from a normal population when successive items are correlated. From these the general formula for the standard deviation of σ_r given in Daw (1945) Appendix 3 was derived, i.e.,

$$SD(\sigma_r) = \frac{\sigma_r}{\sqrt{2n}}\{1 + 2(\rho_1^2 + \rho_2^2 + \rho_3^2)\}^{\frac{1}{2}} \qquad \text{(A16)}$$

where n is the number of values of r_x, σ_r on the right is the estimated population value and the ρ's are the three serial correlations introduced by the process of taking third differences.

Too much reliance must not be placed on formula (A16) as the assumption of normality is somewhat dubious, but it should give a broad indication of the likely spread of values of σ_r (e.g., see § 3.12).

Formula (A16) can be used to determine confidence limits for the values of σ_r. For example the 95% limits would be

$$\text{(Estimated population value of } \sigma_r) \pm 2SD(\sigma_r)$$

Formula (46.85) of Kendall and Stuart (1966) for the variate difference method can be shown, when normality is assumed, to agree with formula (A16) on substituting the values of the ρ's given in formula (A10).

APPENDIX B

CALCULATION OF THE VARIANCE OF SICKNESS

B1. In order to apply the r_x test to the 1972–75 sickness data it is necessary to calculate the variance of sickness at each individual age, the calculations being made separately for each constituent sickness period of the four deferred periods being investigated. Two methods have been used:

Method I is described by Coward (1949) and
Method II is based on the formulae given by Beard (1947).

Method I

B2. Coward (1949) obtained formulae for the variance of sickness, which involved breaking down the graduated sickness rates for each sickness period into rates for each individual week of sickness. His method takes account of the way in which sickness is allocated to individual ages in both the Manchester Unity and the C.M.I. experiences (see § 2.8). Coward uses as his example the Manchester Unity (Whole Society) Experience, 1893–97 for which a breakdown of the graduated sickness rates to individual weeks has already been made in Cmd. 6907 (1913). However for the C.M.I. Sickness Experience 1972–75 these calculations had to be made in respect of the graduated sickness rates.

As there were 14 sets of sickness rates to be broken down to individual weeks, even making the calculations only at quinquennial ages involved some 110 sickness rates. A simple method which could be applied mechanically was therefore needed.

B3. The method of Cmd. 6907, even supplemented by the résumé in Anderson and Dow (1948) p. 279–81, seemed complicated and difficult to apply, and the résumé did not always seem to me to fit in with the method I understand to be described in Cmd. 6907. I therefore tried various finite difference methods involving divided differences. I made many numerical errors and the method often gave impossible (e.g. increasing) sickness rates for some individual weeks. It is amazing what contortions a third degree curve can produce! I came to realize that any method of determining these sickness rates must produce a steadily decreasing set of values.

B4. In what follows a notation will be used which applies only to this Appendix. For a given age and deferred period:

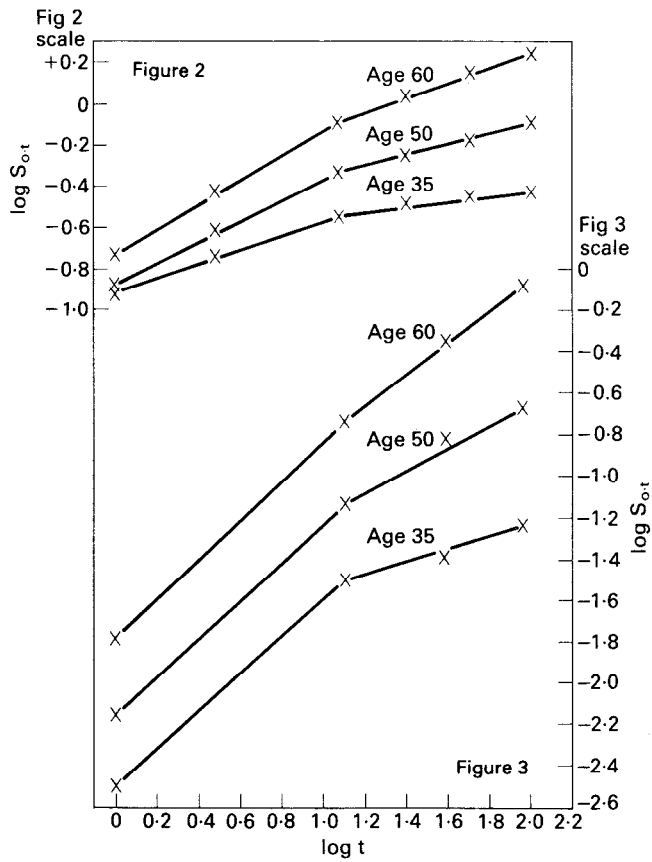
s_t = the rate of sickness for the tth week after the end of the deferred period.

$S_{o\cdot t}$ = the graduated sickness rate from the end of the deferred period to the end of the tth week thereafter.

Thus $S_{o\cdot t} = \sum_{k=1}^{t} s_k$ and $S_{o\cdot 1} = s_1$.

The values of $S_{o\cdot t}$ are obtained by progressive summation of the 1972–75 graduated sickness rates for the particular age of the deferred period. These values are of course limited to certain fixed values of t, which vary for each deferred period.

B5. In searching for a suitable method to determine the values of s_t I drew some graphs. These showed that the graphs of log $S_{o\cdot t}$ against log t could be reasonably represented by two straight lines for deferred periods 1, 4 and 13 weeks and by one straight line for deferred period 26 weeks. Figures 2 and 3 show the graphs for ages 35, 50 and 60 of deferred periods 1 and 13 weeks. The graphs suggested the use of the formula



Plots of log $S_{o\cdot t}$ against log t for D1 (Fig. 2) and D13 (Fig. 3).

$$\log S_{o\cdot t} = a + b \log t \tag{B1}$$

where a and b are constants. Thus the relation between $S_{o\cdot t_1}$ and $S_{o\cdot t_2}$ (where t_1 and t_2 are specified values of t and $t_2 > t_1$) is given by

$$S_{o\cdot t_2} = S_{o\cdot t_1} \times \left(\frac{t_2}{t_1}\right)^b \tag{B2}$$

from which the value of b may be calculated, given $S_{o\cdot t_1}$ and $S_{o\cdot t_2}$.

I thought this formula to be original but later found that Rhodes (1946) had used it for sickness rates, but in a rather more limited situation.

B6. The application of formula (B2) to each pair of consecutive available values of t together with s_1 (or $S_{o\cdot 1}$), which was taken as the graduated claim inception rate, provided a set of values of b (one for each pair) and hence by use of (B2) a series of values of $S_{o\cdot t}$ for $t = 1, 2, 3 \ldots$ For example for deferred period 4 weeks, $S_{o\cdot t}$ was available for $t = 1$, 9, 22, 48 and 100, and formula (B2) was applied to obtain a value of b for each pair of t-values (1, 9), (9, 22), (22, 48) and (48, 100). Then using the appropriate value of b for the pair (1, 9) the values of $S_{o\cdot t}$ for $t = 2$ to 8 were calculated. In this way $S_{o\cdot t}$ for $t = 1, 2, 3 \ldots 100$ were obtained. Taking the first differences of $S_{o\cdot t}$ gave the successive values of s_t ($t = 1$ to 100).

B7. The values of the variance of sickness for the period t_1 to t_2 ($t_2 > t_1$) were obtained by the formulae of Coward (1949):

$$\begin{aligned} U_0 &= \Sigma s_t \\ U_1 &= \Sigma (t - t_1 - \tfrac{1}{2}) s_t \\ U_2 &= \Sigma (t - t_1 - \tfrac{1}{2})^2 s_t \end{aligned} \tag{B3}$$

where the limits of the Σ terms are $t_1 + 1$ to t_2. The variance of the weeks of sickness per person, denoted by $V_{t_1 \cdot t_2}$ is given by the formula

$$V_{t_1 \cdot t_2} = 2U_1 - U_0^2 - \frac{1}{52} U_2 \tag{B4}$$

(The suffix to V can normally be omitted without ambiguity).

The corresponding variance of sickness at age x for exposed to risk E_x is $E_x V_x$ and of the rate of sickness it is $E_x V_x / E_x^2$ or V_x / E_x. The above calculations were made at quinquennial ages and the intermediate values obtained by interpolation on the values of $\sqrt{V}$, which was closer to a straight line than V. Usually first difference interpolation was used for ages up to 55 and second divided differences for ages 55–64.

B8. The successive values of s_t as obtained above usually show a discontinuity at the joint between two sickness periods. Usually there is a gap but sometimes an overlap. Table B1 gives some examples. See also Figure 2.

To see the effect which a method giving a better join between the sections had on the resulting values of V, a laborious method was tried which I thought was on the lines of Cmd. 6907. Out of 14 comparisons which were made the Cmd. 6907

Table B1. *An example of the discontinuities in* s_t *obtained by method I for deferred period 1 week, age 50*

Period of sickness	t	s_t	Period of sickness	t	s_t
S1/3	1	·13527	S13/13	:	:
	2	·06553	(cont)	24	·00613
	3	·05220		25	·00594
	G			G	
S4/9	4	·03633	S26/26	26	·00505
	5	·03174		27	·00490
	:	:		:	:
	11	·02017		50	·00301
	12	·01921		51	·00297
	G			O	
S13/13	13	·00983	S52/52	52	·00339
	14	·00928		53	·00334
	:	:		:	:
				102	·00205
				103	·00203

G = gap; O = overlap.

method gave a lower value of V than did method I in 11 cases (averaging 95% of the method I value) and a higher value (averaging 102%) in 2 cases. In one case it gave an impossible series of s_t with increasing values, but among the other 13 cases there were 5 in which some anomalous values of s_t occurred. However the values obtained for σ_r seemed to be comparatively little affected by small changes in the values of V used, so method I has been adopted for all the sickness periods for which values of σ_r by method I are given in Tables 3.3 and 3.4.

B9. Method I has not been used for sickness period 104/all because Coward's formulae given above apply only to periods of sickness, $t_2 - t_1$, not exceeding 52 weeks. A more complicated formula would be needed for S104/all and it appears that this would involve a breakdown of the total sickness of this period into shorter periods, and that no data are available on which this breakdown could be based. For these reasons no calculations were attempted for this sickness period by method I. However a few were made by method II, giving the values of σ_r for S104/all shown in Table 3.4.

Method II

B10. Table Se 1.2.3 of *C.M.I.R.* **4** (1979) gives numerical values of the moments of weeks of sickness per unit exposed to risk, for quinary age groups of each sickness period. These were calculated from the durations of individual claims in the 1972–75 experience. It is explained that in making the calculations, claims which over-run the year end are truncated at that time and the remainder included in the next year and counted as a separate claim. This is stated to under-

estimate the variance for the duration of sickness claims, and so it does if one is concerned with the whole duration of claims starting in age x. But nevertheless the moments obtained are those applicable to the variance of sickness as treated in the 1972–75 experience and are therefore those needed for a study of the variance of these data. Also the 1972–75 experience includes duplicates and these, I am informed, were included in the moment calculations. Thus these moments would be expected to take account of the effect of duplicates. If therefore they were used to calculate the variance of sickness the resulting values of σ_r should be after taking account of the effect of duplicates, and so would be expected to be less than the corresponding values of σ_r, calculated on the basis of the variances of sickness given by method I, which takes no account of duplicates. This expectation was not realized (§ 3.7).

B11. Beard (1947) and Benjamin and Pollard (1980) p. 153 give the variance of the sickness among E policies as

$$\text{Var (total sickness)} = \text{(total sickness)} \times \frac{m_2}{m_1}$$

where m_1 and m_2 are the first and second moments about zero of the sickness durations of all those who become sick. Hence:

$$\text{Var (sickness rate)} = \frac{\text{total sickness}}{E^2} \times \frac{m_2}{m_1} \qquad \text{B(5)}$$

Table 1.2.3 of *C.M.I.R.* **4** (1979) gives values for quinary age groups of:

$$\text{Sickness rate} = z_y = \frac{\sum_y w}{E_y}$$

$$\text{and } {}_y\mu_2 = \frac{\sum_y w^2}{E_y} - z_y^2$$

where y denotes the age group and w the duration of an individual sickness claim.

$$\text{Now} \quad {}_ym_1 = \frac{\sum_y w}{\text{No. of claims}}, \quad {}_ym_2 = \frac{\sum_y w^2}{\text{No. of claims}}$$

$$\text{so that} \quad \frac{{}_ym_2}{{}_ym_1} = \frac{{}_y\mu_2 + z_y^2}{z_y} \qquad \text{(B6)}$$

and from (B5)

$$\text{Var}(z_y) = \frac{z_y}{E_y} \times \frac{{}_y\mu_2 + z_y^2}{z_y} \qquad \text{(B7)}$$

$$= \frac{{}_y\mu_2 + z_y^2}{E_y}$$

To calculate the variance of the sickness rate at individual ages, the values of the numerators of formula (B7) were allocated to the central age of each age group, and the values for the other individual ages obtained by interpolation as for method I. Division by the exposed to risk for that individual age, gave the variance of the sickness rates at each individual age which were used in the calculation of the values of r_x. Because the values of ${}_y\mu_2$ and z_y^2 were ungraduated the interpolation to individual ages was less satisfactory than in method I.