

INSTITUTE OF ACTUARIES

THE COMPONENTS OF MORTALITY

By H. A. R. BARNETT, F.I.A., A.S.A.

[Submitted to the Institute, 24 January 1955]

The mortality curve is not one simple frequency curve but is made up of several components.—KARL PEARSON

The statistician is never on more dangerous ground than when he passes from the mathematical expression of phenomena to speculation upon their causes.—W. PERKS

I. INTRODUCTORY

No mortality curve appears yet to have been devised which both fits the national statistics at all ages, and also has a simple philosophical explanation. The object of this paper may be stated as being the building up of a new mortality curve from its component parts, as many as possible of these parts having a philosophical background, and the whole being applicable to the national data.

2. Makeham's formula, consisting of one constant term and one increasing throughout life, is capable of simple philosophical explanation, but cannot represent mortality at infantile or childhood ages, nor, indeed, at any ages where mortality decreases, nor does it fit modern mortality over more than a limited range of adult ages without excessive forcing of the data.

3. The family of curves first introduced by Perks⁽⁶⁾ fits over a longer range and may long continue to be the most widely used basis for graduation by curve fitting, but it too does not represent the decreasing mortality at infantile and childhood ages, and may not give a very satisfactory fit to national data. Beard⁽¹⁾ (see p. 416 of his paper) was not wholly satisfied with his attempts at such a fit. Further, there is no simple philosophical explanation of these curves, although, in their simplest form, they usually represent one term decreasing and one increasing throughout life. It is, in the author's opinion, partly the constant D which gives the formulae a somewhat artificial background, and it seems that a more natural representation of the course of mortality at advancing ages should be possible than the introduction of this constant to place some limit on the 'inability to withstand destruction'. It also seems to be artificial, though convenient, to assume the same value of c in the numerator and denominator.

4. The formula suggested by Ogborn⁽⁴⁾ has a philosophical background as described in his written reply to the discussion on his paper, but he was disappointed with the results of the application of his method to national data, which would require more parameters than he felt would justify use of the formula.

5. The formula developed by Phillips⁽⁷⁾ has a philosophical background and gives an expression to which the curve of deaths (as opposed to the rate of mortality) appears to be striving, but does not make any modification to allow for the fact that the basic or ideal curve is never, in fact, attained.

6. As the attempt is being made to fit the statistical facts as shown by the data an alternative title 'The Facts of Death' was considered for the paper,

but this title was discarded because it might seem to imply that evidence of medical facts has been used; this is not so, and, indeed, the object may be further described as an attempt to find out what the statistics themselves show, without tempering the results with any preconceived notion as to what they ought to show. Such conditioning of the results might have taken place had medical evidence been employed. Similarly, the list of references excludes certain works which have followed a strictly medical approach.

7. A supplementary objective, which has appeared in the course of the experiments carried out, is to demonstrate the advantages which might accrue if the Institute's Mortality Data were subdivided according to the cause of death. Such a course has its opponents, and there are admittedly certain difficulties, but it is thought that these can be overcome—both the difficulties and the opponents—and that a subdivided investigation would bring with it valuable lessons. At the worst, we should learn that there is nothing to learn from such an investigation, but the author does not subscribe to this pessimistic view; at the best, we might discover facts which would enrich our science, and this possibility, or indeed probability, should not be discarded merely on grounds of temporary inexpediency. The exciting possibilities are examined in a later section.

II. IDEAS AND DATA

8. The experiments to be described were, in a way, a corollary to some work being performed for the Joint Mortality Investigation Committee (subsequently in this paper referred to as 'the Committee'), and, in fact, many of the considerations in this paper were first set down in the form of a memorandum for the Committee. Although none of the Committee's data is used herein, much of the national data referred to was required for their work in the first place, and the author warmly appreciates and acknowledges the fact that the Committee have raised no objection to the use of the memorandum as a basis of the paper.

9. As a result of a private investigation undertaken early in 1953, in connexion with the controversy and research concerning the causes of lung carcinoma, it appeared that for a great many years the national male rate of mortality from two causes of death only, viz. respiratory tuberculosis and neoplasms of the respiratory system, when combined and plotted graphically against the age, followed closely the shape of the normal curve of error. This may be a fortuitous result so far as these two causes of death are concerned, but it was nevertheless interesting, bearing in mind that a rate, and not a frequency, was being considered; it gave rise to the idea (which is similar to the philosophical ideas expressed by Ogborn⁽⁴⁾ in the reply to the discussion on his paper) that at each age there is an underlying frequency of impaired lives who reach that age having already become likely subjects for death through certain causes—not necessarily confined to the two causes mentioned earlier in this paragraph. Whether or not this frequency of impaired lives would follow the shape of the normal curve would depend on its symmetry, i.e. on whether after the peak, at which deaths of impaired lives would exactly balance new impairments, the fall in the frequency from age to age reflects exactly the rise in the frequency at ages before the peak; but it does seem quite feasible that the ordinate of this frequency curve (with a hump in the middle) when multiplied by whatever may be the annual chance of death applicable to these impaired lives, and divided by the total exposures at the

appropriate age, might give a component part of the rate of mortality of a shape similar to the normal curve.

10. This approach is perhaps open to the criticism that all lives of a certain age might not be subject to the risk of all possible causes of death within a year, but it is not considered that this criticism is valid if the rate of mortality is given its usually accepted definition. In any case, we would not normally have the prior knowledge of which lives were immune from certain types of death within a year, and in the case of some causes the individuals themselves would not be certain either. For example, could any man of 60 say with certainty 'I know I will not die of bronchitis within a year because I am not an impaired life'? It is possible, and indeed likely, that some causes of death might contribute towards more than one of the components of the rate of mortality.

11. The ideas of § 9 suggested the possible value of investigating mortality according to causes of death in broad groups, in the hope that a formula representing the rate of mortality might be built up from its component parts. The author was not aware of any data suitably collected for such an investigation, apart from the Medical Part of the Registrar General's *Annual Reviews* (9). When the work was started, the latest available was the 1949 *Review*, and that for 1950 was published soon after. Before proceeding to describe the experiments, it is essential to set down and bear in mind the disadvantages of using these data, which were, of course, compiled for quite different purposes:

- (a) The exposed to risk are rounded off to thousands;
- (b) The exposed to risk are, in any case, only estimates. It may be recalled that the 1931 population as estimated from the 1921 census differed from that estimated from the 1931 census by as much as 5% in some age groups, and the 1949 and 1950 estimates were built up by a series of approximations covering 18 and 19 years respectively;
- (c) There may be errors or inconsistencies in classification of causes of death, although these may only be small proportionately if the data are not divided into too many subgroups of too small dimensions;
- (d) The effect of migration may have an adverse effect on the census method, 'migration' in this case including transfers to and from the Armed Forces, which would tend to occur in jumps as releases from and intakes to National Service take place;
- (e) There may be misstatements of age on death certificates, not necessarily corresponding to similar misstatements in the census data from which the exposed to risk are derived; for example, there may be a tendency for middle-aged or elderly widows to understate their husbands' ages at death;
- (f) The deaths data for 1950 do not wholly correspond to the exposed to risk. The numerator of the mortality rate includes deaths of non-civilians registered in England and Wales; for the denominator there is the choice between (i) the total population, which includes members of the Armed Forces and Merchant Navy at home and overseas—whereas those overseas are not liable to be included in the deaths recorded and would not necessarily balance out with the Armed Forces of Dominions, Colonies and Foreign Countries temporarily in the country, (ii) the civil population, which clearly would not fit the deaths, and (iii) the home population which includes members of the Merchant Navy overseas; the last has been employed as most nearly fitting the deaths, but it appears that there could be some members of that

population who—apart from the effects of migration—would not, if they died, be included in the numerator. This objection does not apply to the 1949 data, for which the non-civilian deaths are shown separately, and the data used were the civilian population and the civilian deaths;

(g) It is not possible to investigate the mortality rate for ages 85 and over because all these are grouped together; for the same reason it is not possible to apply Hardy's formula for grouped data without giving an upper age limit of 79 to the adjusted data;

(h) The data are available in central form only, and at the higher ages (which it is desirable to be able to include) any attempt to convert to initial form would involve hazardous approximations. For this reason the function operated upon throughout the experiments has been the central rate of mortality; it has been assumed that the central rate for any group applies to the central year of life in that group, and it is thought that the inaccuracies resulting therefrom are insignificant when the foregoing limitations to the data, particularly (b), are taken into account;

(i) The availability of data in quinary groups only may make it difficult to find the precise locations of points of inflexion.

The data, of course, have the advantage that they are national statistics.

12. The foregoing is not intended in any way to be a criticism of the Registrar General and his publication, which is an admirable collection of data, but not in ideal form for the investigation to be described. That it has nevertheless to be so employed is no reflexion on the data, but on the fact that no alternative data are available.

III. EXPERIMENTS WITH MALE DATA FOR 1949

13. When the experiments were begun, the 1949 data were the latest published. It was desired, at that time, to find an expression for the shape of the mortality curve without going to the lengths of evaluating all the constants; by the time the memorandum for the Committee had been prepared, the 1950 data were published, and it was decided to operate upon these more recent data for the purpose of trying to fit figures to the theory—although all the disadvantages described in paragraph 11 apply with equal (or, in the case of (b), greater) effect to the 1950 data.

14. The following conclusions were drawn on examining the 1949 male mortality according to causes of death (see Table 1):

(i) Respiratory tuberculosis and malignant neoplasms in the respiratory system appear on the statistical evidence to form a more regular mortality curve when combined than when considered separately. This combined curve rises to a distinct peak around age 65, and from age 20 to the end of life appears to be the same shape as the normal curve of error.

(ii) Other respiratory diseases and other malignant neoplasms show a tendency for a slackening off in the first differences of the mortality rates. In the first of these two subgroups, although the rates increase throughout life the first differences appear to be 'sluggish' between ages 65 and 80; this may be fortuitous, or may be an inherent feature. In the second subgroup—malignant neoplasms other than respiratory—first differences rise briskly up to age 70 and then fall; the rate itself, however, does not fall until after age 85; it is unlikely that this shape can be attributed to the failings in the data, although inconsistent classification of deaths may account for some of the irregularities.

(iii) The balance, comprising the majority of the causes, appears to follow the shape of a Gompertz curve from age 35 (see column 2 of Table 2), and such a curve can in fact be fitted. The inclusion of 'other respiratory diseases' does not have any appreciable effect on the shape (see column 3 of Table 2). The inclusion of malignant neoplasms other than respiratory does, however, distort the shape (see column 4 of Table 2).

Table 1. Central rates of mortality, subdivided into groups according to causes of death, amongst the male civilian population of England and Wales in 1949

(Rates per thousand)

Age group	Tuberculosis (respiratory)	Malignant neoplasms (respiratory)	Other respiratory diseases	Other malignant neoplasms	All other causes	All causes	Un-natural causes (included in columns 6 and 7)
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
0-	·056	·003	6·705	·056	30·13	36·95	1·352
1-	·028	·003	·335	·057	1·269	1·691	·334
5-	·006	·003	·045	·037	·708	·799	·298
10-	·013	·001	·025	·031	·559	·630	·190
15-	·157	·004	·082	·041	·984	1·268	·380
20-	·395	·018	·084	·068	1·025	1·591	·431
25-	·495	·020	·082	·093	·968	1·657	·298
30-	·520	·040	·125	·126	1·059	1·870	·255
35-	·579	·094	·178	·220	1·413	2·484	·291
40-	·613	·238	·343	·447	2·124	3·766	·292
45-	·802	·601	·747	·824	3·616	6·591	·342
50-	·948	1·014	1·456	1·439	5·668	10·52	·328
55-	1·124	1·525	2·658	2·466	9·723	17·50	·391
60-	1·187	2·047	4·409	4·120	16·76	28·52	·508
65-	1·120	2·228	6·177	6·605	26·84	42·97	·557
70-	·673	1·942	8·470	10·05	45·53	66·67	·771
75-	·472	1·555	12·25	13·17	75·80	103·2	1·456
80-	·252	1·252	17·93	14·75	125·5	159·7	2·429
85-	·185	·984	29·47	13·72	204·5	248·9	4·522

Note. The data from which the above rates were calculated were obtained from Tables 1 and 21 of the Registrar General's *Statistical Review of England and Wales*, 1949 (Part I).

(iv) Some of the middle-age and late middle-age deaths from the 'other malignant neoplasms' and 'other respiratory diseases' groups could be transferred to the normal-curve shaped group (around its peak values) without changing the general shape of the group mortality curve, and possibly the balance could then be added to the Gompertz-shaped group without upsetting that shape. It therefore seems possible that national male mortality above age 35 could be represented by a curve of the form

$$B_1 c_1^x + B_2 c_2^{-x}, \quad (1)$$

The Components of Mortality

where the origin would be approximately at age 65. This conclusion was arrived at in complete ignorance of Thiele's(11) formula, and the author is extremely grateful to Mr M. E. Ogborn for having brought the latter work to light in the historical section of his recent paper(4). The above formula is, in fact, two of the three terms of Thiele's formula, which was published in the *Journal* nearly 80 years before the years to which the data for this paper refer. It may be recalled that Thiele's formula was based on 'a presumed property of the causes of death'.

Table 2. Ratios of central rates of mortality in adjacent groups
(England and Wales, civilian males, 1949)

Central age of group x	m_{x+5}/m_x (excluding all malignant neoplasms and all respiratory causes)	m_{x+5}/m_x (excluding all malignant neoplasms and respiratory tuberculosis)	m_{x+5}/m_x (excluding respiratory tuber- culosis and respiratory malignant neoplasms only)
(1)	(2)	(3)	(4)
37	1.50	1.55	1.61
42	1.70	1.77	1.78
47	1.57	1.63	1.65
52	1.72	1.74	1.73
57	1.72	1.71	1.70
62	1.60	1.56	1.57
67	1.70	1.64	1.61
72	1.66	1.63	1.58
77	1.66	1.63	1.56

IV. UNNATURAL DEATHS

15. If, to the terms of formula (1), is added a curve decreasing from the age of 0, there remains a disturbance between about ages 15 and 30 which appears to be largely attributable to deaths through unnatural causes—particularly road-transport accidents—and to demonstrate this disturbance the unnatural death-rates for males have been included as a final column to Table 1. For the purpose of this paper, the adjective 'unnatural' is taken as applying to anything resulting entirely from human action, intentionally or unintentionally, corresponding to the International Classification Groups E800 to E999; accidental deaths would thus be included, but certain occupational diseases, which are partly due to man-made environment but partly to other factors, would not. In the later years of life, the rates continue to increase in such a way that they can mingle with the 'Other causes' group without difficulty; whether the risk of having an accident increases with the middle and late ages is only one factor, but it is an indisputable fact that, having had an accident, an old person is less likely to recover than a young one, and thus it seems that once the unnatural factor has played its part, natural forces come into play in determining the death-rate. Similarly, at the ages of infancy and early childhood, the rates of unnatural death decrease with age as, indeed, do the rates of death from all causes combined. However, between the ages of, perhaps, 15 and 30 (the precise limits cannot be determined from quinary data), when it might be expected that the chance of

recovery from the operation of unnatural factors would be at a maximum, and comparatively independent of age, the death-rates from these causes are such that it is difficult to merge them with the deaths from other causes. There appears to be an additional curve, perhaps of a shape also similar to the normal curve, which reflects the heavy accident rate at these ages. A further reference to the data would show that much of the disturbance in the rates is due to deaths through transport accidents. A similar consideration of female deaths indicates that the disturbance never rises to such a high peak, and that the unnatural death-rate is never so high for females, although it is high enough to make the disturbance appreciable.

16. Since this part of the paper was completed and its conclusions drawn, but before the experiments had been extended to the 1951 national data, Bowerman's work⁽²⁾ was published, in which he concludes, after reviewing a number of cause of death investigations in various countries, that the dip and trough in the mortality curve in the twenties is, subject to certain limitations, due to accident and tuberculosis. The national data now under consideration show that in England and Wales the feature is not so much a pronounced dip as a flattening of the curve, although the male curve does display a small shallow trough in the middle twenties. (The question of the extent to which tuberculosis accounts for this feature of the male mortality curve is considered further in § 65; there seems to be no reason why this should be regarded as in any way inconsistent with the respiratory tuberculosis deaths at the later ages subscribing to the normal curve shape.)

V. EXPERIMENTS WITH MALE DATA FOR 1950

17. The experiments with the 1949 data had reached the stage described in section III and the memorandum relative to it submitted to the Committee, when the 1950 data were published, and with the 1949 results in mind it was decided to subdivide the 1950 data into the following groups:

- I. All deaths due to malignant neoplasms.
- II. All deaths due to tuberculosis.
- III. All deaths not included in I, II and IV.
- IV. Unnatural deaths.

The deaths in groups I, II and III were further subdivided according to whether the site of the disease causing death was

- (a) The respiratory system, or
- (b) Any other site.

18. It should be mentioned that the causes of death shown in the 1950 data are in accordance with the latest International Statistical Classification. The International Classification numbers are shown, for the convenience of anyone interested, in the headings of Tables 3, 7, 14 and 15.

19. The deaths in age groups and cause groups are shown in Table 3, and the central rates of mortality in Table 4. An additional column is included in Table 4 to show the combined rates due to respiratory tuberculosis and respiratory malignant neoplasms, which again follow a fairly regular series. The same sluggishness in the first differences of the rates due to 'Other malignant neoplasms' and 'Other respiratory diseases' after age 65 is apparent as in the case of the 1949 data. And the same disturbance in the unnatural death rate is seen between ages 15 and 35, after which the rate increases fairly regularly throughout life.

The Components of Mortality

20. Table 5 shows the central rates of mortality with ratios between adjacent values for certain combinations of cause of death groupings, for the purpose of finding whether any section of the data is amenable to a Gompertz fit. As before, the most promising combination is that excluding all malignant

Table 3. Deaths registered in England and Wales in 1950 subdivided into groups according to cause of death (males)

(The figures in square brackets show the International Classification)

Age group	Malignant neoplasms (respiratory system) [160-165] I (a)	Other malignant neoplasms [140-159 and 170-205] I (b)	Respiratory tuberculosis [001-008] II (a)	Other respiratory diseases [470-527] III (a)	Un-natural causes [E 800-E 999] IV	All other causes II (b) and III (b)	Total deaths all causes
0-	1	22	28	1,717	456	9,834	12,058
1-	0	39	17	241	104	465	866
2-	0	44	12	124	133	272	585
3-	0	56	6	56	106	229	453
4-	0	40	10	42	80	131	303
5-	2	108	14	119	395	538	1,176
10-	2	74	11	61	221	434	803
15-	5	110	105	79	506	561	1,366
20-	10	158	341	80	824	655	2,068
25-	40	211	656	143	782	875	2,707
30-	59	255	602	185	566	1,012	2,679
35-	152	471	684	292	671	1,644	3,914
40-	434	800	763	534	742	2,678	5,951
45-	940	1,448	1,077	1,040	848	4,448	9,801
50-	1,495	1,946	1,041	1,598	787	6,655	13,522
55-	1,852	2,842	1,107	2,581	820	9,783	18,994
60-	2,149	3,908	980	3,902	815	15,138	26,892
65-	1,858	5,111	821	4,559	815	20,770	33,934
70-	1,328	5,950	408	4,838	730	26,404	39,658
75-	700	5,137	187	4,448	713	27,864	39,049
80-	252	2,523	58	3,078	507	20,931	27,349
85-	61	977	6	2,064	275	13,641	17,024
Total all ages	11,340	32,230	8,934	31,781	11,905	164,962	261,152

Note. The data for this table were taken from Table 17 of the Registrar General's *Statistical Review of England and Wales, 1950* (Part I).

neoplasms and all respiratory diseases—leaving groupings II (b), III (b) and IV—but it must be remembered that the conclusion from the 1949 data was that part of groupings I (b) and III (a) would eventually be added to the Gompertz component of the mortality curve.

21. Attempts were made to find a fit of formula (1) direct from the data without having to use trial and error methods, but this was found apparently

to be impracticable, perhaps owing to the high moments it was necessary to use. As the result of magnification of statistical and other errors in the data, such a method gave absurd results, and it was considered preferable to start with a trial value of c_1 . Further research might, however, be possible on this question of a more direct method of fitting.

Table 4. Central rates of mortality, subdivided into groups according to causes of death, amongst the male population of England and Wales in 1950

(Rates per thousand)

Age group	$10^6 \times$ reciprocal of estimated home population	Central rates per 1000, cause groupings as in Table 3							
		I (a)	I (b)	II (a)	III (a)	IV	II (b) and III (b)	All groups com- bined	I (a) and II (a)
0-	2.801	.003	.062	.078	4.809	1.277	27.545	33.774	
1-	2.688	.000	.105	.046	.648	.280	1.250	2.328	
2-	2.519	.000	.111	.030	.312	.335	.685	1.474	
3-	2.268	.000	.127	.014	.127	.240	.519	1.027	
4-	2.941	.000	.118	.029	.124	.235	.385	.891	
5-	.6373	.001	.069	.009	.076	.252	.343	.749	
10-	.6935	.001	.051	.008	.042	.153	.301	.557	
15-	.7429	.004	.082	.078	.059	.376	.417	1.015	
20-	.6729	.007	.106	.229	.054	.554	.441	1.392	
25-	.5977	.024	.126	.392	.085	.467	.523	1.618	.416
30-	.6601	.039	.168	.397	.122	.374	.668	1.768	.436
35-	.5907	.090	.278	.404	.172	.396	.971	2.311	.494
40-	.5924	.257	.474	.452	.316	.440	1.586	3.525	.709
45-	.6540	.615	.947	.704	.680	.555	2.909	6.410	1.319
50-	.7734	1.156	1.505	.805	1.236	.609	5.147	10.458	1.961
55-	.9191	1.702	2.612	1.017	2.372	.762	8.992	17.456	2.719
60-	1.053	2.263	4.115	1.032	4.109	.858	15.940	28.318	3.295
65-	1.276	2.371	6.522	1.048	5.817	1.040	26.503	43.300	3.419
70-	1.678	2.228	9.984	.685	8.118	1.225	44.306	66.546	2.913
75-	2.681	1.877	13.772	.501	11.925	1.912	74.703	104.690	2.378
80-	5.917	1.491	14.929	.343	18.213	3.000	123.849	161.825	1.834

Rates for the age group 85 and over are not included because the age distribution is not available.

22. It was decided to use the figures in Table 5 merely to find a first trial value of the Gompertz constant c_1 , and the combination of groupings mentioned in the last paragraph suggested that $c_1^5 = 1.67$ might be suitable; on trial this was, however, found not to fit very well, and it was considered better to try a value more in line with the later values for all groupings combined except I (a) and II (a), the higher crude values of c_1^5 at the earlier ages being satisfied by increasing the values of the second (i.e. the normal curve shaped) part of the curve. It seemed likely that at the ages of 37 and 82 the second part of the curve would only represent a small proportion of the

The Components of Mortality

mortality rate, and the rates from all causes combined at these two ages only indicated a value of c_1^5 equal to 1.603. It was not difficult to superimpose a normal curve with its origin at age $69\frac{1}{2}$. An examination of the mortality at childhood ages, excluding the effects of the extended Gompertz curve, indicated that a harmonic curve with its origin before birth gave a much better fit than either a harmonic curve based on the square root of the age or a decreasing Gompertz curve, apart from the first two years of life where mortality may well be expected to be affected by birth and where special factors can be regarded as giving rise to a separate component of the mortality rate (see Perks's remarks on Ogborn's paper⁽⁴⁾). For the sake of continuity this harmonic curve had to be continued throughout life, and this necessitated

Table 5. Values of $1000 m_x$ for certain combinations of cause of death groupings, and ratios of m_{x+5} to m_x ; England and Wales, males, 1950

Age group	Age last birthday at central age of group x	Groupings II (b) and III (b) combined		Groupings II (b), III (b) and IV combined		Groupings I (b), II (b), III and IV combined	
		$1000 m_x$	$\frac{m_{x+5}}{m_x}$	$1000 m_x$	$\frac{m_{x+5}}{m_x}$	$1000 m_x$	$\frac{m_{x+5}}{m_x}$
35-	37	.971	1.63	1.367	1.48	1.817	1.55
40-	42	1.586	1.83	2.026	1.71	2.816	1.81
45-	47	2.909	1.77	3.464	1.66	5.091	1.67
50-	52	5.147	1.75	5.756	1.69	8.497	1.73
55-	57	8.992	1.77	9.753	1.72	14.737	1.70
60-	62	15.940	1.66	16.799	1.64	25.023	1.59
65-	67	26.503	1.67	27.542	1.65	39.881	1.60
70-	72	44.306	1.69	45.531	1.68	63.633	1.61
75-	77	74.703	1.66	76.615	1.66	102.312	1.56
80-	82	123.849	—	126.849	—	159.991	—

some minor adjustments in the other two sections of the mortality curve, a final fit being made with $c_1 = 1.100$ and $c_2^5 = 1.024$. Details of this curve are shown in Table 6. The table also includes the numbers of unnatural deaths at the young ages, for comparison with the difference between the actual and expected deaths at the ages where the disturbance occurs, and a column showing 5% of the expected deaths, bearing in mind limitation (b) on the data (see section II).

23. In view of the figures in this last column of Table 6 it was not thought to be worth while to carry out any adherence tests in great detail; the table shows that, except at ages 0, 1 and 17-32 (the figures for which are shown in brackets in column 9 of the table) the deviations change sign reasonably often, and that the fit is as good as can be judged bearing in mind the limitations which must be placed on the data. In other words, the evidence is that the process has been curve fitting rather than curve forcing.

24. No attempt has been made to express in algebraic terms the unnatural disturbance at the adolescent and young adult ages; as this only spreads over about 20 years of age, and the data are restricted to four quinary rates, the true shape of this mortality component can only be guessed. It might be

another curve of the normal shape, it might be a parabola, or it might be skew, but without further details age by age it would be somewhat speculative to try to determine its precise shape. Columns 9 and 10 of Table 6 confirm that unnatural deaths largely account for the disturbance in question. That they do not do so entirely may be due to the fit; there is also the possibility that in some cases where the death of an invalid has been hastened by an unnatural cause the primary classification of the death is into some other grouping.

Table 6. A fit of formula (2) to the central death rates amongst the male population of England and Wales in 1950

Age		1000	1000	1000A	1000	Ex-	Actual	Deviations	Un-	5 %
y	x	$B_1c_1^x$	$B_2c_2^{-x^2}$	$\frac{A}{y+1\frac{1}{2}}$	m_y	pected	deaths	(8)-(7)	natural	of
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	deaths	col.
								(10)-(7)	below	(7)
									age 35	
0	-69½	·065	·000	3·600	3·665	1,308	12,058	+ (10,750)	456	65
1	-68½	·071	·000	2·000	2·071	770	866	(96)	104	39
2	-67½	·078	·000	1·385	1·463	581	585	4	133	29
3	-66½	·086	·000	1·059	1·145	505	453		106	25
4	-65½	·095	·000	·857	·952	324	303	52 21	80	16
7	-62½	·126	·000	·545	·671	1,053	1,176	123	395	53
12	-57½	·203	·000	·340	·543	783	803	20	221	39
17	-52½	·326	·000	·247	·573	771	1,366	(595)	506	39
22	-47½	·526	·000	·194	·720	1,070	2,068	(998)	824	54
27	-42½	·847	·001	·159	1·007	1,685	2,707	(1,022)	782	84
32	-37½	1·363	·007	·135	1·505	2,280	2,679	(399)	566	114
37	-32½	2·196	·035	·118	2·349	3,977	3,914	63		199
42	-27½	3·536	·145	·104	3·785	6,389	5,951	438		319
47	-22½	5·695	·476	·093	6·264	9,578	9,801	223		479
52	-17½	9·172	1·228	·085	10·485	13,557	13,522	35		678
57	-12½	14·771	2·502	·077	17·350	18,877	18,994	117		944
62	-7½	23·789	4·021	·071	27·881	26,487	26,892	405		1,324
67	-2½	38·312	5·097	·066	43·475	34,084	33,934	150		1,704
72	+2½	61·702	5·097	·061	66·860	39,849	39,658	191		1,992
77	+7½	99·371	4·021	·058	103·450	38,587	39,049	462		1,929
82	+12½	160·038	2·502	·054	162·594	27,478	27,349	129		1,374

25. Likewise, no attempt has been made to fit a curve to those deaths in the first two years of life which are not accounted for by the Gompertz and harmonic components; this shape could not be determined satisfactorily without details for each month of age—or at least each quarter.

26. In Clarke's centenary paper (3), the distinction was drawn between 'senescent' and 'anticipated' deaths; it was that conception which gave rise to the ideas underlying the experiments now being described, and it is sincerely acknowledged that without Clarke's paper this would never have been written. The division of the deaths is now developed further into five component parts, i.e. (a) senescent deaths, (b) normal anticipated deaths,

(c) childhood anticipated deaths, (d) natal deaths, and (e) early anticipated deaths, chiefly due to unnatural causes. These may be compared with the components arrived at by Pearson(s) who was, however, considering frequencies of deaths rather than mortality rates.

27. The first three of these components may be represented by the formula

$$m_y = B_1 c_1^x + B_2 c_2^{-x} + A/(y + 1\frac{1}{4}), \quad (2)$$

and the values of the constants for the curve evaluated in Table 6 are as follows:

$$1000 \quad B_1 = 48.62$$

$$c_1 = 1.100$$

$$1000 \quad B_2 = 5.250$$

$$c_2^5 = 1.024$$

$$1000 \quad A = 4.500$$

x is measured from the origin of $69\frac{1}{2}$.

y is measured from birth.

28. It has already been remarked that the fact of the data being quinary may make it difficult to fix the precise position of a point of inflexion, and for this reason it is possible that the true origin should be elsewhere than exactly midway between the two age groups 65- and 70-. It may be mentioned that the origin of $69\frac{1}{2}$ for the second term of formula (2) as applied to m_y would correspond to an origin of 70 if a similar formula were to be fitted to μ_y . (The first term can, of course, be measured from any origin, since a change in origin can be adjusted by a change in the value of B_1 ; not so the second and third terms.)

29. It is interesting also to note that if the central rate of mortality for a year of life is regarded as an approximation to the force of mortality at the centre of that year, then the third term gives the force of mortality an infinite value at an age three-quarters of a year before birth. Pearson(s), and more recently Phillips(7), came to a similar conclusion concerning the frequency of deaths in infancy.

VI. EXPERIMENTS WITH FEMALE DATA FOR 1950

30. The deaths for the female data, subdivided into the same groupings as the male data, are shown in Table 7, and the central rates of mortality in Table 8. It is clear that the female mortality curve has, in certain respects, fundamental differences from the male curve; the rate of mortality through respiratory malignant neoplasms not only is of much smaller dimensions, but also increases throughout life up to the age group 75-79; respiratory tuberculosis rises rapidly, from almost negligible amounts below age 15 to a maximum between ages 25 and 30, and thereafter tends to decrease throughout life. The disturbance due to unnatural deaths merely takes the form of a temporary sluggishness; the rate falls up to the age group 10-14, remains practically constant—at a low level compared with the males—from ages 15 to 34, and thereafter increases fairly regularly.

31. Table 9 shows the central rates, together with ratios of adjacent values, for various combinations of groupings, for the purpose of estimating a suitable value for c_1 , if it should prove possible to assign a Gompertz shape

to part of the mortality curve. Columns are included in this table for all causes combined. It is seen that for each combination tried, the ratios continue to increase right up to age 75 or 80, and the implication is that if formula (1) is again applicable to all but the childhood and young adult ages, the origin of the second term occurs very late in life, with the effect that there

Table 7. Deaths registered in England and Wales in 1950 subdivided into groups according to cause of death (females)

(The figures in square brackets show the International Classification)

Age group	Malignant neoplasms (respiratory system) [160-165] I (a)	Other malignant neoplasms [140-159 and 170-205] I (b)	Respiratory tuberculosis [001-008] II (a)	Other respiratory diseases [470-527] III (a)	Un-natural causes [E 800-E 999] IV	All other causes II (b) and III (b)	Total deaths (all causes)
0-	0	28	19	1,389	314	7,009	8,759
1-	0	30	17	238	85	429	799
2-	1	34	6	85	95	233	454
3-	1	51	6	57	64	200	379
4-	0	29	4	34	56	125	248
5-	0	100	12	82	190	411	795
10-	2	60	21	59	65	360	567
15-	5	88	276	79	111	530	1,089
20-	1	81	648	83	122	717	1,652
25-	8	223	806	125	128	999	2,289
30-	26	369	624	115	127	1,115	2,376
35-	55	756	512	236	188	1,647	3,394
40-	114	1,424	426	312	240	2,045	4,561
45-	139	2,286	363	403	341	3,381	6,913
50-	248	3,081	343	598	349	4,831	9,450
55-	280	3,908	269	916	403	7,139	12,915
60-	339	4,785	264	1,415	408	11,559	18,770
65-	383	5,568	234	2,182	479	18,204	27,050
70-	353	6,022	157	3,243	637	26,416	36,828
75-	243	5,323	88	4,066	850	31,801	42,371
80-	110	3,236	41	3,570	836	28,244	36,037
85-	45	1,865	9	3,428	896	25,210	31,453
Total all ages	2,353	39,347	5,145	22,715	6,984	172,605	249,149

Note. The data for this table were taken from Table 17 of the Registrar General's *Statistical Review of England and Wales, 1950* (Part I).

are two curves increasing over the greater part of the range of adult ages. It was difficult to know where to start, and as a shot in the dark the Gompertz curve applicable to the male data was tried, moving it up 5 years in age; several attempts were made to superimpose a second Gompertz curve, but this shape appeared to increase too rapidly at the older ages; a normal curve was superimposed (the origin being at age 112), but, though this gave a reasonable fit at the older ages, the expected deaths at the middle ages were

Table 8. Central rates of mortality, subdivided into groups according to causes of death, amongst the female population of England and Wales in 1950
(Rates per thousand)

Age group	$10^6 \times$ reciprocal of estimated home population	Central rates per 1000, cause groupings as in Table 7						
		I (a)	I (b)	II(a)	III(a)	IV	II (b) and III (b)	All groups combined
0-	2.933	.000	.082	.056	4.074	.921	20.557	25.690
1-	2.825	.000	.085	.048	.672	.240	1.212	2.257
2-	2.632	.003	.089	.016	.224	.250	.613	1.195
3-	2.387	.002	.122	.014	.136	.153	.477	.905
4-	3.086	.000	.089	.012	.105	.173	.386	.765
5-	.6662	.000	.067	.008	.055	.127	.274	.530
10-	.7179	.001	.043	.015	.042	.046	.258	.407
15-	.7194	.004	.063	.199	.057	.080	.381	.783
20-	.6627	.001	.054	.429	.055	.081	.475	1.095
25-	.5949	.005	.133	.479	.074	.076	.595	1.362
30-	.6498	.017	.240	.405	.075	.083	.725	1.544
35-	.5811	.032	.439	.298	.137	.109	.957	1.972
40-	.5848	.067	.833	.249	.182	.140	1.196	2.667
45-	.6223	.086	1.423	.226	.251	.212	2.104	4.302
50-	.6752	.167	2.080	.232	.404	.236	3.262	6.381
55-	.7553	.211	2.952	.203	.692	.304	5.392	9.755
60-	.8410	.285	4.024	.222	1.190	.343	9.721	15.786
65-	.9718	.372	5.411	.227	2.120	.465	17.691	26.287
70-	1.232	.435	7.419	.193	3.995	.785	32.545	45.372
75-	1.873	.455	9.970	.165	7.616	1.592	59.563	79.361
80-	3.597	.396	11.640	.147	12.841	3.007	101.594	129.625

Rates for the age group 85 and over are not included.

Table 9. Values of $1000 m_x$ for certain combinations of cause of death groupings, and ratios of m_{x+5} to m_x ; England and Wales, females, 1950

Age group	Age last birthday at central age of group x	Groupings II (b) and III (b) combined		Groupings II (b), III and IV combined		Groupings II, III and IV combined		All groupings combined	
		$1000 m_x$	$\frac{m_{x+5}}{m_x}$	$1000 m_x$	$\frac{m_{x+5}}{m_x}$	$1000 m_x$	$\frac{m_{x+5}}{m_x}$	$1000 m_x$	$\frac{m_{x+5}}{m_x}$
35-	37	.957	1.250	1.203	1.263	1.501	1.178	1.972	1.352
40-	42	1.196	1.759	1.519	1.690	1.768	1.580	2.667	1.613
45-	47	2.104	1.550	2.567	1.520	2.793	1.480	4.302	1.483
50-	52	3.262	1.653	3.901	1.638	4.133	1.595	6.381	1.529
55-	57	5.392	1.803	6.389	1.762	6.592	1.741	9.755	1.618
60-	62	9.721	1.820	11.255	1.802	11.477	1.787	15.786	1.665
65-	67	17.691	1.840	20.277	1.841	20.504	1.830	26.287	1.726
70-	72	32.545	1.830	37.325	1.842	37.518	1.837	45.372	1.749
75-	77	59.563	1.706	68.771	1.708	68.936	1.706	79.361	1.633
80-	82	101.594	—	117.442	—	117.589	—	129.625	—

much too low. It became clear that a lower value of c_1 was necessary, nearer the crude values shown in the last column of Table 9 between ages 35 and 60, where it might be expected that the normal curve or second Gompertz would give relatively low values. This conclusion having been reached it did not take long to find a fit with $c_1^5 = 1.525$, with normal and harmonic curves superimposed as in the case of the male curve. As before, the procedure was to find a normal curve which, when added to the Gompertz curve, reproduced the older rates fairly faithfully; a harmonic curve was again found to be the best shape which, when added to the Gompertz curve values at the childhood ages (other than the first two years of life), reproduced the rates at those ages; finally, when the harmonic curve had been extended throughout life, some minor adjustments were made to the other two curves. The fit of formula (2) is shown in Table 10, the constants being as follows:

$$\begin{aligned} 1000 \quad B_1 &= 185.0 \\ c_1^5 &= 1.525 \\ 1000 \quad B_2 &= 82.00 \\ c_2^5 &= 1.025 \\ 1000 \quad A &= 3.300 \end{aligned}$$

x is measured from the origin of 92.

y is measured from birth.

32. Again, formula (2) does not, nor was it expected that it would, encompass all the deaths in the first two years of life, and again there is a disturbance from ages 15 to 35, the deviations for all these ages being shown in brackets in column 9 of Table 10. But in the case of the females, this disturbance cannot wholly—or even largely—be attributed to unnatural deaths. The main factor here appears to be the hump in the rate of mortality from respiratory tuberculosis. Column 10 of Table 10 shows the combined deaths due to this cause and to unnatural causes, for comparison with the number of deaths not included within the formula (2) curve. Some might consider that the similarity of the age-range over which the disturbance occurs for both sexes is too coincidental to be due to separate causes; the author's opinion, however, is that the cause of death statistics speak for themselves and that there is no one cause underlying the phenomenon applicable to each sex.

33. Subject to the previous paragraph, and to the limitations on the data, the curve again appears to have been fitted rather than forced.

34. It was considered interesting to try to trace the causes of death which might be chiefly represented by the normal curve of formula (2)—corresponding to respiratory tuberculosis and neoplasms in the case of the males. It was found that the only main group where the mortality seemed to be persistently heavier for females was 'Diseases of the nervous system and sense organs', and this group, together with malignant neoplasms and respiratory diseases, more than covers the normal anticipated deaths corresponding to column 4 of Table 10. It is tempting to speculate whether we might not have expected the males to have stronger nervous systems while the females have stronger respiratory systems! It will have been noticed that the subdivision of the data suggested in section V has not helped in the fit of formula (2) to the female data, and it is possible that if a cause of death investigation were to be made for females it would be advisable to record separately whether the site of the disease causing death was the nervous system.

VII. SUMMARY OF CONCLUSIONS FROM THE
1949 AND 1950 DATA

35. The experiments with the 1950 data have been described at some length, and it may be useful to consolidate the ideas expressed so far, arising from the considerations which have been given to the data for 1949 and 1950. The first conclusion is that the data are consistent with the suggestion

Table 10. A fit of formula (2) to the central death rates amongst the female population of England and Wales in 1950

Age		$1000 B_1 c_1^x$	$1000 B_2 c_2^{-x^2}$	$\frac{1000A}{y + \frac{1}{4}}$	$1000 m_y$	Ex- pected deaths	Actual deaths (all causes)	Deviations (8) - (7)	Deaths due to respira- tory tuber- culosis and unnatural causes between ages 15 and 35	5 % of col. (7)
<i>y</i>	<i>x</i>									
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
0	-92	·079	·000	2·640	2·719	927	8,759	+ (7,832)		46
1	-91	·085	·000	1·467	1·552	549	799	(250)		27
2	-90	·093	·000	1·015	1·108	421	454	33		21
3	-89	·101	·000	·776	·877	367	379	12		18
4	-88	·110	·000	·629	·739	239	248	9		12
7	-85	·142	·000	·400	·542	814	795	19		41
12	-80	·216	·000	·249	·465	648	567	81		32
17	-75	·330	·000	·181	·511	710	1,089	(379)	387	36
22	-70	·503	·000	·142	·645	973	1,652	(679)	770	49
27	-65	·767	·000	·117	·884	1,486	2,289	(803)	934	74
32	-60	1·169	·000	·099	1·268	1,951	2,376	(425)	751	98
37	-55	1·783	·000	·086	1·869	3,217	3,394	177		161
42	-50	2·719	·000	·076	2·795	4,779	4,561	218		239
47	-45	4·147	·004	·068	4·219	6,780	6,913	133		339
52	-40	6·324	·030	·062	6·416	9,502	9,450	52		475
57	-35	9·645	·193	·057	9·895	13,101	12,915	186		655
62	-30	14·708	·963	·052	15·723	18,695	18,770	75		935
67	-25	22·430	3·744	·048	26·222	26,982	27,050	68		1,349
72	-20	34·205	11·374	·045	45·624	37,047	36,828	219		1,852
77	-15	52·163	26·992	·042	79·197	42,291	42,371	80		2,115
82	-10	79·549	50·043	·040	129·632	36,038	36,037	1		1,802

that the rate of mortality is made up of five component parts; these components are the respective rates of senescent death, normal anticipated death, childhood anticipated death, natal death and early anticipated death.

36. The second conclusion is that the first three of these components can be represented by the formulae

$$m_y = B_1 c_1^x + B_2 c_2^{-x^2} + A/(y + \frac{1}{4}) \quad (2)$$

or

$$\mu_y = B_1 c_1^x + B_2 c_2^{-x^2} + A/(y + \frac{3}{4}), \quad (2a)$$

where *y* is measured from birth and *x* from some other origin.

37. The third conclusion is that it would be useful if, in the future, the Committee's deaths data were to be divided into four groups, viz.

- I. Deaths due to malignant neoplasms.
- II. Deaths due to tuberculosis.
- III. Deaths not included in I, II and IV.
- IV. Unnatural deaths.

The first three groups would be subdivided according to whether the site of the disease causing death was (a) the respiratory system, (b) the nervous system, or (c) any other site; for male data, (b) and (c) could be combined. There might also be some advantage in subdividing Group IV according to whether or not death was due to a transport accident. The uses to which the subdivided data might be put are considered in the next section. It is, of course, realized that with the exception of annuitants the Committee's data are practically limited to males; also that the incidence of deaths by cause may differ between the Committee's and the national data.

38. It would be necessary for those scrutinizing the death certificates for this purpose to be aware of the instructions given for certification of deaths. These are described fully in Supplement 3 to the *Bulletin of the World Health Organization*, and more briefly in an article by W. P. D. Logan, M.D., Ph.D., in *B.M.J.* 6 June 1953, **1**, 1272. It is clear from these instructions that, for statistical purposes, the cause most likely to be required is the last to be recorded in Part I, i.e. the one which arose first, which is described in the instructions as 'underlying'. It is not suggested, however, that the other stated causes should be completely ignored, and it would therefore be desirable to record the full stated causes in each case, to enable those undertaking the investigation to determine whether any of the other causes are relevant. Although it would not take long to record the full causes each time, large offices might prefer to take microfilms of death certificates for later use by the Committee.

VIII. USES OF THE SUBDIVIDED DATA

39. The primary purpose of this paper is not to find an alternative to the Perks formula for use when graduating by curve fitting. Having found a not too complex curve which appears to represent the national data, we should do well, it might be thought, to let the matter rest there. It is, however, submitted that future data, subdivided as suggested, would have certain definite uses, to say nothing of exciting possibilities, which would probably outweigh the inconvenience of recording particulars of the deaths in rather more detail than at present.

40. In the first place, it is thought that a formula fitting the data naturally is likely to be useful for purposes of extrapolation. Projecting into the future by extrapolating on constants is an appealing proposition, but its success is likely to be limited if the formula used has to be forced to the data since the constants may be to a certain extent interdependent; even in the case of the Perks formula, there is some interdependence between the constants B and D , which between them fix the limit to be placed on the rate of mortality, and it is possible that a fairly regular secular improvement in the rates, as represented by a Perks fit, might produce apparently irregular changes in both B and D . In formula (2) however, whether or not a further term is added to represent early anticipated deaths, the constants are completely independent, apart from the fact that the value of B_1 depends on the position of the origin x .

The Components of Mortality

41. It should be mentioned here that Pollard⁽⁸⁾ has described various methods of projection which have been tried, and more recently Starke⁽¹⁰⁾ has described some experiments using a formula, similar to Perks, but which had, up to a point, to be forced to the data. Pollard also considered that the 'cause of death' method had possibilities for extrapolation, but did not mention that if too many groups each containing too few causes are used, it is possible to produce absurd results.

42. Secondly, particulars of causes of death amongst assured lives' or annuitants' data would almost certainly throw much light on the direct effects of class selection, self-selection, and medical selection. Selection, in either permanent or temporary form, is probably merely the process of partial elimination of a number of causes of death, not necessarily entirely confined to normal anticipated deaths. It is considered that this use alone would justify the subdivision of the deaths data.

43. Thirdly, if an alternative curve should be desired for graduation purposes, either of Thiele form or of a double Gompertz form, it would be almost impossible in some cases to arrive at suitable first-trial values of the exponential constants without having available the natural component parts of the data. Such a curve might be desired with a view to having standard tables which, while based on experience, would at the same time have possibilities for extrapolation (see § 40). Or it might be desired if it were considered necessary to use a curve which could give true representation over the range where mortality drops as age increases.

44. Fourthly, the data in this form would enable an investigation to be made into the true nature of the mortality curve without having to rely on data unintended for, and not wholly suited to, such a use. It is agreed that such an investigation would not necessarily have nation-wide application, but it would be very informative and, for the purpose of having an approximation to national experience for comparison with assured lives, it is possible that industrial assurance data could be used. It appears that none of the disadvantages set out in section II of this paper would apply, except for errors or inconsistencies in classification of causes of death. Errors should be rare and become still rarer with the further advance of medical science. Inconsistencies are likely to be confined to differences of opinion as to which cause of death is the primary cause, and to some extent this disadvantage would be overcome if the full stated causes were recorded.

45. It may be mentioned in passing that such an investigation would make imperative the exclusion of duplicates, at least from the deaths, and the recording of full details of causes might facilitate the removal of duplicates as between different offices.

46. A final advantage which might accrue with closer scrutiny of the death certificates is that a statistical investigation would be possible which would throw light on the effect on national mortality data of misstatements of ages on death certificates—which although unlikely to be individually large may consist of numerous small errors possibly deliberate and biased (e.g. widows taking a year or two off their husbands' ages). It would, of course, be easy to compare the age recorded on the death certificate with the proved age.

47. It need hardly be added that the earlier the data are collected in the recommended form, the earlier will these advantages accrue.

IX. COMPARISON OF FORMULAE (1) AND (2) WITH PERKS

48. It has been hinted in the previous section that one of the advantages of subdividing the deaths according to cause might be the possibility of graduating by a curve other than Makeham or Perks. The third term of formula (2) is comparatively small after age 35 and could perhaps be 'lost' in the first term by decreasing c_1 slightly. So far as ages under 35 are concerned, we must not lose sight of the fact that the childhood anticipated deaths are more considerable, and also that the early anticipated deaths would have to be represented in any graduation; it is, however, at these lower ages that it is thought that the flexibility of the new curves will be useful, and meanwhile it may be instructive to compare the first two terms of formula (2), or in other words formula (1), with the Perks form.

49. The difficulty in fitting a Makeham curve over a wide range of ages is that there is a tendency for the slope of the mortality curve to slacken off at the higher ages. This has been counteracted by the introduction of the Perks constant D ; but if this peculiarity in the mortality curve is regarded as the result of taking a number of causes of death whose rates of destruction increase regularly throughout life, and superimposing other causes which have a maximum effect in late middle life (or perhaps later still in the case of females), the second term of formula (1) may prove to be a powerful substitute for the constant D .

50. That there is some similarity between formula (1) and Perks is obvious, since both give an adjustment to an exponential function in order to slacken off the slope. In Perks's formula, this slackening off continues throughout life, but in formula (1) the adjustment wears off before the limit of life. Attempts have been made to demonstrate algebraically the similarity between these adjustments given by the two curves by expanding the formulae in terms of ascending powers of e^x and comparing the coefficients. Unfortunately, all that this comparison has shown is that the curves are similar if the adjustments are small in relation to the exponential part of the formula, a conclusion which was so evident from the start that it does not seem worth while recording the algebra.

X. EXTENSION OF THE EXPERIMENTS TO THE 1951 DATA

51. The experiments and conclusions had reached this stage at the time the Registrar General's *Statistical Review* for 1951 became available. It was tempting to hold up the conclusions until the latest data had been investigated, but upon reflexion it was decided to record faithfully what had been done to date. The 1951 data should be superior, in particular in so far as disadvantage (b) of section II does not apply to nearly so great an extent. At the same time, too much should not be expected of the data, bearing in mind the other disadvantages and the fact that the exposed to risk are still only estimates.

52. The complaint is frequently heard that when one is fitting a curve to mortality data the figures never do quite what is hoped of them. In the experiments just described the technique was adopted of not expecting anything in particular from the figures, but of allowing them to show the way, without ever letting them go so far as to take control. This arrangement worked excellently for the 1949 and 1950 data, and under the circumstances

the author proceeded to the investigation of the 1951 data with every confidence that the figures would not let him down provided he was prepared to follow where they led.

53. As the 1951 population estimates were based on the census taken in that year, it seemed reasonable to assume that the data as a whole would be more reliable than the data for 1949 and 1950. It was considered that it would now be worth while applying King's formula, to find the exposed to risk and deaths for the central age-year of each five-year group. On taking out the differences of the male data, it became clear that fourth differences could not be ignored, and therefore a further term in King's formula was retained (see King *J.I.A.* 43, 114—formula ii a). Including this term the formula is

$$u_0 = \frac{1}{5} (w_0 - \frac{1}{25} \Delta^2 w_{-1} + \frac{1}{3125} \Delta^4 w_{-2}).$$

When the adjustments were calculated, it appeared that the term in Δ^4 was sufficiently small, so far as the populations were concerned, for sixth and higher differences to be ignored. This, however, was not so for the deaths, where the term in Δ^4 was sometimes so large that it seemed that the terms in Δ^6 and even Δ^8 would have to be calculated, and in order to avoid such a long process (with no guarantee that the results would justify the labour) it was decided to use Table 13 of the Registrar General's *Statistical Review (Medical)*, which gives the deaths by individual ages but not according to cause of death. (A preliminary cause-group investigation gave indications similar to the 1950 results.) The above modification of King's formula as applied to the figures in Table 1 of the *Review* enabled estimates of the mid-1951 population to be made at every fifth age, which could be used as the exposed to risk corresponding to every fifth number of deaths as obtained from Table 13 of the *Review*. This procedure may be criticized for using only one-fifth of the available data of deaths, but it was considered that it might lead to a suitable curve which would be useful as a first trial if population estimates could be obtained age by age. Details of this curve, and a similar curve found for the female data, are not thought to be worth publishing, but it may be said that they were useful stepping stones to the curves finally fitted to the 1951 data, also that they gave rise to no suspicion that the conclusions drawn from the earlier data would be materially changed.

54. An approach was made to the General Register Office for mid-1951 population estimates age by age, which it was hoped might have been prepared from the census data. The reply was that even if such figures were available they would be too conjectural, and an offer was made instead of a tabulation prepared from the 1% sample data. The author warmly acknowledges this offer, which was accepted. The tabulation gave the age distribution, for each sex separately, of the 1% data, but two features had to be considered in preparation for its use. As the data were populations on the census date (8 April 1951) some adjustment would be necessary to find the mean exposed to risk, age by age, over the year; and as the age distribution was subject to sampling errors, an appropriate allowance for this fact would need to be made when calculating the standard deviation of the expected deaths, particularly at the higher ages where any sampling error in the exposures has a relatively large effect on the expected deaths. Data for all ages over 94 had been amalgamated in the tabulation.

55. If P_x is the 1% population at age x last birthday, θ_x the deaths during the year, and n the interval from the beginning of the year to the census date, it was considered from first principles that, subject to sampling errors, the central exposed to risk would be given by

$$\cdot 01 E_x^c = \frac{1}{2} (1-n)^2 P_{x-1} + (\frac{1}{2} + n - n^2) P_x + \frac{1}{2} n^2 P_{x+1} - \cdot 01 (\frac{1}{2} - n) \theta_x.$$

There are two inaccuracies in the final term, which assumes that deaths are evenly distributed over the calendar year, and which strictly speaking should consist of three separate terms in θ_{x-1} , θ_x and θ_{x+1} , but since the whole term is small compared with $\cdot 01 E_x^c$ (except at the highest ages where, in any case, sampling errors have a very large effect), it was thought that minute refinements would be out of place. Substituting for n , this formula gives

$$\cdot 01 E_x^c = \cdot 267551 P_{x-1} + \cdot 696405 P_x + \cdot 036044 P_{x+1} - \cdot 0023 \theta_x.$$

56. The standard deviation of the expected deaths as derived from a sample exposed to risk is in the nature of the standard deviation of the product of two variables, E_x^c being subject to the sampling errors appropriate to a 1% sample, and m_x assumed to be subject to the usual random errors. There is no reason to suppose that the two variables are correlated. The author was unable to find the expression for the standard deviation of the product of two uncorrelated variables in any statistics text-book, but an expression was obtained from first principles. From this, it was deduced that the variance of the expected deaths at age x last birthday in a population, as found by reference to a 1% sample of the exposed to risk, was approximately equal to

$$\text{expected deaths} \times (1 + 99m_x).$$

This assumes that the sample exposed to risk is large enough for certain terms to be ignored, which would be so in the case of the data in question. In fact, the calculation of column 13 of Table 11 included these terms, but they were clearly negligible and were therefore ignored when compiling column 13 of Table 12.

57. This completed the preliminary work and, as already stated, the rough curve found from the data for every fifth age formed a useful first trial. An examination of the 1951 male data age by age revealed two modifications of the interim conclusions derived from the 1949 and 1950 data. In the first place, the curve for the early anticipated deaths is quite skew, having a considerable effect between the ages of 13 and 37 and having a peak at age 22; the curve actually fitted was of the form $yB_3c_3^{-z^{2/3}}$, where y is measured from birth and z from age 21, but it is not suggested that there is any philosophical or scientific reason why it should take this form, apart from the fact that it does represent a shape with the required degree of skewness, running off asymptotically to zero in both directions when the forces in question cease to have any effect.

58. The second feature shown was that, although a number of different shapes were tried, and several different sets of constants for each shape, all the curves tried were significantly lower than the crude curve from ages 30 to 35 inclusive, or perhaps 29 to 36. This is probably attributable to the effects of the 1939-45 war. Those aged 29-36 in 1951 were aged 17-24 in 1939, and it is not unreasonable to suspect that this group would represent the ages feeling the greatest effects of the war, as regards not only the elimination of

a proportion of the best lives, but also the transfer of another proportion to a subnormal category.

59. The curve fitted to the male data, apart from the first two years of life and the feature described in § 58, was of the form

$$m_y = B_1 c_1^x + B_2 c_2^{-x^2} + A/(y + 1\frac{1}{4}) + y B_3 c_3^{-z^{1/4}}, \quad (3)$$

where

$$1000 \quad B_1 = 41.78$$

$$c_1^5 = 1.656$$

$$1000 \quad B_2 = 6.000$$

$$c_2^{25} = 1.145$$

$$1000 \quad B_3 = .036$$

$$c_3^{-1} = .660$$

$$1000 \quad A = 4.400$$

x is measured from age 67,

y is measured from birth,

z is measured from age 21.

Details are shown in Table 11, which also gives crude values of m_y for comparative purposes. It will be seen that, except at ages 0, 1 and 30–35 inclusive, the fit is satisfactory, apart from the fact that the probability of a value of 116.7 or more for χ^2 (87 values) is only about $2\frac{1}{2}\%$. It should, however, not be overlooked that the deviations are not entirely caused by random errors, since in national data there is always a certain amount of age misstatement, even though the incidence—and the size of individual misstatement—is almost certainly decreasing with time. It is considered that much of the effect of these local errors will be removed if the deviations are examined in quinary groups, and this is accordingly done in the last columns of the table, which demonstrate that if we allow for the effect of local errors the fit is excellent. It should be mentioned that since nothing has been deliberately or consciously minimized in the course of the fit, no reduction has been made in the number of degrees of freedom; the fact that the fit is not necessarily ‘the best’ does not matter because the curve has not been forced to the data. It should also be noted that in this instance the mean deviation test does not give a satisfactory summary of the curve’s adherence to the data; on this test alone, the curve could be accepted for the whole range including ages 30–35, and the reason for this false answer to the question is that the deaths, and deviations, are very small at these ages as compared with the deaths and deviations between, say, ages 48 and 88, with the result that what are really quite significant deviations at the earlier ages fade into insignificance when averaged with the later ages.

60. When the 1951 preliminary female data were investigated for every fifth age, and also in quinary age groups to take account of the cause of death subdivision (which had been made in accordance with the recommendations of § 37), it again appeared that the normal anticipated mortality curve might be either a second Gompertz curve, or a curve of the normal shape. When, however, the full information age by age up to 94 was examined it was clear that a double Gompertz would give rates which were much too high above age 87.

Table 11. A fit of formula (3) to the central death rates amongst the male population of England and Wales in 1951

Age y	1 % central exposed to risk	Actual deaths (whole popn.)	Crude $m_y = (3)/$ $100 \times (2)$	1000 $B_1 c_1^*$	1000A $y + 1 \frac{1}{2}$	1000 $B_1 c_1^* - x^2$	1000 $y B_1 c_1^* - x^2/y$	m_y graduated (5) + (6) + (7) + (8)	Expected deaths (2) x (9) x 100	(3)-(10) Actual minus expected	Σ (11) Accumulated deviations	Variance of number of deaths	χ^2 (11) x (11) (13)	Grouped actual deaths	Grouped expected deaths	Grouped deviations (actual minus expected)	Variance of group deaths	χ^2 (17) x (17) (18)
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)	(19)
0	3343.9	(11,773)	.03521	.04848	3.52000	(.0035685)	(.0020092)	(.10 %)	(10 %)	+	+	598	.14					
1	3475.9	(918)	.00264	.05363	1.05556	(.0020092)	(.0020092)	(.75 %)	(75 %)	-	-	485	.60	1,339	1,344	5	1,497	.02
2	3718.6	535	.00144	.05932	1.35385	.0014132	.0011009	526	526	9	9	414	.02					
3	3977.0	421	.00106	.06561	1.03529	.0011009	.0009107	438	438	17	17	300	.85					
4	4168.2	293	.00092	.07258	.83810	.0009107	.0007843	279	279	3	3	255	.14					
5	3561.0	385	.00083	.08028	.70400	.0007843	.0006957	239	239	16	16	222	.22					
6	3430.6	233	.00068	.08880	.60690	.0006957	.0006316	209	209	6	6	198	.99	1,072	1,081	9	1,151	.07
7	3308.5	216	.00065	.09823	.53333	.0006316	.0005844	187	187	7	7	176	.82					
8	3203.9	173	.00054	.10866	.47568	.0005844	.0005499	167	167	14	14	158	.10					
9	3028.4	155	.00051	.12019	.42927	.0005499	.0005264	150	150	12	12	154	.06					
10	2843.2	154	.00054	.13295	.39111	.0005264	.0005153	147	147	4	4	160	.10					
11	2801.0	144	.00050	.14706	.35918	.0005153	.0005209	152	152	3	3	170	.59					
12	2914.8	156	.00054	.16267	.33208	.0005209	.0005492	161	161	4	4	186	.05	812	786	26	828	.82
13	2927.0	171	.00058	.17994	.30877	.0005492	.0006053	176	176	10	10	207	.04					
14	2907.9	187	.00064	.19903	.28852	.0006053	.0006901	194	194	11	11	235	.15					
15	2811.1	197	.00070	.22016	.27077	.0006901	.0007905	218	218	3	3	271	.355	1,181	1,225	44	1,340	1.44
16	2728.8	224	.00082	.24353	.25507	.0007905	.0009274	248	248	6	6	307	.355					
17	2678.2	217	.00081	.26938	.24410	.0009274	.0010530	278	278	31	31	320	.38					
18	2643.2	245	.00093	.29797	.22857	.0010530	.0011736	287	287	33	33	372	.348					
19	2442.8	298	.00122	.32960	.21728	.0011736	.0012769	331	331	11	11	424	.94					
20	2592.4	367	.00142	.36459	.20706	.0012769	.0013571	374	374	36	36	447	2.74	1,931	1,945	14	2,210	.09
21	2752.8	354	.00129	.40329	.19775	.0013571	.0014453	412	412	20	20	497	1.16					
22	2782.1	428	.00154	.44609	.18925	.0014453	.0014597	435	435	41	41	520	.09					
23	2850.8	371	.00130	.49344	.18144	.0014597	.0014617	455	455	24	24	507	.348					
24	2982.8	411	.00138	.54582	.17426	.0014617	.0014577	444	444	39	39	522	.00					
25	3113.2	462	.00159	.60376	.16762	.0014577	.0014541	457	457	1	1	543	.74	2,396	2,339	57	2,673	1.22
26	3048.5	486	.00145	.66784	.16147	.0014541	.0014503	475	475	20	20	581	1.45					
27	3143.5	456	.00145	.73873	.15575	.0014503	.0014690	508	508	18	18	629	.322					
28	3200.1	455	.00140	.81714	.15043	.0014690	.0014960	549	549	11	11	612	4.76					
29	3459.8	537	.00155	.90388	.14545	.0014960	.0015399	532	532	45	45	513	2.53					
30	3672.7	594	.00162	.99982	.14080	.0015399	.0016020	443	443	54	54	494	1.82					
31	3452.7	586	.00170	1.10595	.13643	.0016020	.0016869	424	424	56	56	591	3.58					
32	2704.4	479	.00173	1.22334	.13233	.0016869	.0017933	503	503	110	110	697	5.34					
33	2512.4	454	.00181	1.35319	.12847	.0017933	.0019235	587	587	146	146	805	.78					
34	2804.3	549	.00196	1.49682	.12482	.0019235	.0020701	669	669	176	176	900	.16					
35	3052.2	648	.00212	1.65570	.12138	.0020701	.0022619	737	737	222	222	1,013	.14					
36	3218.4	694	.00216	1.83145	.11812	.0022619	.0024741	816	816	283	283	1,160	5.38					
37	3256.2	749	.00230	2.02585	.11503	.0024741	.0027184	917	917	308	308	1,300	.65					
38	3298.1	828	.00251	2.24088	.11210	.0027184	.0029981	1,006	1,006	332	332	1,453	1.16					
39	3374.9	838	.00248	2.47874	.10932	.0029981	.0033168	1,098	1,098	353	353	1,740	.10					
40	3354.9	977	.00291	2.74184	.10667	.0033168	.0036790	1,281	1,281	253	253	1,938	6.36					
41	3399.8	1,057	.00319	3.03288	.10414	.0036790	.0040508	1,386	1,386	224	224	2,163	9.72					
42	3482.6	1,204	.00372	3.35480	.10173	.0040508	.0045546	1,498	1,498	183	183	2,436	.53					
43	3388.9	1,275	.00376	3.71090	.09944	.0045546	.0050798	1,620	1,620	196	196	2,765	1.30					
44	3289.4	1,353	.00411	4.10479	.09724	.0050798	.0056720	1,780	1,780	85	85	3,142	1.07	5,956	6,269	313	8,594	11.40
45	3206.5	1,665	.00519	4.54049	.09514	.0056720	.0063385	1,941	1,941	145	145	3,671	2.62					
46	3139.1	1,720	.00548	5.02244	.09312	.0063385	.0070869	2,170	2,170	60	60	4,188	8.53					
47	3002.7	1,883	.00615	5.55555	.09110	.0070869	.0079250	2,361	2,361	84	84	4,992	.11					
48	3002.6	2,072	.00677	6.14524	.08934	.0079250	.0088607	2,677	2,677	58	58	5,417	1.92					
49	2979.6	2,550	.00856	6.79753	.08756	.0088607	.0099022	2,753	2,753	98	98	6,044	.00					
50	3020.7	2,700	.00894	7.51905	.08585	.0099022	.011057	2,904	2,904	240	240	6,591	5.25					
51	2779.8	2,855	.01027	8.31717	.08421	.011057	.012334	2,986	2,986	51	51	7,824	.43					
52	2626.0	2,903	.01105	9.19999	.08263	.012334	.013739	3,337	3,337	74	74	8,839	.77	14,909	14,957	252	30,868	2.06
53	2421.0	3,172	.01310	10.1765	.0811	.013739	.015281	3,541	3,541	259	259	9,978	.53					
54	2428.6	3,279	.01350	11.2507	.0796	.015281	.016966	3,748	3,748	201	201	11,120	1.64					
55	2317.2	3,279	.01415	12.4516	.0782	.016966	.018802	3,911	3,911	262	262	13,532	.31					
56	2209.2	3,675	.01663	13.7732	.0769	.018802	.020796	4,454	4,454	73	73	16,278	4.45					
57	2079.9	4,046	.01945	15.2352	.0755	.020796	.022957	5,010	5,010	66	66	18,094	2.06					
58	2141.7	4,519	.02110	16.8523	.0743	.022957	.025293	5,200	5,200	134	134	19,919	.08					
59	2182.3	4,741	.02172	18.6411	.0730	.025293	.027816	5,342	5,342	203	203	23,385	.43					
60	2056.1	5,007	.02435	20.6108	.0718	.027816	.030539	5,552	5,552	260	260	26,317	1.18					
61	1920.3	5,302	.02761	22.8085	.0707	.030539	.033478	5,841	5,841	193	193	29,741	.09					
62	1916.3	5,952	.03106	25.2295	.0696	.033478	.036652	6,468	6,468	40	40	33,676	1.03					
63	1834.3	6,191	.03375	27.9074	.0685	.036652	.040084	6,822	6,822	100	100	37,675	.22					
64	1764.7	6,555	.03771	30.8697	.0674	.040084	.043804	7,104	7,104	50	50	41,840	.04					
65	1702.0	7,008	.04118	34.1463	.0664	.043804	.047845	7,339	7,339	187	187	47,707	.02					
66	1621.7	6,867	.04234	37.7708	.0654	.047845	.052246	7,775	7,775	87	87	54,201	.02					
67	1534.0	7,434	.04846	41.7800	.0645	.052246	.057054	8,200	8,200	150	150	59,740	.72					
68	1488.2	7,816	.05252	46.2147	.0635	.057054	.062323	8,379	8,379	55	55	64,681	1.27					
69	1437.3	8,233	.05728	51.1202	.0626	.062323	.068111	8,396	8,396	14	14	68,795	.43					
70	1344.5	8,171	.06077	56.5464	.0618	.068111	.074488	8,253	8,253	19	19	79,188	.06					
71	1232.7	8,109	.06578	62.5485	.0609	.074488	.081528	8,769	8,769	476	476	90,592	.23					
72	1108.0	8,802	.07944	69.1877	.0601	.081528	.089315	9,245	9,245	73	73	99,926	.65					
73	1075.6	8,840	.08219	76.5316	.0593	.089315	.097941	9,381	9,381	144	144	103,29						

Table 12. A fit of formula (4) to the central death rates amongst the female population of England and Wales in 1951

Age y	1 % central exposed to risk	Actual deaths (whole popn.)	Crude m_y = (3)/(100 × (2))	1000 $B_1 c_1^*$	1000 A $y + 14$	1000 $B_2 c_2^{-2}$	1000 $B_3 c_3^{-3}$	m_y graduated (5) + (6) + (7) + (8)	Expected deaths (2) × (9) × 100	(3)-(10) Actual minus expected	Σ (11) Accumulated deviations	Variance of number of deaths	$\frac{x^2}{(11) \times (11)}$ (13)	Grouped deviations (actual minus expected)	Variance of group deaths	$\frac{x^2}{(15) \times (15)}$ (16)
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)
0	3265.1	(8,450)	.025880	.02062	2.40000			.0024296	(9.1 %)							
1	3323.1	(774)	.002329	.03281	1.33333			.0013661	(59 %)							
2	3561.1	(450)	.03634	.001264	.92308			.0009594	(76 %)							
3	3857.9	(353)	.000915	.04024	.70588			.0007401	(82 %)							
4	3990.1	(299)	.000749	.04457	.57143			.0006160	(82 %)							
5	3434.8	187	.000544	.04937	.48000			.0005294	182	5	5	192	.13			
6	3285.7	157	.000478	.05468	.41379	.00000		.0004685	154	3	8	161	.06			
7	3278.5	131	.000400	.06056	.36364	.00004		.0004242	139	8	0	145	.44			
8	3077.4	126	.000409	.06707	.32432	.00029		.0003917	121	5	5	126	.20			
9	2765.0	98	.000354	.07429	.29268	.00139		.0003684	102	4	1	106	.15			
10	2642.3	102	.000386	.08228	.26667	.00497		.0003539	94	8	9	97	.66			
11	2773.7	89	.000321	.09113	.24490	.01408		.0003501	97	8	1	100	.64			
12	2906.8	97	.000334	.10093	.22642	.03267		.0003800	105	8	7	109	.59			
13	2882.1	119	.000413	.11179	.21053	.06410		.0003864	111	8	1	115	.56			
14	2779.8	109	.000392	.12382	.19672	.10937		.0004299	120	11	10	125	.97			
15	2791.1	132	.000473	.13713	.18462	.16625		.0004880	136	4	14	143	.11			
16	2754.0	164	.000595	.15188	.17391	.22987		.0005557	153	11	3	161	.75			
17	2688.7	154	.000573	.16822	.16438	.29438		.0006270	169	15	18	179	1.26			
18	2731.2	195	.000714	.18632	.15584	.35453		.0006967	190	5	13	203	.12			
19	2819.0	235	.000834	.20636	.14815	.40673		.0007612	215	20	7	231	1.73			
20	2940.7	211	.000715	.22856	.14118	.44929		.0008190	242	31	24	262	3.67			
21	2999.4	280	.000934	.25314	.13483	.48212		.0008701	261	19	5	281	1.28			
22	3005.8	270	.000898	.28037	.12903	.50619		.0009156	275	5	10	300	.08			
23	2977.6	278	.000934	.31053	.12371	.52300		.0009572	285	7	17	312	.16			
24	3032.4	293	.000966	.34393	.11881	.53419		.0009969	302	9	26	332	.24			
25	3102.1	323	.001041	.38093	.11429	.54127		.0010365	322	1	25	355	.00			
26	3160.1	364	.001152	.42190	.11009	.54551		.0010775	341	23	2	377	1.40			
27	3196.4	368	.001151	.46728	.10619	.54787		.0011213	358	10	8	398	.25			
28	3322.8	403	.001213	.51755	.10256	.54891		.0011690	388	15	23	433	.52			
29	3519.5	446	.001267	.57322	.09917	.54967		.0012221	430	16	39	482	.53			
30	3775.8	458	.001213	.63488	.09600	.54990		.0012808	484	26	13	545	1.24			
31	3593.6	488	.001358	.70317	.09302	.54998		.0013462	484	4	17	549	.03			
32	2876.9	446	.001550	.77881	.09023	.55000		.0014190	408	38	55	465	3.11			
33	2607.7	431	.001653	.86258	.08759	.55000		.0015002	391	40	95	449	3.56			
34	2896.7	471	.001626	.95537	.08511	.55000		.0015905	461	10	105	534	.19			
35	3140.7	526	.001675	1.05813	.08276	.55000		.0016909	531	5	100	620	.04			
36	3399.0	618	.001818	1.17195	.08054	.55000		.0018025	613	5	105	722	.03			
37	3459.7	675	.001951	1.29801	.07843	.55000		.0019264	660	9	114	793	.10			
38	3463.6	697	.002012	1.43763	.07643	.54998		.0020604	715	18	96	861	.38			
39	3443.1	721	.002094	1.59228	.07453	.54990		.0022167	763	42	54	930	1.90			
40	3478.3	806	.002317	1.76355	.07273	.54967		.0023860	830	24	30	1,026	.56			
41	3368.5	812	.002411	1.95325	.07101	.54891		.0025732	867	55	25	1,088	2.78			
42	3505.9	929	.002650	2.16335	.06936	.54787		.0027806	975	46	71	1,243	.70			
43	3436.0	1,008	.002934	2.39606	.06780	.54551		.0030094	1,034	26	97	1,342	.50			
44	3330.6	1,046	.003141	2.65379	.06630	.54127		.0032614	1,086	40	137	1,437	1.11			
45	3369.7	1,163	.003451	2.93925	.06486	.53419		.0035383	1,192	29	166	1,610	.52			
46	3276.6	1,291	.003940	3.25542	.06340	.52300		.0038419	1,259	32	134	1,738	.59			
47	3256.1	1,353	.004155	3.60559	.06218	.50619		.0041740	1,359	6	140	1,921	.02			
48	3170.4	1,544	.004870	3.99343	.06091	.48212		.0045365	1,438	106	34	2,084	5.39			
49	3121.3	1,565	.005014	4.42299	.05970	.44929		.0049320	1,539	26	8	2,290	.30			
50	3250.3	1,738	.005347	4.89875	.05854	.40673		.0053640	1,743	5	13	2,669	.01			
51	3171.8	1,880	.005927	5.42560	.05742	.35453		.0058376	1,852	28	15	2,922	.27			
52	3006.9	1,880	.006252	6.00932	.05634	.29438		.0063600	1,912	32	17	3,116	.33			
53	2899.1	2,054	.007085	6.65572	.05530	.22987		.0069409	2,012	42	25	3,395	.52			
54	2818.2	2,177	.007725	7.37165	.05430	.16625		.0075922	2,140	37	62	3,748	.37			
55	2706.0	2,271	.008392	8.16459	.05333	.10937		.0083273	2,253	18	80	4,110	.08			
56	2694.1	2,574	.009554	9.04282	.05240	.06410		.0091593	2,468	106	186	4,706	2.39			
57	2672.5	2,631	.009845	10.0155	.0515	.0327		.010100	2,699	68	118	5,398	.86			
58	2712.8	2,968	.010941	11.0929	.0506	.0141	.0000	.011158	3,027	59	59	6,371	.55			
59	2593.0	3,104	.011971	12.2861	.0498	.0050	.0001	.012341	3,200	96	37	7,110	1.30			
60	2571.2	3,303	.012846	13.6076	.0490	.0014	.0003	.013658	3,512	209	246	8,261	5.29			
61	2422.3	3,464	.014300	15.0714	.0482	.0007	.0007	.015121	3,663	199	445	9,146	4.33			
62	2393.4	3,978	.016021	16.6025	.0474	.0000	.0017	.016742	4,007	29	474	10,648	.08			
63	2357.0	4,274	.018133	18.4881	.0467	.0041	.0041	.018539	4,370	96	570	12,391	.74			
64	2285.8	4,713	.020619	20.4768	.0460	.0092	.0092	.020532	4,693	20	550	14,232	.03			
65	2244.1	5,069	.022588	22.6794	.0453	.0199	.0199	.022745	5,104	35	585	16,507	.07			
66	2144.4	5,237	.024422	25.1190	.0446	.0413	.0413	.025205	5,405	168	753	18,802	1.49			
67	2049.5	5,747	.028041	27.8209	.0440	.0825	.0825	.027947	5,728	19	734	21,576	.73			
68	2023.2	6,065	.029977	30.8135	.0433	.1584	.1584	.031015	6,275	210	944	25,542	1.73			
69	1962.5	7,217	.036775	34.1280	.0427	.2924	.2924	.034463	6,763	454	490	29,837	6.91			
70	1863.5	7,075	.037966	37.7990	.0421	.5187	.5187	.038360	7,148	73	563	34,294	.16			
71	1744.5	7,222	.041399	41.8649	.0415	.8846	.8846	.042791	7,465	243	806	39,089	1.51			
72	1696.7	8,158	.048082	46.3682	.0410	1.4502	1.4502	.047859	8,120	38	768	46,593	.03			
73	1635.2	8,607	.052636	51.3558	.0404	2.2853	2.2853	.053682	8,778	171	939	55,429	.53			
74	1511.5	9,274	.061356	56.8800	.0399	3.4615	3.4615	.060381	9,127	147	792	63,686	.34			
75	1379.6	9,404	.068165	62.9984	.0393	5.0397	5.0397	.068077	9,392	12	780	72,691	.00			
76	1240.5	9,569	.077138	69.7749	.0388	7.0532	7.0532	.076807	9,535	34	746	82,095	.01			
77	1051.8	9,286	.088287	77.2803	.0383	9.4883	9.4883	.086807	9,120	156	590	87,592	.28			
78	930.3	9,409	.101139	85.9331	.0379	12.2692	12.2692	.097900	9,108	301	289	97,384	.93			
79	841.2	9,171	.109023	94.8000	.0374	15.2500	15.2500	.11009	9,261	90	379	110,196	.07			
80	767.3	8,962	.116799	104.997	.037	18.220	18.220	.12325	9,457	495	874	124,849	1.96			
81	646.3	8,397	.129024	116.291	.036	20.925	20.925	.13725	8,870	473	1,347	129,393	1.73			
82	537.9	8,021	.149117	128.801	.036	23.999	23.999	.15194	8,173	1						

61. After a trial fit of a Gompertz, plus a normal, plus a harmonic had been made to the 1% data, the exposed to risk having been found in the same way as for males, it was necessary to determine the shape of the early anticipated deaths, and here the figures led into new ground. This component of the female curve appeared to reach effectively from about age 13 to 56; it also appeared that a symmetrical curve with its origin at $34\frac{1}{2}$ could be fitted more easily than any skew curve, but that it was practically flat-topped with a steep drop at either side, a bowler hat rather than a cocked hat. It was found that this could be represented by the curve $B_3c_3^{-x^2}$. It is not quite clear why the curve which, at the younger ages, is caused largely by tuberculosis and partly by unnatural deaths, should persist into middle life. The effects of childbirth and pregnancy deaths are not apparently of a high enough order to cause the phenomenon, but there may well be a large number of pregnancy deaths which are attributed primarily to other causes; this possibility is examined further in §65.

62. The curve fitted to the female data, excluding the first five years of life, was of the form

$$m_x = B_1c_1^x + B_2c_2^{-x^2} + A/(y + 1\frac{1}{4}) + B_3c_3^{-x^2}, \quad (4)$$

where

$$1000 \quad B_1 = 158.0$$

$$c_1^5 = 5/3$$

$$1000 \quad B_2 = 25.00$$

$$c_2^{-25} = .610$$

$$1000 \quad B_3 = .550$$

$$c_3^{10^4} = 1.022$$

$$1000 \quad A = 3.000$$

x is measured from age 84,

y is measured from birth,

z is measured from age $34\frac{1}{2}$.

Details are shown in Table 12, and the fit seems to be quite acceptable.

XI. COMPARISON BETWEEN MALE AND FEMALE CURVES

63. The fundamental differences between the male and female mortality curves appear to be threefold. First, the origin of the normal anticipated death curve is found much later in life in the female curve, with the result that an examination of the ratio m_{x+5}/m_x up to age $x=80$, using either graduated or ungraduated rates, would show that whereas the ratio tends to decrease over quite a large range of ages for the male curve (suggesting a curve of the Perks shape), it continues to increase well beyond the allotted span of life for females. The main causes of death applicable to this feature for males have already been seen to be respiratory tuberculosis and respiratory neoplasms. For females, the only important group having a deaths distribution in any way similar to the normal anticipated curve is diseases of the nervous system and sense organs (I.C. numbers 330-398). It is not considered worth while showing figures, especially because all ages over 84 would have to be grouped together, but it may be significant that this is also the one group

The Components of Mortality

where female deaths at all the high ages are appreciably more numerous than male deaths.

64. Secondly, the balance at the infantile ages unfitted by the curves, and presumed to represent natal deaths, shows either that the harmonic curve does not represent childhood anticipated deaths so well for females as it does for males, or that the effect of birth on male mortality is both more severe and more immediate than its effect on female mortality; this is demonstrated in Table 13, although it is, of course, realized that the natal

Table 13. Hypothetical life tables showing the survivors from 100,000 births if there were no deaths other than natal deaths in the early years of life

Age y	m_y^N crude central natal mortality rate	q_y^{N*} initial natal mortality rate	l_y	$d_y^N = l_y \times m_y^N$
Males				
0	·03135	·03086	100,000	3,086
1	·00063	·00063	96,914	61
2	—	—	96,853	—
Females				
0	·02345	·02318	100,000	2,318
1	·00096	·00096	97,682	94
2	·00030	·00030	97,588	29
3	·00017	·00017	97,559	17
4	·00013	·00013	97,542	13
5	—	—	97,529	—

* Given by the formula $\frac{1}{q_y^N} = \frac{1}{m_y^N} + \frac{1}{2} + \frac{m_y^N}{12}$.

mortality rates are derived from crude data. The fact that a decreasing Gompertz gave a worse fit at the childhood ages, and that the alteration of the scale to the square root of the age would still not give a fit which reconciles the data of the first five years of life with the other childhood ages, led to the belief that the postponement to, at the latest, the fifth year of life of certain deaths which, in the case of males, would have occurred at the latest in the second year of life was the most likely cause of this phenomenon in the female statistics. The data were not sufficiently extensive for a formula to be found for this component of the mortality rate, but in any event such a curve would be discontinuous.

65. Thirdly, the early anticipated deaths are of quite different shapes for the two sexes; Table 14 shows that the unnatural causes account for a very large proportion of this group of male expected deaths, while the remainder can be more than accounted for by tuberculosis deaths at the younger ages; Table 15 shows that the early anticipated female expected deaths are, from ages 15 to 49, well in excess of the total of deaths through tuberculosis or unnatural causes, but that if deaths at these ages due to pregnancy, child-birth and diseases of the genito-urinary system are included, the total of the group is again more than accounted for. The balance of deaths from these causes will, of course, fit into the other components of the mortality curve.

66. Subject to these basic differences, together with the effects of what might be called 'temporary war selection', both curves appear to be striving towards the form of formula (2).

Table 14. A comparison of the male deaths from certain causes in England and Wales in 1951 with the expected early anticipated deaths in certain age groups based on the rates shown in column (8) of Table 11

(The figures in square brackets show the International Classification)

Age group	Expected early anticipated deaths	Deaths from unnatural causes [E800-E999]	Deaths due to respiratory tuberculosis [001-008]	Deaths due to other forms of tuberculosis [010-019]
(1)	(2)	(3)	(4)	(5)
10-14	64	270	10	39
15-19	541	483	61	64
20-24	1,048	862	242	41
25-29	885	830	448	54
30-34	389	663	472	57
35-39	127	644	529	57

Table 15. A comparison of the female deaths from certain causes in England and Wales in 1951 with the expected early anticipated deaths in certain age groups based on the rates shown in column (7) of Table 12

(The figures in square brackets show the International Classification)

Age group	Expected early anticipated deaths	Deaths due to unnatural causes [E800-E999]	Deaths due to tuberculosis (all forms) [001-019]	Deaths due to pregnancy, child-birth, etc. [640-689]	Deaths due to diseases of the genito-urinary system [590-637]	Sum of columns (3), (4), (5) and (6)
(1)	(2)	(3)	(4)	(5)	(6)	(7)
10-14	62	62	75	—	36	173
15-19	400	121	230	15	49	415
20-24	748	130	458	86	58	732
25-29	890	139	622	135	62	958
30-34	866	136	575	110	105	926
35-39	929	197	475	142	138	952
40-44	935	228	390	50	201	869
45-49	809	339	371	19	234	963
50-54	447	375	327	4	307	1,013
55-59	61	397	270	1	341	1,009

XII. THE COMPONENTS OF THE RATE OF MORTALITY

67. The interim conclusion of § 35 may now be modified to the following final statement of the apparent components of the rate of mortality:

Normal components

- (i) The rate of senescent deaths, an exponential curve.
- (ii) The rate of normal anticipated deaths, in the shape of a normal curve.
- (iii) The rate of childhood anticipated deaths, apparently in the shape of a harmonic curve.

Components due to abnormalities or disturbances

(iv) The rate of natal deaths, starting at a high level at birth, and decreasing rapidly, with no remaining effect after the first few years of life; a discontinuous curve.

(v) The rate of early anticipated deaths, of skew cocked-hat shape for males and symmetrical bowler-hat shape for females, due to a number of causes, largely unnatural.

(vi) The additional rate of mortality due to temporary disturbances, the only evidence of any such disturbances in the 1951 data being some remaining 'war selection' for males.

68. The conclusions of § 37 remain unchanged after considering the 1951 data, apart from the fact that a female investigation might well include a further subdivision to show separately deaths due to pregnancy, childbirth, etc. If the Institute will agree to a continuous cause of death investigation along the suggested lines, an important object of this paper will have been achieved.

XIII. EPILOGUE

69. References have been made to Clarke(3), Pearson(5) and Phillips(7), but I have not overlooked that all these were primarily concerned with the curve of deaths rather than the rate of mortality. I have no desire to express a firm opinion as to whether μ_x or ϕ_x is the basic function, any more than I would hazard a guess whether the chicken or the egg came first. What does seem clear from Phillips's paper is that ϕ_x is a somewhat troublesome function to use in practice, and even if it is proved to be the basic function we shall not be able to dispense with q_x , m_x or μ_x . Because of the complex relationship between μ_x and ϕ_x no way has been discovered of translating any of the formulae given in the present paper into ϕ form, although such translations may well form the subject of future research by young actuaries recently qualified; indeed, if this paper provokes further research it may be said that another object will have been achieved. In this connexion, it may be recalled that when Phillips replied to the discussion on his paper he expressed the hope that young actuaries undertaking research on similar subjects would operate on ϕ_x rather than μ_x ; my hope is not that they will operate on ϕ_x , nor that they will operate on μ_x , nor that they will prejudge any issue in any way, but that they will embark on their research with an open mind, prepared to follow where their data lead. It is for us to give them the best data we can, so that even if they cannot attain the complete truth they may at least be able to approach it asymptotically.

REFERENCES

- (1) BEARD, R. E. Some notes on graduation. *J.I.A.* 77, 382.
- (2) BOWERMAN, W. G. The dip in mortality in the twenties of age. *T.S.A.* 5, 117.
- (3) CLARKE, R. D. A bio-actuarial approach to forecasting rates of mortality. *Proceedings of the Centenary Assembly*, 2, 12.
- (4) OGBORN, M. E. On the nature of the function expressive of the law of human mortality. *J.I.A.* 79, 170.
- (5) PEARSON, K. The chances of death.
- (6) PERKS, W. On some experiments in the graduation of mortality statistics. *J.I.A.* 63, 12.
- (7) PHILLIPS, W. A basic curve of deaths. *J.I.A.* 80, 289.
- (8) POLLARD, A. H. Methods of forecasting mortality using Australian data. *J.I.A.* 75, 151.
- (9) THE REGISTRAR GENERAL'S *Statistical Review of England and Wales, Part I, Medical*, for the years 1949, 1950 and 1951.
- (10) STARKE, L. G. K. Time-changes in the mortality rate: an experimental formula. *J.I.A.* 78, 171.
- (11) THIELE, T. N. On a mathematical formula to express the rate of mortality throughout the whole of life, tested by a series of observations made use of by the Danish Life Insurance Company of 1871. *J.I.A.* 16, 313.

ABSTRACT OF THE DISCUSSION

Mr A. Duval, in opening the discussion, said that no one could complain that the author had not provided enough to talk about. He had devised a new graduation formula—or perhaps a modification of a very old one—and had put forward a philosophical basis for the main terms in that formula. He had suggested that, in a mortality investigation, analysis by cause of death might give useful results, particularly for forecasting and for investigating the effects of selection. He had also made specific proposals for the grouping by causes of death of the data collected by the Joint Mortality Investigation Committee.

He did not propose to discuss the effectiveness of the new graduation formula beyond saying that it appeared to be a powerful one, in that it gave what seemed to be adequate fits on four different population mortality experiences, which were notoriously hard to fit. They were quite different experiences since, apart from the differences between male and female mortality, the 1951 experience included an influenza epidemic which made it quite a different shape from the 1950 experience. The main part of the curve took four parameters, with two more for the early adult deaths and another for the childhood deaths, which seemed to bear out Perks's remark* that with modern mortality, whatever type of curve was used, four parameters were required to fit the main adult mortality and if there was a hump in the 20's two more would be needed.

But the author made it clear that his primary purpose in writing the paper was not to put forward a new graduation formula for practical use but rather to break down the mortality rate into its component parts, the main components having a philosophical background. The author did not say what he meant by a philosophical background, but presumably he was thinking of some kind of reason or explanation for the particular component's varying with age in the way that it did; he was looking for something in the nature of human life itself or in its environment that would make that type of variation probable in all, or virtually all, mortality experiences.

He regarded both objects as highly desirable and was glad that the author had attempted them. Apart from the fact that any greater knowledge of mortality was worth having, if it were possible to break down the mortality rate there would be obvious practical advantages in making estimates for the future—and in their practical work all actuaries were concerned with making estimates of future mortality, whether they made a formal projection of the rates or not—and if they had some idea why the mortality rate had the shape that it had, that again would give them greater confidence in making those estimates.

Frankly, however, he was doubtful whether the author had gone very far towards achieving his objects. He had started with the happy idea of separating out the deaths attributable to causes for which the mortality rate did not increase throughout life, and found that by grouping two of them together, respiratory cancer and tuberculosis of the lung, a 'cocked hat' shape was obtained broadly similar to the normal curve. But then his process was to take the experience from the remainder of the causes of death and try to fit a Gompertz curve to it. He found that he could fit a Gompertz curve if he left out quite a few of the deaths, which he then transferred to the other class. It was important to understand that the other class (the normal anticipated deaths) included far

* *J.I.A.* 79, 199.

more than respiratory cancer and respiratory tuberculosis deaths. The author pointed that out, but the actual figures were surprising. The fit of the 1950 male mortality was described in § 22-25. The normal term in the formula gave a rate of 4.0 per thousand at age 62, compared with 3.3 from those causes, and in the 70-75 age-group 5.1, compared with 2.9 from those causes. In other words, the exponential curve excluded many deaths other than those specifically excluded in obtaining the rate.

So far so good, but the author went on to suggest that the exponential term might be regarded as covering the senescent deaths as defined in Clarke's centenary paper, and the normal term might be regarded as covering the anticipated deaths, or some of the anticipated deaths. Clarke defined senescent deaths as deaths occurring because the natural life-span had run out—mainly deaths from degenerative diseases—and he thought, from an examination of the causes of death, that those senescent deaths might range from 5 % at the younger ages up to the whole of the deaths at ages 80-85 or so. The exponential term in the author's formula gave something like two-thirds the total deaths at age 30 for the males and one-half the total deaths at age 30 for the females. Certainly nothing like that number of deaths occurred from degenerative diseases at those ages, and it was difficult to imagine any definition of senescent deaths that would include such a high proportion of the total deaths at the young adult ages.

The term for the 'normal anticipated deaths', as the author called them—the deaths covered by the normal term in his formula—gave for the female experience for 1950 a death-rate of 11 per thousand at age 72, rising to 50 per thousand at age 82 and to 82 per thousand at age 92, after which it started to fall away. For 1951 the author used the same formula and obtained figures of roughly $1\frac{1}{2}$ at age 72, 23 at age 82 (which was almost the peak), and 7 at age 92. Those rates were said to represent the normal anticipated deaths, the deaths occurring before the expiry of the normal life-span from some extraneous cause. In 1951 there had been an influenza epidemic which made the total deaths more than in 1950 at those ages, but it was difficult to see how any definition of senescent and anticipated deaths could justify the anticipated deaths going down from 82 per thousand at age 92 in 1950 to 7 per thousand at age 92 in 1951, or senescent deaths going up by more than the amount by which anticipated deaths went down. The actual effect of an epidemic was a little difficult to describe. Presumably it caused some deaths that would otherwise have been senescent deaths to occur sooner and brought about an increase in anticipated deaths, but on no conceivable explanation did it seem possible to get figures anything like those that the author had produced.

The fact threw considerable doubt on the author's remarks in section VIII of the paper, and in particular § 40, where it was suggested that 'a formula fitting the data naturally is likely to be useful for purposes of extrapolation'. The author went on to say that 'Projecting into the future by extrapolating on constants is an appealing proposition', but he did not agree; it seemed to him to depend on far too many assumptions, most of which were probably not true. The author continued '... but its success is likely to be limited if the formula used has to be forced to the data since the constants may be to a certain extent interdependent'. That was a view with which the speaker agreed.

The author suggested that the constants in his formula were not in any way interdependent, but surely his very method of fitting a Gompertz curve first and then adjusting it to get the normal curve to fit was bound to make his constants to some extent interdependent? The figures for the normal curve terms on the

female rates for 1950 and 1951 bore that out. How was it possible to extrapolate from those normal curve terms which he had just mentioned? Nobody would attempt to extrapolate from only two years' experience, particularly when one year included an influenza epidemic and one did not, but the same problem would be present to a lesser degree in any experience.

The philosophical background to the exponential curve was Gompertz's and was well known, but the philosophical background to the normal curve was given, presumably, in § 9 of the paper. Although he had read that paragraph many times, he still could not see how it provided any kind of philosophical basis for believing that the normal anticipated deaths would follow a normal curve. Obviously, if those deaths occurred from impairment, and a person became impaired before he died, then, if the rate of becoming impaired followed a normal curve, and death followed shortly afterwards, the death-rate would follow a normal curve; but was there any philosophical basis for believing that the rate of becoming impaired was a normal curve?

It was interesting to compare that philosophical basis with that used by Prawitz in the Scandinavian investigation, where he had suggested that for certain causes of death, including those mentioned by the author but also certain others, the deaths from a given cause might occur entirely amongst those who had some predisposition from birth to that cause of death; and that on the assumption that those deaths did occur only amongst that group of the population and that the death-rate from other causes was the same for the predisposed group as for the other group, it was possible to get a skew cocked-hat curve, if the mortality amongst the predisposed class from that particular cause of death followed broadly the usual shape of the mortality curve for all causes of death. The rate came down at the older ages because by that time the predisposed class had been substantially reduced by the effect of death from that particular cause and then represented only a very small proportion of the total population.

He found it difficult to see why any kind of philosophical background should be expected for the anticipated deaths, particularly when they included deaths from tuberculosis. The generation mortality curve for tuberculosis was a curve of entirely different shape from the period mortality curve, and a mortality curve that assumed that people had always lived in the current environment would again be a curve of entirely different shape. It was surely only the third of those curves that could be expected to have any kind of philosophical basis or any kind of basis resting on the nature of the disease or the nature of the environment or the nature of human life.

He thought that the whole idea of senescent deaths contained a flaw in principle. Senescence was the wearing out of the human body, the body gradually becoming less and less able to resist the onslaughts made on it. It had been defined by one eminent biologist as 'that change of the bodily faculties and sensibilities and energies which accompanies ageing and which renders the individual progressively more likely to die from accidental causes of random incidence'.* As such, it was not very different from Gompertz's inability to withstand destruction. It could not be measured directly, but only through some attribute—there was no absolute scale of measurement.

It could be measured by certain physical attributes—there was a well-known saying that 'A man is as old as his arteries'—or by some attribute of a group of

* P. B. Medawar, *An Unsolved Problem of Biology*. An inaugural lecture delivered at University College, London, 6 December 1951.

lives—the total death-rate had been used for that purpose—but the trouble was that all those measures were to some extent faulty, because they all measured something else as well. Nobody died from senescence alone; there had to be something that pushed him over the edge, however weak that person had become and however senile. In the current state of medical knowledge, however, that was perhaps not of great practical importance. When a person was so senile that he could not carry on the ordinary affairs of life and could not eat and drink without dying it might as well be said that he had died from senescence; but it made difficult the allocation of epidemic deaths between anticipated deaths and senescent deaths, and also the treatment of deaths from such causes as cancer and peptic ulcer. There were at least two kinds of senescence: the natural wearing out of the body and the unnatural wearing out caused by the strains to which it was subjected.

Where, then, had they got to? It had been realized that the problem of mortality was a problem in biology and not merely an exercise in statistics. In essence, it was the problem of finding out why people died when they did. Investigation by cause of death might help in the solution of the problem, but he doubted whether it would do so *at that time*, because of mixed and changing populations and of rapidly changing environments. It might be that they would get more help towards solving the problem by an investigation of animal mortality in more stable and more homogeneous environments, but then there would be the difficulty of trying to translate an animal mortality rate into something applicable to human mortality. If any investigation by cause of death was going to be made, he would prefer something on the lines suggested by Clarke in his centenary paper to that suggested by the author. Even if it did nothing towards solving the ultimate problem of mortality it should at least be of some use for short-term forecasting.

It had to be concluded regretfully that, although the two main terms in the author's formula gave a reasonable fit to mortality from age 35 or 40 upwards, there was no evidence in the paper to support the author's view that the Gompertz term could be regarded as covering senescent deaths and the normal term as covering the normal anticipated deaths, on any reasonable definition of 'normal' and 'senescent' deaths. Such evidence as there was in the paper seemed to point against that.

The author had adopted the modern fashion of heading his paper with a quotation—two quotations, in fact. The second was taken from Perks's classic paper. Had the author noticed that the two sentences which followed that quoted were also relevant to the subject under discussion? They read:

In the case of mortality, we have a little knowledge of the forces operating, and there is the temptation, to which it is dangerous to yield, to allow this knowledge to influence our statistical method. After all, it is a question whether scientific words of doubtful precision written under the heading 'primary cause of death' do more than scratch the surface of the mystery of death.

He did not know whether Perks still held those views which he had expressed nearly 25 years earlier, particularly the remark about the danger of allowing their knowledge of the forces operating to influence their statistical method, but perhaps he would agree that the first step in a new field rarely did more than scratch the surface of the problem, and that even to scratch the surface, as the author might have done, was still very worth while.

Mr M. E. Ogborn referred to three previous attempts which had been made to deal with mortality statistics. The first was the formula devised by Thomas Young in 1826 to represent the curve of deaths throughout life. The next, which was mentioned in the paper, was the one due to Thiele, which the author had done a service by refurbishing, and the third was Karl Pearson's attempt to represent the curve of deaths by a series of mathematical curves.

The formula produced by Thomas Young was of interest in the current discussion. Young had been one of those who liked to shroud their methods in mystery. He was not an actuary as such but primarily a medical man, so that he could not be accused of being a typical actuary in representing his calculations as a mystery! It was known that he also devised a new nosology—a method for the classification of disease—and it was a likely guess that he built up his formula by a consideration of the causes of death in much the same way that the author had done.

There were two possible lines of approach to the formula proposed by the author, both of which he had mentioned. The first was to consider a combination of two groups of lives in a mixed group, each of the two groups having its own separate mortality experience. The alternative was to consider the group as a whole, supposed subject to two independent mortalities which were additive.

To take the first alternative, the combination of two groups each with its separate mortality experience, was to adopt the approach of §9, which assumed that the group contained a proportion of people who were impaired lives subject to a higher mortality than the normal—an impairment not necessarily from birth, but possibly from some later age, an impairment which separated them from the others by introducing or intensifying causes of death to which the others were not or were not much subject. If that approach were adopted in the mixed group, each of the independent groups produced a component rate of mortality which was not a true rate; it produced the rate of mortality of the group multiplied by the proportion of the group to the whole. The component rates of mortality, therefore, were each products: the product of the rate of mortality for the particular group and the proportion of the lives in that group subject to the mortality. Since the group with the higher mortality would tend to die out first, one of the components would be of the cocked-hat type, and it might take the shape of the normal curve; but the implication was that the other one, the Gompertz curve, was not subject to true Gompertz mortality. The component would be a combination of a rather flatter curve of mortality with an increasing proportion subject to that mortality. So the position was not so simple as it seemed.

If, on the other hand, the second view were taken, considering the mortality of the group as a whole subject to independent additive mortalities, he thought that if the formula was to have any meaning the mortalities had to be independent. The author seemed to recognize that and claimed that his constants were independent, but it was difficult to agree with him. If, for example, the results for the male mortality and the female mortality were compared, it was evident that there was a big discrepancy in the age at which the normal curve reached its maximum. The function of the normal curve in the formula was quite different for the male mortality as compared with the female mortality.

As another illustration of the point, in the formula which the speaker had himself produced (*J.I.A.* 79, 183 and 190) the adult mortality was in fact represented by two components, one of which steadily increased throughout life while the other was of a cocked-hat shape, but not symmetrical. It was interesting

that the component which was steadily increasing throughout life finished at age 114 for the assured lives and at age 110 for the population data, while the component which had the cocked-hat shape reached its maximum at about age 88 in both experiences. That had been based on data which were many years old, and the result might be different for the data which the author had used, but it did suggest that even if an earlier age were found with the author's data it was impossible to claim that the constants were independent. The fitting of the formula was to be regarded as a whole; an adjustment in one section of the formula would lead to an adjustment in another section.

In §42 reference was made to some of the advantages which might accrue from an analysis of deaths, and with that paragraph the speaker heartily agreed. He thought that the study of the causes of death could throw considerable light on problems of selection and also on problems of forecasting, but he did not think that the study of the causes of death led to much understanding of the mathematical representation of mortality.

Mr R. D. Clarke thought that it might be of interest if he were to draw attention not so much to the similarities but to the difference of approach between the paper which he had presented to the Centenary Assembly and the paper under discussion. He himself had begun with the fundamental notion that the natural life-span of a man was an inherent thing, and that each individual came into existence with, as it were, a little number in his gene-complex giving his maximum length of life. He had based that largely on the evidence provided by Raymond Pearl in his study of the ancestry of the long-lived, which showed that longevity was an hereditary factor. It had then occurred to the speaker that if only those hidden numbers could be read a frequency distribution would be revealed that would be a limiting form of the curve of deaths. He had therefore used what was fundamentally a curve of deaths approach, and the senescent deaths were, in his view, best represented in that way.

That was rather different from the approach used by the author, who used the Gompertz formula to represent the death-rate from senescent causes. Incidentally, the speaker admitted that the term 'senescence' might not have been the happiest one to use. The opener had already referred to the various meanings that it could have and the misunderstandings to which it could give rise. In 1952, M. Bourgeois-Pichat published a paper in the French journal *Population* in which he put forward a theory that seemed to the speaker extremely similar to his own, and in that paper the terms 'endogenous' (for senescent deaths) and 'exogenous' (for the others) were used. It might be that such scientifically neutral terms would reduce the likelihood of misunderstanding.

Returning to his main theme, he thought that the difference between the approach from the curve of deaths and that from the force of mortality was rather more serious than the author had suggested, and it led to a different mathematical interpretation. On grounds of general reasoning—he had not tested it out—it seemed to him that the Gompertz formula might have more relation to the anticipated deaths. Gompertz interpreted c as the capacity to resist destruction. Age diminished that capacity, and therefore there was a greater likelihood of falling a victim to disease; and deaths from disease, as distinct from deaths from the mere running down of the organism, were what made up the bulk of the anticipated deaths. He had one small piece of evidence to put forward, namely the fact that the assured lives experience of the C.M.I. could be represented by a Gompertz curve over the range 35 to 75 fairly well;

and that was the range in which anticipated deaths from disease predominated. Below 35 the accidents were relatively more numerous and tended to swamp the other deaths, while after 75 the senescent deaths became progressively more important, and that, he suggested, was why the curve started to depart from Gompertz. However, he put forward the idea only tentatively.

There could be no doubt, of course, that the great drawback to all work of that kind was the difficulty of drawing aside the veil to get down to the real causes of death. He had found that difficulty in trying to sort out senescent deaths. For example, a large number at old age went under 'bronchitis', but at younger ages deaths from bronchitis were obviously of a different kind. The same could be said of cancer and many other causes. The only causes of death that could be regarded as senescent, apart from old age itself, were in his view—he might be treading on dangerous ground there—myocardial degeneration and general arteriosclerosis.

He would like to ride a private hobbyhorse of his own by referring to the use of the word 'philosophical', which had become usual in discussions on mortality. The remark was not directed at the author, but it seemed to him that what people did when they tried to give an explanation of mathematical formulae was to give a biological or scientific explanation—not a philosophical one. Philosophy was concerned with such topics as the theory of knowledge and logic, and those were not the kind of problem with which people were concerned when they tried to find the interpretation of a mathematical formula representing mortality.

He might seem to have developed his own themes instead of concentrating on the paper. He felt that work of the kind which the author had laid before them was of enormous value to all those who studied mortality, and he would like to pay a whole-hearted tribute to the author for the immense amount of work which he had put into the paper and which the speaker, for one, would find most useful to have on permanent record.

Mr H. Prawitz (a visitor) said that he felt honoured by the kind invitation to be present at an Institute meeting and to take part in the discussion.

Two different streams, seemingly incompatible, could be distinguished in British papers on mortality, the one practical, the other philosophical. He knew that it had become the modern attitude to smile a little at the idea of a philosophical background, but he felt that the philosophical point of view should never be forgotten, and he did not agree that the practical and philosophical approaches were in fact incompatible.

In his opinion there was one, and only one, philosophical background for all scientific work, and that was to describe the observed facts in the simplest possible way. He did not propose then to develop the reasons for that opinion, and it might perhaps be unnecessary to do so in a country where empiricism had been created by great philosophers such as Locke, Hume and Berkeley; but the question arose of what should be regarded as being simple.

Partly it was a matter of personal taste. It was well known, for example, that the Makeham formula was very simple and practical to use when dealing with assurances on joint lives, but that was a simplicity in use, not in description. In dealing with the graduation of mortality it was important to use simple formulae for the convenient determination of constants. It was not, however, that sort of simplicity that he had in mind, nor did he mean that the formula used had to occupy little space. In his opinion a description was to be regarded as simple if it embraced all the different sides of the matter. If observed facts from one point

of view were described in one way, and from another point of view in another way, he would not call such a description simple, even if each one of the two were simple. If, for instance, they succeeded in describing the observed rates of mortality by a simple formula, he would not call it simple if it was not in accordance with biological experience or with medical experience.

He regarded the author's investigation as most interesting. The graduation formulae were in one sense simple and gave a good fit, but from the wider point of view that he had just defined they were perhaps not so simple. Was it really likely that the mortality from a certain cause of death would decrease with increasing age? Could it be imagined, for example, that the power of the human body to withstand destruction (to use an old expression) would decrease in regard to some causes of death and at the same time increase in regard to others? That was, of course, possible, and it was known, for instance, that it was possible to acquire immunity against tuberculosis. In Sweden it had been found that mortality from tuberculosis was represented by a two-peak curve. The decrease after the first peak might be a consequence of the acquisition of immunity, but, since according to medical experience old people would generally not become immune, it was difficult to explain the decrease after the second peak in that way. So far as other causes of death were concerned it might also be possible to find explanations of decreasing mortality, but he did not think such explanations likely.

If, then, they took the view that in old age the power of the human body to withstand destruction by certain diseases did not increase, and that attacks against the body would not be less frequent, there was in his opinion only one way to explain the decrease in mortality at older ages, and that was, as the opener and Mr Ogborn had already indicated, to assume heterogeneity in the population in relation to some causes of death. If part of the population had a higher mortality from a given cause of death than the rest of the population, that part would be reduced more rapidly than the rest, and the result would be a decreasing mortality from that cause of death in old age. A description which took account of such heterogeneity would lead to more complicated mathematical formulae, but from his point of view would be more simple.

There were some points of detail which were of interest to him. Comparing the author's investigation with a similar investigation that had been carried out in Sweden, he found an important difference. With regard to causes of death showing decreasing mortality in old age, the author's curve had peaks at a considerably lower age than the Swedish, and they were more symmetrical. He wondered whether that difference was material, or whether possibly it might be fictitious. Sometimes deaths were recorded with the cause as unknown, or with ill-defined causes. That heading played no great role in Swedish statistics.

There was another of the same type, called 'senility' or 'old age'. According to modern ideas, old age was not a cause of death, and by careful diagnosis it would always be possible to state the real cause of death. As a consequence of that conception, 'senility' played an ever-shrinking part, but it had once had a considerable role in Sweden. In 1917, of all deaths at ages 70 or over no fewer than 46 % were classed as caused by senility, but in 1950 the percentage had shrunk to 16. Since the total mortality at those older ages had not decreased much, the change caused an increase in the mortality from other causes of death. To avoid such fictitious changes, they had spread the deaths from senility amongst the other causes, with some exceptions, in proportion to the deaths observed

from each cause. Such an apportionment would not be appropriate if the percentage of deaths from senility was high, and therefore they had restricted their investigation to a comparatively late period.

He would like to ask the author whether, in the British national statistics, the deaths from senility or from unknown or ill-defined causes formed a considerable part of the deaths at old ages. If so, that circumstance might explain the difference between the British and Swedish curves. The percentage of deaths from senility always increased with age, and if that percentage were considerable it would affect the curves just in the way observed.

The author had suggested that in future the deaths should be divided into eleven groups, and apparently wanted the cause of death to be fully recorded on the death certificate. It was, the speaker said, no business of his to give advice, but he would like to mention what had been found to be advantageous in Sweden. In Sweden the published national deaths data were divided into more than a hundred groups of causes of death. In treating mortality they gathered them into six groups for males and seven for females, but they had an intermediate stage of division into twenty-five groups. No intensive treatment was given to those twenty-five groups, but it proved to be advantageous for the preparatory work to have that subdivision, and it did not increase the labour considerably to do the gathering in two steps. If in future they were to find that another grouping was better, it would be much easier to use that half-way product than to go back to the raw materials.

Lastly, he would like to describe how they intended to carry out the investigation of the mortality according to cause of death in Sweden in future. They intended to investigate continuously, which they had not done before, the causes of death of the assured population. In Sweden each person had an identification number which followed him from birth to death. By an arrangement made with the Government statistical office, that office would give the number on the death cards. The companies would also put the number on their death cards and leave some columns blank for inserting the cause of death. At the end of each year the companies' cards were to be sent to the Government statistical office, and the cards would then be sorted by identification numbers and the cause of death which was shown on the office's cards would be reproduced on the cards of the companies. In that way exactly the same cause of death would be recorded on the companies' cards and in the national statistics. That might be important for eventual comparisons of mortality among assured lives with national mortality. He did not know whether it was possible to have anything of that kind in Britain.

He would like to conclude with one of the quotations with which the author had prefaced his paper:

The statistician is never on more dangerous ground than when he passes from the mathematical expression of phenomena to speculation upon their causes.

He was wholly in agreement. It was dangerous, but, to make progress, they had to brave the danger, and therefore he was glad that the author had done so.

Mr F. M. Redington said that, not being a statistician, he claimed the right to pass from the mathematical expression of phenomena to speculation upon their causes without any danger whatever, and there were three points that he wanted to make. First of all, he shared Mr Clarke's attitude to senescent deaths, in the sense that each individual had a gene-constitution which affected his

potential longevity in a normal, or perhaps, to be more accurate, a perfect, environment.

He felt sure that the μ_x curve, for senescent deaths, would not be a Gompertz curve. He would expect that, following almost all other natural phenomena, the curve of deaths would be bell-shaped. Assuming that the curve of deaths was something like a normal curve, the μ_x curve, instead of increasing continuously like the Gompertz curve, started very low and increased rapidly until about the mean of the distribution was reached, and then flattened off quite considerably. The mortality experience of extreme old age tended to confirm that hypothesis.

Secondly, he would not expect that the anticipated deaths could be expressed in a simple manner as an additive term of μ . He thought that the anticipated and the senescent deaths would be interlocking in their effects. Almost any disease killed some, but damaged others and affected their mortality later in life. It was interesting, however, that the author's main anticipated cause of deaths was cancer of the respiratory system, which was very much a killer; probably not a great number of people who suffered from it went on until later in life. Influenza or diabetes, however, killed a number at all ages but also left many damaged; and therefore he would not expect to be able to express the effect of such diseases as a simple expression for anticipated μ added to a simple senescent μ , because he was sure that they would interlock.

He thought that the author had produced a graduation but not a biological justification. He doubted whether philosophical bases for mortality were possible, for many reasons, among them that which the opener had mentioned, that the whole picture was radically changed for a generation mortality table. The treatment of disease changed every year and the deaths from such things as tuberculosis changed every year. While he could not expect any success in a philosophical basis for the shape of the anticipated mortality, he believed that it was possible to get some idea of the shape of the mortality curve for the senescent deaths. He regarded that as most important, and perhaps it was the main point to which the Institute might pay attention over the next years, because he thought that there was some evidence to go upon and any success would reveal something of the range of possibilities and might imply a limit for annuitant mortality.

Mr R. E. Beard found himself in close agreement with much of the ground covered by the opener. In particular, he felt that they ought to think more in terms of biology, and perhaps of biochemistry, than of pure statistics.

It was when Phillips presented a paper some twenty years earlier that the speaker had first considered the curve of deaths, and, in spite of the difficulties, he still believed that they should concentrate on $(\mu l)_x$ and not so much on μ_x .

He felt that the so-called philosophical approach to the subject was not really adequate. It would be a rare occurrence if, after a mathematical form for a set of statistics relating to some physical phenomenon had been found, anything significant about the reasons underlying the particular phenomenon could be deduced from that mathematical form; and he thought that so far as mortality statistics were concerned, and particularly those relating to human life, that was especially true.

That linked up with the thoughts expressed by Mr Prawitz. Even if a theory of life that led to the form to be expected for the distribution of deaths or for μ_x was known, it would be extremely hard to find statistics to which the formula would apply without some adjustments, and therefore a simple expression for the

statistics would be unlikely to reflect any simple underlying law. An example from another field was provided by the theories of so-called accident proneness, where it was possible to get a set of statistics and fit a satisfactory curve to them, but that curve did not disclose anything about the nature of accident proneness; in fact some recent studies had led to the conclusion that a large proportion of earlier work on that subject was of little value.

A more pertinent example was to be found in the author's paper, where in §3 he stated that he knew of no simple philosophical explanation for a Perks formula. The speaker had been interested in Perks formulae for some years and had looked for mathematical models. He had found three and there might be another hundred and three all having some sort of reasonably acceptable basis. He had not published his three because he was not yet convinced that they had any meaning, although ultimately one might be found.

The first model was to assume that the mortality function, μ_x in that case, was basically Makeham in form—Makeham because he was not satisfied with the philosophy put forward for the Gompertz formula; it seemed to him that it was little more than general reasoning. If the B parameter of the Makeham formula was distributed across the population in a Pearson type III form—which linked up with certain things that had been suggested in the discussion about individuals having life numbers in their genes—the resulting mortality of the population would be Perks in form. Some people might like that way of explaining the Perks formula.

The second model, which was one that he had discovered more recently, was linked up with the so-called 'shot' hypothesis, which again linked up with the idea that individuals had life numbers. On the assumption that individuals in a population had a number of units of resistance to destruction, and that those units were lost, on a probability model, continuously throughout life, at a rate which depended on the number of units possessed by the individual, and that as soon as the total units fell below a certain value the individual died, then it would be found that starting with the proper initial distribution the mortality was Perks in form. Over the course of generations—and again that linked up with the paper by Phillips on time and its application to mortality—individuals would have developed different degrees of resistance to destruction, and it was reasonable to expect that during their living period the resistance was weakened in some way. Those with a high amount of resistance were more likely to lose some than those with a lower amount. That model might appeal to some people as a more satisfactory explanation than plotting figures and deducing a curve therefrom.

The third model resembled the Gompertz hypothesis. It was merely to say that the rate of change of mortality at an instant was proportional to the product of the value at that instant and the difference of the value from a fixed quantity. It was a very simple model, but did not seem to offer much scope for development.

There might be other models even more acceptable. But there was a more important reason for not adopting those models then, and that was simply that μ_x was the ratio of the ordinate of a frequency curve to the area of the tail; in other words, it was $(\mu l)_x / l_x$. It was known that that ratio in a fairly wide range of statistical distributions was well represented by a Perks or the logistic form. The fact that Makeham had fitted some past statistics and Perks more recent data was merely an accident of what was being measured, and neither was more than a convenient mathematical approximation to what they were trying to measure. That was to his mind a good reason for not agreeing with the author's last paragraph. It was better to experiment with $(\mu l)_x$ rather than μ_x .

That led him to the point raised by the author in §40 about the use of his proposed technique for extrapolations. The speaker agreed with the opener about the dangers of that process. Even if the formula had been fitted to the data for a number of years, there would be grave dangers in using it to go far forward. He had avoided using a mathematical expression, and instead had extrapolated on the statistical parameters of the curve of deaths. The method had been illustrated in a note (*J.I.A.* 78, 341) where annuitants' mortality over the past 100 years was used as the basis of an extrapolation. Possibly, however, the method was lost under the extensive mathematics that had been used for convenience in computing.

The part of the paper where the author found it convenient to link up deaths from tuberculosis and cancer of the lung was of particular interest. He had referred to a possible association between them in a previous discussion (*J.I.A.* 79, 205). Some twenty years earlier he had read a monograph entitled *Tuberculosis, Cancer and Zinc, an Hypothesis*, by Dr D. C. Cruikshank, in which by tabulating the mortality rates from cancer and tuberculosis in various parts of the world with the zinc content of the soil, Dr Cruikshank showed that there was a link between them and finally deduced that there was probably some reciprocal relationship between cancer and tuberculosis. The statistics did not stand up to critical tests, but the ideas of the paper had always remained in the speaker's mind, and at various times he had come across pieces of evidence which suggested that there was a relationship between the incidence of tuberculosis and cancer. The most recent information he had read related to certain carcinogenic substances in which zinc played a very important part.

His own work had always been on the basis of arithmetic, and he had refrained from philosophical speculations. Nevertheless he felt that some day a model would be discovered descriptive of mortality. He believed that it would have three basic components; the first part would be concerned with infantile mortality, in which discontinuity at birth would play a large part. To his mind there was a big change at birth in the basic metabolism and the method of feeding of the animal organism, and any formula would have to reflect in some way that change. The next component would be the ageing process, and that might well be continuous from the time of conception. One of the models already described might be adequate, but he thought that it was possible to devise a method based not on general observations nor on common-sense reasoning but built up on some biochemical basis. Finally, there would be a third component which—and there he turned to biology again—related to the main function of any animal, reproduction.

Human beings were unusual among animals in living after the reproductive period. Reproduction would show a rising activity from about age 10 to a maximum at some age and a declining tendency thereafter. He thought it was not a coincidence that the age incidence of tuberculosis was linked with the rising reproductive activity, and the incidence of cancer, particularly lung cancer, with the declining activity; all three were related, and it was known that sex hormones, some tubercular substances and some carcinogenic substances were of related chemical structure. By such associations, he thought, they would obtain some idea of the mechanism of mortality, and more progress was likely to follow than merely by looking at statistics, fitting curves to them and trying to deduce something from the resulting mathematical expression, which might be only one of many forms of representing the statistics.

Those were his own views and he put them forward as being the easiest way

of criticizing the paper. The author had done excellent work, but he could not go far with him, because he did not think the paper took the right road for progress in the study of mortality.

Mr R. G. Barley, in closing the discussion, remarked that the meeting had been very interesting, and it was clear that there was no lack of ideas on the subject. Early in the paper there was a reference to the 'philosophical background', and, like one or two other speakers, he was a little unhappy whenever he saw those words. He thought that if the term 'natural philosophy' had remained in use—and he had always been sorry that it had gone into disuse—they might understand what was meant a little better. He had been pleased to hear Mr Clarke say that what was wanted was a biological explanation.

It seemed to be impossible to talk about the subject without referring to Perks. When discussing Ogborn's paper, Perks had made a very clear distinction between curve-fitting and the search for a satisfactory theory of mortality. The speaker would have liked to have seen the phrase 'a satisfactory theory of mortality' rather more frequently used instead of the phrase 'philosophical background'.

He had been interested in Mr Prawitz's remarks on the philosophical background, and particularly in his views on simplicity. Often they tended to be frightened of something that had more than two or three terms in it, fearing to be told that it was too complicated. On the same subject of the philosophical background, he had been interested in Mr Beard's remarks about units of resistance, because, as some members were aware, he himself had always been an advocate of examining mortality other than human mortality. It was well known that a cat had nine units!

The title of the paper was *The Components of Mortality*, and the word 'components' had a suggestion about it of taking something to pieces. He had thought a good deal about that before coming to the meeting, and had perceived even more than before, as he listened to the discussion, the similarity to a small boy taking the dining-room clock to pieces and being very uncertain how to put it together again. There were many different ways in which the splitting up could be done, and several of them had been mentioned that evening. He for one was always inclined to subdivide according to constitution, and that, he thought, was why they split mortality into male and female; they did so because the two sexes were entirely differently constituted. Probably originally the distinction between male and female mortality had been made instinctively, without any real thought about the reason for it.

He had always found it a little difficult, therefore, to visualize what was really meant by splitting up according to causes of death, particularly because he had been 'brought up' during the past few years by a medical officer who insisted that people did not die from what they suffered from; and he thought that it was what they suffered from rather than the cause of death on the death certificate that would really have something to tell them.

He wished to refer to something that had not so far been mentioned, and to congratulate the author on having persuaded whoever had to be persuaded to put all the figures in the paper. It was very good to be able to see what the author had really been doing and to have that demonstration of the tremendous amount of arithmetic that was necessary to make even a tentative examination of the problem under consideration. Whatever they might individually think, they were all bound to agree that the author had told them exactly what he had done, and it would be of use to them in the future.

The author knew, and they all knew, that his paper could not be the last word in an examination of mortality; but the test of a good paper was that it opened up a number of new avenues. He was certain that members of the Institute, and particularly those with plenty of time on their hands to do a lot of arithmetic, would read the paper carefully and find in it a tremendous number of ideas which they would follow up. He hoped that the author himself would follow some of them up and present the Institute with another paper later on.

The President (Mr J. F. Bunford), in proposing a vote of thanks to the author, recalled the saying that the fruit of interest in any field was never reaped without some spade-work. In the paper they were made aware of the amount of experimental work that had had to be done in the course of the author's examination of the 1949 and 1950 national data and the 1951 1% sample data. Some of the harvest from that field appeared in the shape of the conclusions in the middle of the paper, as confirmed or modified at the end, and in the proposals for dealing with the different causes of death.

Another part of the harvest was the most interesting discussion that had taken place on the inexhaustible subject of mortality. It had touched on biology, philosophy, science and so on. He had noticed on other occasions that the quality of the discussion frequently varied inversely with the size of the audience.

It was his duty to refer primarily to the author, but he thought that the members present would excuse him, and indeed applaud him, if he broke with tradition to the extent of thanking their friend Mr Prawitz very much for his contribution, which he had crossed the seas to present to them.

The last word in the paper reminded him that he once knew, or thought that he knew, what an asymptote was. That had been in the high and far-off times (to use Rudyard Kipling's phrase) when the author had been one of the students in his Part I class, and he had been the author's first tutor. He had learnt from the author then, and the process still continued. It was therefore a special pleasure to thank him in the name of the Institute for his great and whole-hearted industry in preparing the paper and for the clarity of its presentation, and to invite him to reply to any points in the discussion that seemed to call for an immediate response.

Mr H. A. R. Barnett, in reply, thanked the President for his kind remarks. He had expected and would have liked to hear a little more criticism. He was disappointed, for example, that no one had said anything about his suggestion for fitting the mortality at the youngest ages. He agreed with some of the main criticism, which was concerned with the labels that he had tied on to the different components that he had suggested, and he was not trying to pass on that criticism when he said that he had deliberately tried to keep to the definitions previously used by Mr Clarke. He agreed that to tie the label 'senescence' to the part which appeared to be fitted best by Gompertz was not the happiest choice.

Many of the opener's points had been based on a comparison of the figures in the paper relating to the 1949, 1950 and 1951 data, and he did not consider that they were really valid comparisons. His work on the 1949 and 1950 data had been purely exploratory. It was not certain what the data really were because the exposed to risk in those two years were unknown. Any apparent discrepancy between 1951 and the previous years merely showed that there was probably something wrong with the 1949 and 1950 fits, because he could not know exactly what data he was fitting.

On the question of senescence he agreed with the remark that each individual carried a label with a certain number on it, but he did not agree that there could not be a senescent death at a young age; some people had low numbers, which might have been derived from something which occurred before birth. At the same time, he did not consider that predisposition to a certain disease necessarily dated from birth or before. He understood the opener to say that they were trying to find why people died when they did. He did not agree; he thought they were interested in finding out when people died. The Why was only incidental, but it was necessary to explore the Why in order to understand the When.

Mr Prawitz's point about mortality from certain causes decreasing in old age had been dealt with adequately by Mr Ogborn when he pointed out that the mortality rate was the build-up, the product, of the mortality rate applicable to a certain group and the proportion of the total lives in that group.

On the point made by Mr Redington and Mr Ogborn of the constants being interlocking, there was or might be some dependence, but it was not so marked as in some of the other formulae that had been suggested for mortality. To misquote George Orwell, 'All constants are dependent, but some are more dependent than others'.

Mr Beard had mentioned extrapolating and the possibility of using data from year to year. He, personally, would never suggest that. He had found, in some of his investigations, that one epidemic was likely to occur in each four-year period. He did not mean that they came round exactly once in four years, but in a four-year period there would be a year with an epidemic; there would probably be one epidemic and there might be two. He had found that four years was quite a good cycle to work on, and, no matter what formula was used, with a four-year cycle there would be a good chance of success in extrapolating.

Finally, Mr Barley had mentioned the dining-room clock. He, also, had experienced difficulty in putting it together again, but he submitted that he had done it in such a way that the parts fitted.

Mr A. Duval has written as follows in amplification of his remarks:

My remarks on the nature of senescence were not very clear and it is evident that they were misunderstood by some subsequent speakers. I think it is important to understand that if the hypothesis is correct that a person has an inherited natural life span, it arises because of an inherited innate rate of senescence (as defined in my remarks in the discussion), the end of the natural life span being reached when senescence has progressed so far that the body can no longer sustain life. Superimposed on this innate senescence is an acquired senescence resulting from the strains and diseases to which the body is subjected during life. Even the deaths from degenerative diseases are, therefore, probably occurring before the end of the natural life span, but they are also occurring amongst that section of the population with the longer natural life spans, since those persons with shorter natural life spans will have greater rates of senescence and hence be more likely to succumb to disease and become 'anticipated' deaths. (Mr Clarke, of course, mentioned in his centenary paper the possibility that the anticipated deaths might be a selective force.) Improvements in medical science and in social conditions that reduce the anticipated deaths may result in the survival to the older ages of lives with a shorter natural life span, but such improvements will almost certainly also reduce the wear and tear (acquired senescence) on the lives that would have survived in any case. It

will not, of course, be possible to determine the net effect on the mortality rates at the older ages because there are so many other factors that are also changing which affect those rates. It is, of course, likely that senescence varies in kind as well as in amount. Mr Prawitz's theory of predisposition can be regarded as an extreme example of variation in kind of senescence.

Mr H. A. R. Barnett writes as follows in amplification of his reply to the discussion:

I made no attempt to assign a biological meaning to 'senescent', which merely means ageing. The opener justifiably assumed that I was following Mr Clarke's definition faithfully by regarding senescent deaths as being those occurring from degenerative diseases, and I am sorry that I did not make this clear in the paper. I had merely intended the label to apply to those deaths which were due to forces operating with increasing intensity throughout life. I deliberately use the word forces rather than 'causes' because I agree with Mr Barley that we would like to be more concerned with what people suffer from than the causes stated on the death certificates. I hold the view that the stated cause should only be a rough guide to our investigations, and that is one reason why the number of cause groups should not be too large; not only does the wood contain numerous varieties of trees, but also there are many hybrid varieties and any attempt to analyse the wood by counting the trees of each variety would be unlikely to succeed; but if the trees are classified in large families the effort may not prove in vain. I intended to convey the idea that stated causes of death do not necessarily keep to their water-tight compartments, or components, when I wrote the last sentence of §10.

Similarly, I did not intend the normal anticipated deaths to be regarded as wholly applying to any specific causes, although the 'killers' that I had cited do indicate closely the shape of the normal curve representing this component; the height of the peak is not determined until certain other unspecified deaths are included. As I have already said, I think the apparent discrepancy the opener has shown between the figures for 1950 and 1951 is due entirely to the fact that the 1950 graduation was not a true fit.

Possibly the best improvement in the labels would simply be to call the first and second components, respectively, 'deaths from late killers' and 'deaths from normal earlier killers', retaining the word 'normal' to indicate the shape and also to make it clear that the abnormalities described in the fifth and sixth components are excluded.

Mr Clarke's suggestion that it may be the anticipated deaths that can be represented by a Gompertz curve, 'while after 75 the senescent deaths become progressively more important', is interesting, but I doubt its soundness because the implication would seem to be a combined mortality curve steeper than Gompertz after age 75. So far as I know this has never accorded with the facts.

Taking Mr Ogborn's first approach to my formula, I have considered what the effect would be if the basic mortality rate for the 'normal earlier killers' is a Gompertz curve—a very steep Gompertz killing off most of the predisposed lives before old age—and if the resultant proportions that the surviving predisposed lives at any age bear to the total population at that age are in the shape of a normal curve. Indeed, bearing in mind that these proportions would increase to a maximum and then fall away, the 'proportion curve' might well be bell-shaped. This would give a mortality rate, being the product of a normal curve and a

Gompertz curve, which would itself be in the shape of a normal curve,* and would constitute my second component. The proportion of 'unpredisposed' lives would still be very large at all ages, and therefore no effective distortion would ensue if the mortality rate for the later killers were still to be represented by a Gompertz. Thus the first two components of my formula may be the result of what is basically a double Gompertz.

In this connexion, it is interesting to recall that the normal earlier killers do not have nearly such a destructive effect on females; when I did some preliminary work on the data I thought the first two components of the female curve formed a double Gompertz, but subsequently when I made a more detailed investigation it was necessary to modify this by replacing one of the Gompertz components by a normal curve with its peak very late in life.

Mr Prawitz asked about deaths from unknown or ill-defined causes. In England and Wales in 1951, out of over 280,000 male deaths, 50 were classified as 'ill-defined and unknown causes of mortality' (I.C. 795) and are therefore not important. 3,535 were classified as 'senility without mention of psychosis' (I.C. 794), of which 1197 applied to the age group 80-84 out of a total of over 30,000; this would be a considerable percentage, from the point of view of the graduation, had the deaths in cause groups at these high ages played any part in the fitting of the curves. It was at the earlier ages that I found the cause groups so important, and here the ill-defined causes (including senility) were not very many—certainly not sufficiently numerous to give rise to any suspicion that the drop in the mortality rate from the 'normal earlier killers' after the later sixties is spurious.

I think it exceedingly doubtful that these senility deaths would include any respiratory tuberculosis or neoplasms, but their existence makes it still more important that we only classify our deaths in broad groups, and not too many of these.

It is difficult to comment on the difference between the conclusions of my investigation and those of Mr Prawitz without knowing more about his data. I suspect that neoplasms of the respiratory system have a less marked effect on mortality in Sweden, and if so that could account for the difference Mr Prawitz has pointed out.

I do not find any of Mr Beard's explanations of the Perks formula wholly satisfactory. His first explanation takes into account something that happens to the B parameter, and I agree that if we write the formula as

$$A + \frac{Bc^x}{1 + Dc^x}$$

it can be rewritten as

$$\frac{A + (B + AD)c^x}{1 + Dc^x} = \frac{A + B'c^x}{1 + Dc^x} \quad (\text{say})$$

which is one of the more usual forms. But this is only because the identical exponential constant c is used in numerator and denominator and to my mind this is a weakness, though a convenience, of the formula. Alternatively, just what does the term $A/(1 + Dc^x)$, which decreases with age for the usual values of c and D , represent? If it is the tail of the childhood curve, it should fit the childhood ages but I doubt whether it would do this.

* Vide *Frequency Curves and Correlation* (Elderton), 3rd ed. p. 120.

Similarly, where does the constant A fit into Mr Beard's second explanation? And could he try any one of his three explanations on a layman and come away satisfied that it was simple?

My main criticism of the Perks formula from a theoretical angle (I have no criticism from the purely practical angle) is that it does not take account of natal mortality, childhood mortality, or the peculiar things that happen in the twenties, and I think from some of Mr Beard's remarks that he agrees. It is because I do not think that any simple 'law' of mortality will be found which applies throughout life that it seems to me to be so important to break it down to its components. It may be that over age 30 or 35 a Perks curve will prove to fit better than my exponential plus normal, but this I doubt in view of the strong influence of neoplasms of the respiratory system in the sixties. Surely this hump cannot just be lost in an all-embracing law purporting to cover most of the adult ages, and it is my view that the fact that Perks curves fit so well over a certain range is to some extent, as Mr Beard has suggested, accidental-convenient, eminently useful from the practical point of view, but accidental.

After reflection on the remarks of three of the speakers about the interdependence of the constants in my formula, I still feel that the degree of dependence would be negligible for extrapolation purposes, whereas the dependence of the Perks constants B and D would be considerable.

I think Phillips's approach may yield results in the propounding of a law of mortality, and I believe the function ϕ may have this use—and this only. I cannot see how this complex function can possibly have any practical use—complex because it depends not only on the force of mortality at a certain instant, but also on all the forces of mortality at earlier ages. And so I reiterate, in disagreement with one speaker, that I hope future research will not close its mind to any function.

I have tried to steer a middle course between Phillips, whose purely theoretical approach covers the whole of life, and Perks whose practical approach covers a limited range only.

Finally, I should like to return to my recommendation that causes of death should be recorded in our mortality data, by summarizing the opinions expressed on this matter in the discussion. The opener would prefer an investigation on the lines suggested by Mr Clarke in his centenary paper. Mr Ogborn is in favour of my suggestion for some but not all of my reasons. Mr Prawitz is obviously in favour, and if the Institute were to sanction a detailed investigation such as was carried out in Sweden I should be delighted. I think Mr Redington agrees by implication, in view of the importance he attaches to getting an idea of the shape of the mortality curve for senescent deaths. Mr Beard and Mr Barley did not commit themselves, but in view of the former's interesting remarks on cancer and tuberculosis, and the latter's admission that he was in favour of splitting things up according to constitution, I do not think they can be ranged against the suggestion. To sum up, the majority, if not all, of the speakers are in favour of a cause of death investigation, and no member felt sufficiently strongly against it to say so at the discussion. May I therefore make a further plea for the Council to give it their consideration.