JIA 115 (1988) 721-725

THE IMPROVING ACCURACY OF THE BASIC DATA AND THE EXPOSED TO RISK OF THE ENGLISH LIFE TABLES (UPDATED TO E.L.T. NO. 14)

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

1. INTRODUCTION

1.1 The publication last year of the English Life Table No. 14 enables my earlier paper (Daw, 1982) to be extended to include the latest table. This paper applied Redington and Michaelson's r_x test to E.L.T. Nos. 8 to 13 and the corresponding figures for E.L.T. No. 14 have now been calculated. The Government Actuary's Department kindly supplied me with the unpublished exposed to risk figures which were used.

1.2. The r_x test was originally put forward in 1940 and is briefly described in the Appendix to Daw (1982). However, in his collected papers Frank Redington (1986) "re-presents the idea in its basic simplicity". r_x is the third difference of the ungraduated central rates of mortality divided by its standard deviation. The end result of the r_x test is a value for the variance of r_x (denoted by σ_r^2) which should be around unity and any significant excess over unity is regarded as a measure of the inaccuracies of the data and methods of constructing the table. My 1982 paper found a progressive reduction in the values of σ_r^2 from E.L.T. No. 8 to No. 13, indicating the improving accuracy of data and methods. Table 1 of that paper set out certain information in respect of the six E.L.T.'s examined and Table A below gives the corresponding information for the new E.L.T. No. 14 and repeats that already given for E.L.T. No. 12 and No. 13. As in the 1982 paper the basic data considered here are for ages 20 to 90 only and the investigation is not concerned at all with the method of graduation or the graduated rates.

2. ENGLISH LIFE TABLES NO. 14

2.1 Table B gives the values of σ_r^2 based on the data for ages 20 to 90 for E.L.T. No. 14 compared with those previously given for E.L.T. No. 12 and No. 13. The Appendix to Daw (1982) gave approximate 5% confidence limits for σ_r^2 when calculated from data for ages 20 to 90 as being 1·59 and ·55. On this basis the value of σ_r^2 in Table B for E.L.T. No. 13, males, falls between these limits and that for females not far above the upper limit. In the case of E.L.T. No. 14 the position is reversed and the females fall within the limits while the males are negligibly above the upper limit. However, it seems to me that it cannot be argued that the errors and inaccuracies of the data and methods have been largely eliminated until some values of σ_r^2 less than 1·0 begin to appear.

Table A. *Methods used in obtaining data for the English Life Tables*

English Life Table No. (and date of census)	Information asked regarding age at: Census	Death	Exposed to risk at each age	Method of Graduation
12 (23.4.61)	Age in years and completed months	Age last birthday in years	Estimated as accurately as possible from census population and the tabulations of deaths in the Registrar-General's Statistical Reviews	Mathematical formula
13 (25.4.71)	Date of birth	Date of birth	As for E.L.T. No. 12 except for one quite minor improvement	Cubic splines
14 (5.4.81)	Date of birth	Date of birth	Derived from "a reliable series of OPCS population estimates" as at 30 June in 1980, 1981 and 1982. The figures for ages in the late fifties and the sixties were then adjusted to take account of the irregularities in the annual number of births from 1915 to 1923 (see Daw (1982) Fig. 1) and of the deaths in these age groups during 1980 to 1982. See also § 5.	Cubic splines

Table B. *Values of σ_r^2 using data for ages 20 to 90*

E.L.T. No.	Census Year	Males	Females
12	1961	4·48	2·94
13	1971	1·25	1·93
14	1981	1·61	1·32

3. VALUES OF σ_r^2 FOR 10-YEAR AGE GROUPS

3.1 On studying the values of r_x at individual ages for the three life tables included in Table B, I noticed that the larger values (say $|r_x| > 2{\cdot}0$) occurred much more frequently in the last 10 ages of r_x (i.e., $x = 78$ to 87) than at the younger ages. (Because of taking third differences of ungraduated m_x in the calculation of r_x, the last value is r_{87}, although data for ages to 90 are used.) Accordingly values of σ_r^2 were calculated for 10-year age groups of r_x and these are given in Table C. For what they are worth, the approximate 5% confidence limits for σ_r^2 based on 10 values of r_x are 2·79 and ·11.

Table C. *Values of* σ_r^2 *for 10-year groups of* r_x

E.L.T. No.	Age Group 28–37	38–47	48–57	58–67	68–77	78–87
Males						
12	·68	2·28	2·52	7·05	5·74	11·39
13	·24	·35	·71	1·91	·49	3·83
14	·80	1·96	·48	1·24	·37	5·66
Females						
12	·32	·82	3·27	5·43	4·50	4·43
13	·41	·76	1·38	3·79	2·08	4·45
14	·64	·22	1·51	·65	1·41	3·91

3.2 Table C shows that with few exceptions:

(i) For each sex and E.L.T., σ_r^2 for ages 78 to 87 is substantially greater than those for any other age group (one exception) and is also far above the upper confidence limit.

(ii) For E.L.T. No. 13 and No. 14 σ_r^2 for all age groups except the highest lies within the 5% confidence limits (one exception).

(iii) For E.L.T. No. 12 only the two or three youngest age groups give values of σ_r^2 within the 5% confidence limits and the values of σ_r^2 are always greater than those for E.L.T. No. 13 and No. 14 (three exceptions).

3.3 In Daw (1982) it was suggested that the lower values of σ_r^2 in Table B for E.L.T. No. 13, as compared with E.L.T. No. 12, were due to asking, for the first time, for the date of birth (instead of for age) at census and death registrations. Table C gives the impression that this improvement has been spread over all the 10-year age groups and is not confined to any particular section of the age range. Also the improvement shown in E.L.T. No. 13 is maintained in E.L.T. No. 14.

3.4 Bearing in mind the large variability of σ_r^2 when based on only 10 ages, there does not seem to be any noticeable difference between the pattern of values in Table C for E.L.T. No. 13 and No. 14.

4. CONSIDERATION OF DATA FOR AGES 20 TO 80

4.1 Because of the results in Table C the values of σ_r^2 were calculated using the data for ages 20 to 80, i.e., omitting the last ten ages; these values are given in Table D and the approximate 5% confidence limits for them are 1·64 and ·52.

4.2 As would be expected, all the values of σ_r^2 in Table D (ages 20 to 80) are less than the corresponding ones in Table B (ages 20 to 90). However, for E.L.T. No. 12 both values are still greater than the upper 5% confidence limit. The values of σ_r^2 for E.L.T. No. 13 and No. 14 all lie within the 5% confidence limits but, even more interesting, three out of the four values are less than 1·0 (see comment at end of § 2.1). However, it would be unrealistic to say that over the age range 20 to

Table D. *Values of σ_r^2 using data for ages 20 to 80*

E.L.T. No.	Census Year	Males	Females
12	1961	3·29	2·68
13	1971	·81	1·49
14	1981	·90	·86

80 all the errors and inaccuracies of the data and methods have been eliminated; some must remain. Perhaps it can be said that the remaining errors are either unimportant or are undetectable by the r_x test, with all its own inaccuracies and assumptions, which may also affect the confidence limits given above.

5. EXPOSED TO RISK AT AGES 20 TO 90

5.1 Table A briefly describes the new method of calculating the exposed to risk for E.L.T. No. 14 which differs from that used 10 years earlier for E.L.T. No. 13. The exposed to risk for No. 13 was arrived at from the 1971 census population and the deaths in 1970–72. It seems likely that the 30 June populations, on which the No. 14 exposed to risk was based, were determined by using the 1981 census and 1980–82 death figures in a very similar way. However, for No. 14 certain adjustments were made which were not included in the No. 13 exposed to risk. These were adjustments for net migration and for the uneven incidence of births from 1915 to 1923 (see Table A). Also additions were made to the 30 June estimates (i) to take account of under enumeration at the census, and (ii) to cover those who were normally resident in this Country but were absent on census night and not included on any census form. O.P.C.S. (1983, 1984a) describe the methods used. The additions set out above will have the effect of bringing the exposed to risk for No. 14 more in line with the registered deaths to which they are related and would appear to decrease the crude central rates of mortality. The decrease would not be expected to result in substantial increases in the irregularities of the third differences of the mortality rates on which σ_r^2 is based. Thus there seems to be little reason to expect the change of method between E.L.T. Nos. 13 and 14 to have any appreciable effect on the values of σ_r^2; this is in fact what has been found above. However, I think it is a pity that the exposed to risk for No. 14 was not also calculated by the No. 13 method. This would have shown the effect of the changed method on the rates of mortality produced and also enabled the values of σ_r^2 resulting from the two methods to be compared.

6. ACCURACY OF AGES

6.1 The results of §§3 and 4 indicate that over the age range 20 to 80 the methods of constructing the two latest E.L.T.'s are accurate within the limits of

detection by the r_x test. But for ages of over 80 this standard has not yet been attained.

6.2 No comparisons have been made of the age information given at the 1971 or 1981 censuses with the corresponding birth registrations. Such studies for earlier censuses have shown a higher proportion of errors at the older ages. For example, Table 5.3 in Gray and Gee (1972) gives error percentages for the highest age group shown (65 and over) which are about double the all-ages figures quoted from the same table in Daw (1982), p. 459.

6.3 A substantial excess in the number of persons recorded as being centenarians has been found at the 1971 and 1981 censuses (O.P.C.S., 1981, 1984b). This must have resulted in deficiencies in the enumerated population at younger ages. For example, as mentioned in Daw (1982), p. 460, Thatcher (O.P.C.S., 1981) found that a small proportion of old people (mostly over age 75) gave dates of birth at the 1971 census which were earlier than the true date, in many cases by exactly 20 or 30 years. It therefore seems reasonable to think that the significantly high values of σ_r^2 shown in Table C for the oldest age group of the two latest E.L.T.'s may be caused by age errors.

6.4 The age information given at death registration appears to be more reliable than that recorded at censuses (e.g., Thatcher, 1987). From this it would seem that any improvements which can be made in the accuracy of the dates of birth recorded at census by the oldest persons may result in the number of centenarians coming more within the range of possibility. At the same time greater accuracy may be brought to the age range 80 to 90. Then perhaps the r_x test might begin to produce values of σ_r^2 for future E.L.T.'s which lie within the 5% confidence limits for the age range 20 to 90, or perhaps even higher.

REFERENCES

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

OFFICE OF POPULATION CENSUSES AND SURVEYS. (1972). Gray, P. & Gee, F. A. *A quality check on the 1966 ten per cent. sample census of England and Wales.* H.M.S.O., London.

OFFICE OF POPULATION CENSUSES AND SURVEYS. (1981). Thatcher, A. R. Centenarians. *Population Trends*, **25,** 11–14. H.M.S.O., London.

OFFICE OF POPULATION CENSUSES AND SURVEYS. (1983). Population definitions. *Population Trends*, **33,** 21–25. H.M.S.O., London.

OFFICE OF POPULATION CENSUSES AND SURVEYS. (1984a). Population estimates 1961–81: a new series. *Population Trends*, **35,** 30–33. H.M.S.O., London.

OFFICE OF POPULATION CENSUSES AND SURVEYS. (1984b). Centenarians: 1981 estimate. *Population Trends*, **38,** 12–13. H.M.S.O., London.

REDINGTON, F. M. (1986). 'The r_x test' in *A Ramble through the Actuarial Countryside*, 19–23. Chamberlin G., (ed), Staple Inn, London.

THATCHER, A. R. (1987). Mortality at the highest ages. *J.I.A.* **114,** 327.