Cross-Correlation of Numbers of Baptisms and Burials in Sixteenth-Century Parish Registers: an Exploratory Analysis*

Michael J. Slater[1]

Abstract

In order to estimate life expectancy at birth for a single parish the statistical technique of cross-correlation has been applied to parish register data. Two adjacent ancient parishes of Giggleswick and Horton in Ribblesdale, both now in North Yorkshire, and five other parishes for comparison, have been studied, mainly for the period of Elizabeth I. Estimates of life expectancy may further be used to estimate population sizes. Life expectancy in Giggleswick was in the low 30s and for Horton parish was 20 to 30 years. Populations of about 1400 and 400 to 700 respectively are calculated. Credible results are also found for Colyton (Devon), Odiham (Hampshire), Oswaldkirk (North Yorkshire), Shepshed (Leicestershire) and Southill (Bedfordshire) for which parish registers from 1538 and 1541 are available and for which other studies have been made.

Introduction

This article reports on the exploratory use of a technique known as cross-correlation to try to estimate the expectation of life at birth in early parish registers from the sixteenth century. The method is applied to the parish registers of Giggleswick and Horton in Ribblesdale in North Yorkshire. The motivation for this paper was an analysis of the population history of Giggleswick using the parish register, which dates from the mid sixteenth century. The Giggleswick parish register is being analysed from the point of view of women's lives, but a secondary consideration was to estimate the life expectancy at birth and parish population at the time of Elizabeth I. Since the parish register is not detailed enough for family reconstitution, and the parish rather small, an alternative technique was investigated: the cross-correlation of numbers of baptisms and burials. Other methods of estimating population size have been considered and found wanting. Using the 1379 tax lists and sixteenth-century lists of tenants for Giggleswick manor (not the whole parish) one can estimate population size but the results depend on assumptions about household size of, say, 4.5 persons per tenant and an uncertain predicted increase in population size between 1379 and a point in time 200 years later. The relationships between the average annual numbers of baptisms or burials, population size and life expectancy at birth, assum-

* https://doi.org/10.35488/lps105.2020.120.
1 Michael J. Slater: mandmslater@mypostoffice.co.uk.

ing a stationary population, give different estimates for life expectancy for Giggleswick and rely on uncertain independent estimates of population.

Giggleswick is a small village neighbouring Settle in North Yorkshire, and Horton in Ribblesdale is a smaller, more isolated, more scattered community further north up the valley of the River Ribble. The ancient parish of Giggleswick comprised the townships of Giggleswick and Settle and the much smaller townships of Langcliffe, Stainforth and Rathmell. The parish registers for Giggleswick and Horton in Ribblesdale are available for the reign of Elizabeth I, but the parish registers of other neighbouring parishes (Clapham, Ingleton and Long Preston) are, unfortunately, not available for this period.[2] Early wills and manorial records have been collected in recent years to better understand the local history of these communities.[3]

Estimating life expectancy in sixteenth-century English parishes

The literature on the topic of life expectancy in historical England suggests a country-wide range of around 27 to 38 years for life expectancy at birth in about 1550, but details of how life expectancy can be calculated for individual parishes are unknown to this author.[4] The extent of variation between parishes is also unknown. Various approaches to the estimation of life expectancy are possible, but all have drawbacks. One approach involves the time-consuming method of family reconstitution, but this depends on having a parish register with adequate data to distinguish individuals. It has been applied to England from 1580 onwards, but not to the earlier period.[5]

Another approach makes use of model life tables, such as those produced by Ansley Coale and Paul Demeny.[6] These life tables are largely based on modern human populations, and so should be used with the *caveat* that the assumed similarity in age patterns of mortality in modern human populations may not apply to Tudor England. Large numbers of records are also needed to obtain trustworthy average figures. The idea is that there is a systematic relationship between the proportions of deaths at any age and the expectation of life at birth.[7] Therefore, if we can estimate the proportion of deaths in some specific age groups, we should be able to estimate the expectation of life at birth. Table 1 gives the

2 R.W. Hoyle (ed.) *The Parish Register of Giggleswick. Volume 1, 1558–1668* (Yorkshire Parish Register Series 147) (Leeds, 1986); R.W. Hoyle (ed.) *The Parish Register of Giggleswick. Volume 2, 1669–1769* (Yorkshire Parish Register Series 151) (Leeds, 1986); North Yorkshire County Record Office PR/HHR 1 Horton Parish Registers 1556–1671.

3 See www.dalescommunityarchives.org.uk [accessed 17 November 2020].

4 E.A. Wrigley and R.S. Schofield, *The Population History of England, 1541–1871: a Reconstruction* (Cambridge, 1989), p. 528, give values of 33.75 for 1541, 37.99 for 1551, 27.77 for 1561, 38.22 for 1571 and 41.68 for 1581. See also A.E. Laurence, *Women in England 1500–1760* (London, 1994).

5 See E.A. Wrigley, R.S. Davies, J.E. Oeppen and R.S. Schofield, *English Population History from Family Reconstitution, 1580–1837* (Cambridge, 1997).

6 A.J. Coale and P. Demeny (with B. Vaughan) *Regional Model Life Tables and Stable Populations*, 2nd edn. (Princeton, NJ, 1983).

7 The expectation of life at birth is an average value and, as such, involves the loss of detailed information since the distribution of ages at death is so skewed. The average fails to reveal substantial infant mortality and the fact that some people live to 70 and 80 years.

Table 1 An example age at death distribution with a life expectancy at birth of 30.2 years

Age group (years)	Proportion of deaths
0–4	0.39
5–9	0.04
10–14	0.02
15–19	0.02
20–24	0.02
25–29	0.03
30–34	0.04
35–39	0.05
40–44	0.04
45–49	0.04
50–54	0.04
55–59	0.04
60–65	0.06
65–69	0.05
70–74	0.05
75–79	0.04
80 and over	0.03

Source: A.J. Coale and P. Demeny (with B. Vaughan), *Regional Model Life Tables and Stable Populations*, 2nd edn (Princeton, NJ, 1983), p. 44. See also A. Hinde, *England's Population: a History since the Domesday Survey* (London, 2003), p. 101.

proportion of deaths taking place at different ages for a life expectancy at birth of 30.2 years.[8] It should be recognized that the large fraction of deaths in the age group 0–4 years has a dominant impact on the life expectancy calculation and, for high mortality populations, is probably the most sensible age group to use. But the resulting average age at death is therefore very sensitive to the under-estimation of infant deaths. The under-reporting of infant deaths for whatever reason will result in invalid values of the average age at death.

The distribution of ages at death can in theory be determined from parish registers using baptism data (in the absence of birth dates) and burial dates for linked individuals. Unfortunately, for Giggleswick parish, there are so many families with the same surname and so few Christian names in use that it is very difficult to be sure that an individual with a known baptism date can be associated for certain with particular burial with a burial date. In some cases the parish scribe has added comments to the burial record such as 'infant', 'servant', 'young boy or girl', 'old man', 'widow of …' or 'wife of …', which make identification more secure. However, in many cases an individual disappears from the records after baptism, due to leaving the parish for a variety of reasons or information not being recorded in the register. Because of this the confidence level in any estimate of life expectancy or any value using the reconstitution method or the distribution of ages at death method for a single parish is not high, particularly if numbers are small. If the

8 A. Hinde, *England's Population: a History since the Domesday Survey* (London, 2003), p. 101, presents proportions of deaths at different ages based on Coale and Demeny, Regional Model Life Tables, for various values of the expectation of life at birth.

number of people dying at advanced ages cannot be determined accurately then the fraction of infant burials will also not be accurate and any average will also be in error. The limited data for Giggleswick show poor agreement with the numbers in Table 1, in part because numbers are so statistically small for Giggleswick. There is probably also a bias towards younger people in determining burial dates. Another method of determining life expectancy was therefore sought in the light of all these uncertainties.

Cross-correlation in theory

As an alternative approach the technique of cross-correlation of baptism and burial numbers has been tried. The great advantage of this approach is that it only requires details of the annual number of baptisms and burials, rather than record linkage, to achieve results. Cross-correlation has been used in studies of mortality and medical matters, but it is thought that the method has not been applied to parish register data. In this work cross-correlation calculates the best estimate of the average time delay between baptism and burial (which is the same as the life expectancy at birth) for a cohort by comparing the pattern of variation of number of baptisms year by year with patterns of numbers of burials some time later using the correlation coefficient.[9] The results for each trial fitting are reported as *cross-correlation coefficients*, varying in magnitude from –1 (the poorest possible match) to +1 (a perfect match). The closest similarity between annual increases and decreases of the numbers of baptisms and the burials at a later time is sought.

To help understand the theory underlying this approach, consider the following simplistic approach. Suppose that each child baptised over a range of years were to die at the same age (rather than a range of possible ages). This age will be the expectation of life at birth. Suppose this is age 35 years. Then the number of baptisms in any year will virtually coincide with the number of burials 35 years later. So if we were to calculate the correlation coefficient between the series of varying annual numbers of baptisms and the corresponding series of burials 35 years later, we will find that the two series will virtually coincide, and the pattern of increases and decreases will be repeated with a cross-correlation coefficient near 1.0 if there is no major inflow or outflow of people. To estimate the expectation of life at birth under such a situation, therefore, all that is required is to calculate the cross-correlation between the number of baptisms and the number of burials, say, 20, 21, 22 … 45 years later (the range should encompass the hypothesised expectation of life). The cross-correlation coefficient will be low for all ages except one, at which it will be very high. That age is the estimated expectation of life at birth. By computing the correlation coefficient we take advantage of the fact that there is variation in the number of baptisms each year: it is this variation that allows the method to work. If the number of baptisms were exactly the same in each year the approach would not tell us anything.

9 The correlation coefficient is described in any standard statistics textbook such as L. Foster, I. Diamond and J. Jefferies, *Beginning Statistics: an Introduction for Social Scientists* (London, 2015).

Of course, in fact the age at which each of those born in a particular year will die is decided by future events unknown and irregular. There will be a distribution of ages at which they die. The life expectancy is the average of this distribution. In general, regardless of the distribution of ages at death, if all these individuals remain in the parish and die there the pattern of the annual variation in baptisms might be expected to influence the pattern of burials some years later. The simplistic case of the same age at death for each person is obviously not applicable because there *is* a distribution of ages at death for all those born in each year. It might be expected that this distribution does not change much over decades (except for short-term instances of plague, epidemics and starvation due to crop failure), in which case a good estimate of the average of the distribution might be found using cross-correlation. If the distribution of ages at death remains the same for each year in the cohort then although the correspondence of patterns might be less than perfect, it is conjectured that an influence on later burials will remain.

Cross-correlation calculations can be made very easily using the formula *CORREL* available in the Excel 2016 spreadsheet software, or *Data Analysis, Correlation* in earlier versions of Excel, or Open Office Calc. An array of a set (cohort) of 10 years of numbers of baptisms is compared with sets of 10 years of numbers of burials taking place say one year later after the first baptism date, then year by year up to 40 years later (after the last baptism date), postulating life expectancy increasing one year at a time, to cover the range of likely life expectancies say up to 40 years. This entails having a 50-year range of burial numbers. A cross-correlation coefficient of a value near 1.0 would show near complete association of baptisms with burials for some given life expectancy, but is most unlikely given the complexity of the factors which affect the association between the number of baptisms and the number of burials. Negative values indicate poor association.

There is a theoretical problem with the cross-correlation approach. As indicated in the simplistic example described above, it should work well if everyone dies at the same age. It is likely to work tolerably well if the distribution of numbers of deaths by age is unimodal, by which we mean has a single peak. But the distribution of numbers of deaths in human populations is far from unimodal. It normally has a mode at very young ages and another mode at older ages. An interesting question is, therefore, whether the cross-correlation method will work under such circumstances.

An initial consideration is that, because infant mortality was so high, it is likely that the peak cross-correlation will occur between baptisms and burials one year later. We are really interested in the peak within the age range where the life expectancy in sixteenth and seventeenth century England might reasonably be expected to lie, say between ages 20 and 50 years. This is a much lower peak.

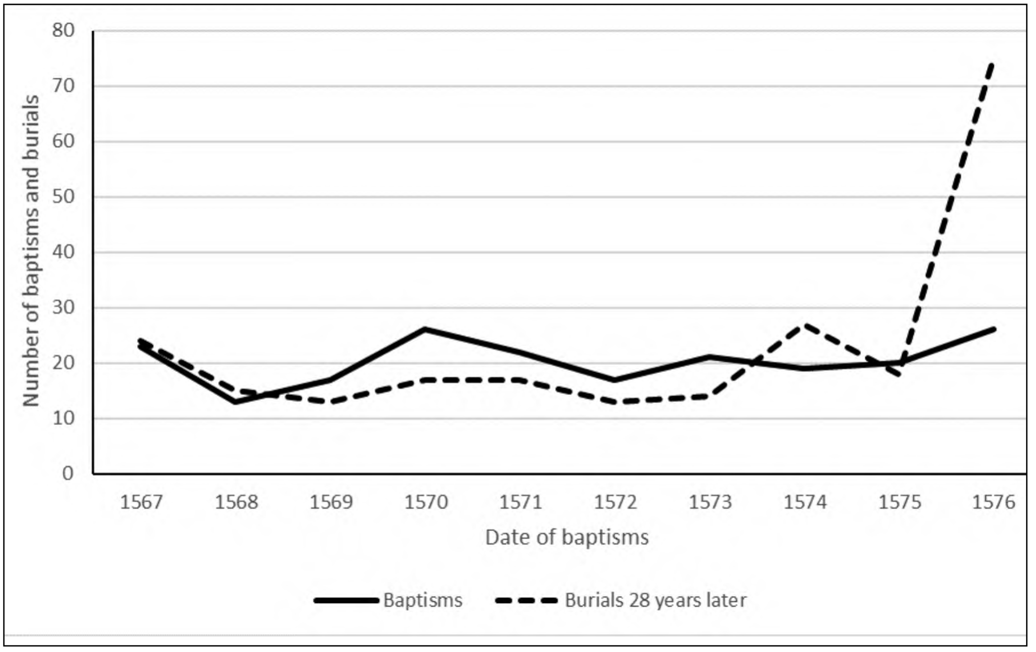
An analytical mathematical proof that the life expectancy at birth can be estimated using cross-correlation may not be possible when the age at death distribution is complex, so a spreadsheet test simulation assuming no inflows or outflows of population was used instead to check the method and to try to establish proof of the concept. The limitation then is that in the case of migration the results will be affected to some degree. The initial test calculations of the numbers of burials were made using the assumed age at death distribution and the resulting life expectancy at birth listed in Table 1, based on the series of Giggleswick baptisms. The age groups of Table 1 were used. The first simulations were made using a long list of actual numbers of baptisms and the calculation on a spreadsheet of the expected numbers of deaths in all following years for 80 years, assuming the age at death distribution, and hence life expectancy, of Table 1. The baptisms for each year were considered, and their deaths distributed over the next 80 years according to the age at death distribution. Thus, for example, if there were 100 baptisms in 1558, then we should expect 39 to die in the period 1558–1563, 2 to die in the period 1563–1568, and so on. This means having values of the fractions of baptisms who die at ages up to 80 years and extending calculations to account for all those dying in old age. Unfortunately, if the age at death distribution is spread over 80 years the calculation of numbers of burials in any year requires knowledge of the previous 80 years of baptisms, which might seem a serious limitation, but is only needed for the simulation exercise. Cross-correlation between the baptism series and the resulting expected burial series was carried out to see if it indicated the same life expectancy (Table 1), as revealed shown by a peak in the cross-correlation coefficient value at a specific gap in years between the baptism and burial series. One can also compare simulated calculated total deaths with known real totals.

The simulations made for this age at death distribution were unsatisfactory because of the averaging of baptism data over 5-year periods, as used for the age fractions at death, which leads to smoothing of estimates of numbers of burials, making it difficult for cross-correlation to find a clear match.

It is important to recognise that the variation in baptism numbers from year to year is important for the method to work. Strong variation in numbers of both baptisms and burials is essential for the method to give useful results. For Giggleswick the standard deviation of baptism numbers between 1558 and 1626 is 13 on an average annual number of 51 baptisms (that is the *coefficient of variation* is 0.25 or 25 per cent). For burials, the coefficient of variation is 0.19 or 19 per cent.[10] The method also relies on the expectation of life at birth and the age pattern of deaths not changing. Changing disease patterns or pathogenicity over a long period will have some effect on the outcome. The first one or two years of any register often seem to show fewer entries than is usual later on, so it would be advisable to start calculations a year or two after the start of the register. The individuals counted in the lists of baptisms are not completely the same as those counted in the lists of burials. Some people moved away, but some moved into the parish: women generally married in the parish of their birth but then moved away; men might have married in another parish and not returned. Younger men might not have been able to inherit a tenement from their father and had to move elsewhere. It can only be assumed that people moving between (local) parishes had much the same lifestyle and life expectancy. The population might have changed up or down in any period chosen but we assume that the population changes

10 The coefficient of variation is the ratio between the standard deviation and the mean.

Figure 1 **Comparison of numbers of baptisms in Horton in Ribblesdale in the years 1568–1577 with burials 20 years later**



Source: Parish registers of Horton in Ribblesdale, North Yorkshire.

through migration were sufficiently gradual that the patterns of annual baptisms and burials can still be compared.

Despite these reservations, the technique of cross-correlation is so much less labour intensive than a family reconstitution method that it seems worth exploring its potential in more detail. It can also be used when parish register data are insufficiently detailed to be very useful for record linkage, as is often the case in early times.

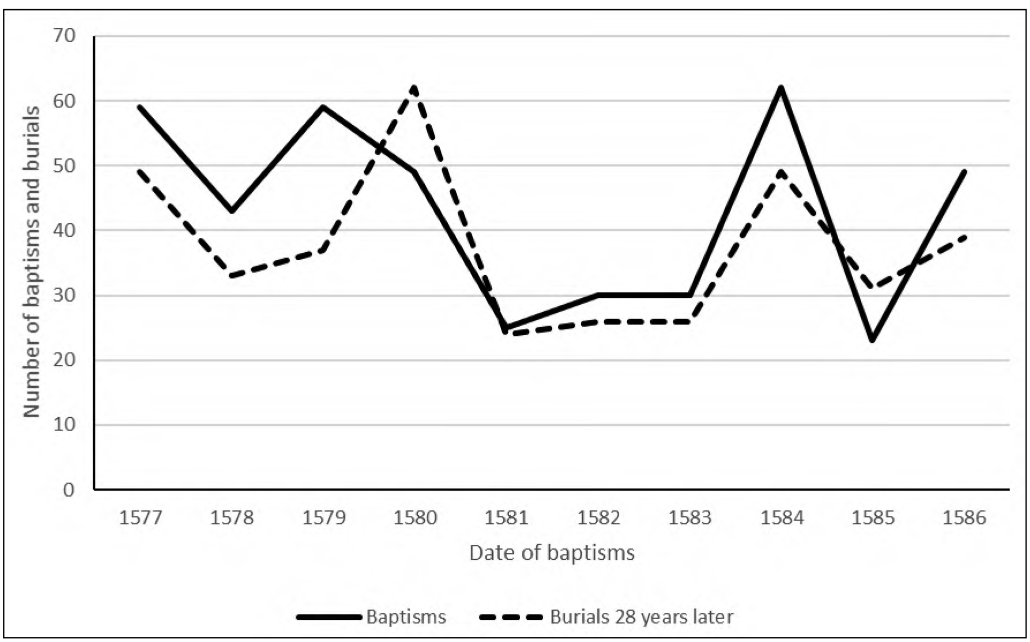
Multiple peaks

Instead of a single clear highest peak value of the cross-correlation coefficient within a sensible age range, sometimes several very similar peak values at different ages for each cohort are found which complicates matters. It becomes necessary to decide which, if any, of the peaks might be valid as an estimate of the expectation of life at birth.

To try to understand the numbers that emerge from the analysis, we should consider that the best results will potentially be found when the temporal patterns of baptisms and burials are identical with numbers of baptisms and burials coinciding. If the temporal patterns are similar, but the actual numbers do not coincide (as for example when a peak in burials occurs due to some epidemic or a subsistence crisis), then the cross-correlation coefficient will be smaller. Figure 1 shows the example of Horton in Ribblesdale in 1597, when the number of burials increased dramatically, probably due to a poor harvest, as 1597 was a difficult year in many areas of northern and western England.[11] If the temporal patterns differ, goodness-of-fit gets worse and again the cross-correlation coefficient will be smaller. These two aspects (differences in the temporal patterns, and similar temporal patterns but different numbers) can be difficult to separate unless the patterns are inspected visually; similarity of temporal patterns is here emphasized over closeness of actual numbers, which can be affected by migration and epidemic years.

One way to look at the similarity of temporal patterns in the numbers of baptisms and burials at some interval of time later is to visually inspect comparison plots of numbers of baptisms and burials at the later appropriate time (Figures 1 and 2). If the temporal pattern is similar, then when the number of baptisms increases from one year to the next, the number of burials should increase from the corresponding year to the next. For example, if the number of baptisms is greater in, say, 1586 than in 1585, then if a 25-year gap is being analysed it should be expected that the number of burials in 1611 will be greater than in 1610. Doing this for all years, one can easily count the number of substantial mismatches in increases and decreases over consecutive years. A mismatch involves an increase in baptisms from one year to the next in conjunction with a decrease in burials from the corresponding year to the next (or *vice versa*). It is common to find a peak cross-correlation coefficient for data which show several mismatches. In cases where there are two very similar peak cross-

Figure 2 **Illustration of mismatches: comparison of numbers of baptisms in Giggleswick, North Yorkshire, in the years 1577–1586 with burials 28 years later**



Source: Parish registers of Giggleswick, North Yorkshire.

11 See A.B. Appleby, *Famine in Tudor and Stuart England* (London, 1978).

correlation coefficients for different gaps between the baptisms and burials, the gap with the fewest mismatches is here generally taken to be a superior estimate of life expectancy. For the few cases with many similar peaks it is more difficult to assess which one is valid.

The number of possible sequences of annual changes for a sequence of 10 numbers of baptisms (a 10-year cohort) is 2^9 (512), so if for example a baptism sequence of 10 is compared with 50 sequences of 10 burials one year at a time, then there is a chance (50 in 512, i.e. about 1 in 10) of seeing a peak which is not Figuvalid for life expectancy estimation. The possibility of seeing an invalid peak is enhanced if there is just one mismatch, i.e. about 2^8 (256) and 2^7 for two mis-matches (128). If only 30 sequences are tried (for postulated life expectancies from say 20 to 50 years) then possibilities of a match by chance are reduced. A test was carried out for the Giggleswick 1577–1586 baptism cohort (28 year life expectancy) by arbitrarily changing data to change the number of mismatches to see the effect on cross-correlation coefficient (Table 2). Clearly, the value of the cross-correlation coefficient declines quickly when there are two or more mismatches in a ten-year cohort.

The width of the cohort in years is important since, for a short cohort of width, say three years, only 2^2 (i.e. 4) annual changes are possible so many multiple fits may be found: for say 25-year cohorts there are 2^{24} (around 10,000,000) changes possible and for 30 years of trial sequences of 25 in a cohort the chance of seeing an invalid peak will be only 30 in 10^7 but with more mismatches likely. Clearly, the larger the width of the cohort the longer the sequence of data required. The number of 25-year cohorts possible to assess is no more than two within the Elizabethan period, rather than a possible six with a 10-year cohort. A cohort size of ten years appears to be a good compromise.

Testing on real data

Cross-correlation was then applied to real data from several English parishes with the discussion in the preceding section in mind. A baptism cohort was chosen and cohorts of burials varied to find the highest cross-correlation coefficient peak and hence an estimate of life expectancy at birth.

Table 2 Effect of changing numbers of mismatches for 1577–1586 birth cohort in Giggleswick, comparing baptisms with burials 28 years later

Number of mismatches	Cross-correlation coefficient
0	0.81
1	0.73
2	0.50
3	0.40
4	0.17
5	0.46
6	0.36

Source: Parish registers of Giggleswick, North Yorkshire.

Table 3 **Peak values of the cross-correlation coefficient for 10-year baptism cohorts, year by year and decade by decade: Giggleswick, North Yorkshire**

Baptism cohort	Estimate of life expectancy at birth	Value of peak cross-correlation coefficient	Population estimate
10-year cohorts			
1558–1567	23	0.70	621
1559–1568	33	0.62	
1560–1569	33	0.66	
1561–1570	33	0.66	
1562–1571	33	0.62	
1563–1572	no match		
1564–1573	27	0.53	
1565–1574	27	0.53	
1566–1575	32	0.53	
1567–1576	32	0.58	
1568–1577	32	0.53	864
1577–1586	28	0.73	756
1631–1640	33	0.51	1,518
1654–1663	21	0.69	966
1700–1709	39	0.54	1,237
1720–1729a	26	0.59	858
25-year cohorts			
1558–1583	33	0.32	878
1568–1593	35	0.40	931
1660–1685	29	0.56	771

Source: Parish registers of Giggleswick, North Yorkshire.

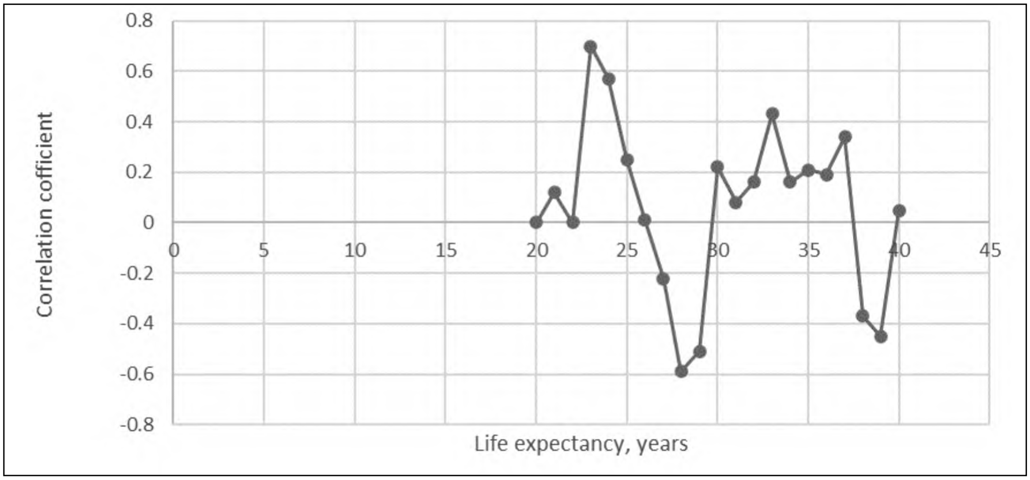
Notes: a. For 1720–1729 there was a cross-correlation of 0.61 with a gap of 45 years, but this had two mismatches rather than the one in the case of 26 years.

Giggleswick

The Giggleswick parish register starts at 1558 but there are data gaps 1627–1630 and 1640–1660 which constrict the periods which can be analysed. The years quoted in this work are the Julian calendar years (March to March) as used in the registers. There were peak burials in 1587, 1597, 1603 and 1608 (due to infectious epidemics of one or two years' duration perhaps, some of which are age-related) which can lower the cross-correlation coefficient. The results are shown in Table 3. Figure 3 shows a typical plot of the variation of the cross-correlation coefficient.

One can reasonably conclude that the life expectancy at birth was about 32 years in Elizabeth's reign, and over the next 100 years or so. Trials were also made for male and female baptisms and burials with similar but less clear results (lower correlation coefficients). They could not be distinguished so any suggestion that women had longer life expectancies than males cannot be supported.

Figure 3 **A typical plot of the cross-correlation coefficient against postulated life-expectancies ranging from 20 to 40 years: Giggleswick, baptisms 1558–1567**



Source: Parish registers of Giggleswick, North Yorkshire.

The variation from year to year in the estimated life expectancies at birth makes it clear that the cross-correlation approach cannot estimate the expectation of life accurately. However, it seems able to give an approximate estimate.

The next stage of the analysis is to use the estimated expectation of life to estimate the total population of Giggleswick. In a stationary population the crude death rate is equal to the reciprocal of the expectation of life at birth.[12] The crude death rate is equal to the annual number of deaths divided by the estimated mid-year population. So, mathematically, if the crude death rate is d, and the population is P, and the expectation of life at birth is e_0,

crude death rate = $d/P = 1/e_0$,

so

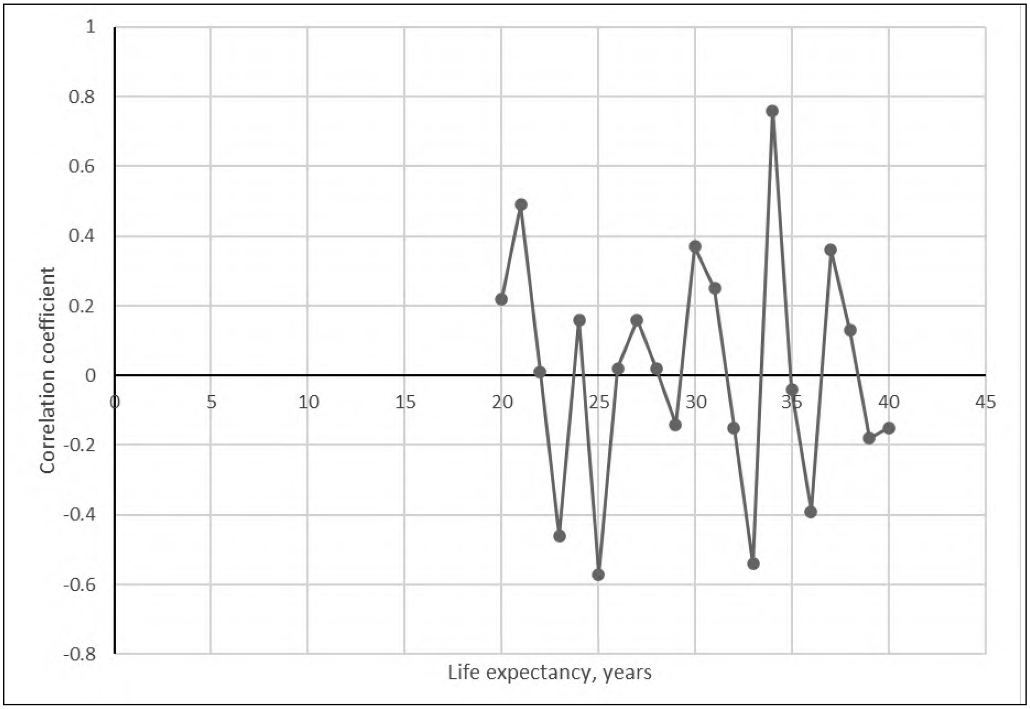
estimated population, $P = de_0$,

so multiplying the annual number of deaths by the estimated expectation of life at birth produces an estimate of the population total.

The average annual number of burials in Giggleswick from 1558 to 1603 was 27. The burial rate for 1603–1640 was 46 per year; 1653–1700 was 46 per year; 1700–1750 was 33 per year. Population estimates assuming a stationary population can therefore be made over these periods, and these are also shown in Table 3. The 1379 Poll tax records are available

12 A stationary population has a crude death rate equal to its crude birth rate and, hence, a population growth rate of zero. It is also effectively closed to migration. See A. Hinde, *Demographic Methods* (London, 1998), pp. 164–6.

Figure 4 Plot of the cross-correlation coefficient against postulated life-expectancies ranging from 20 to 40 years: Horton in Ribblesdale, baptisms 1700–1709



Source: Parish registers of Horton in Ribblesdale, North Yorkshire.

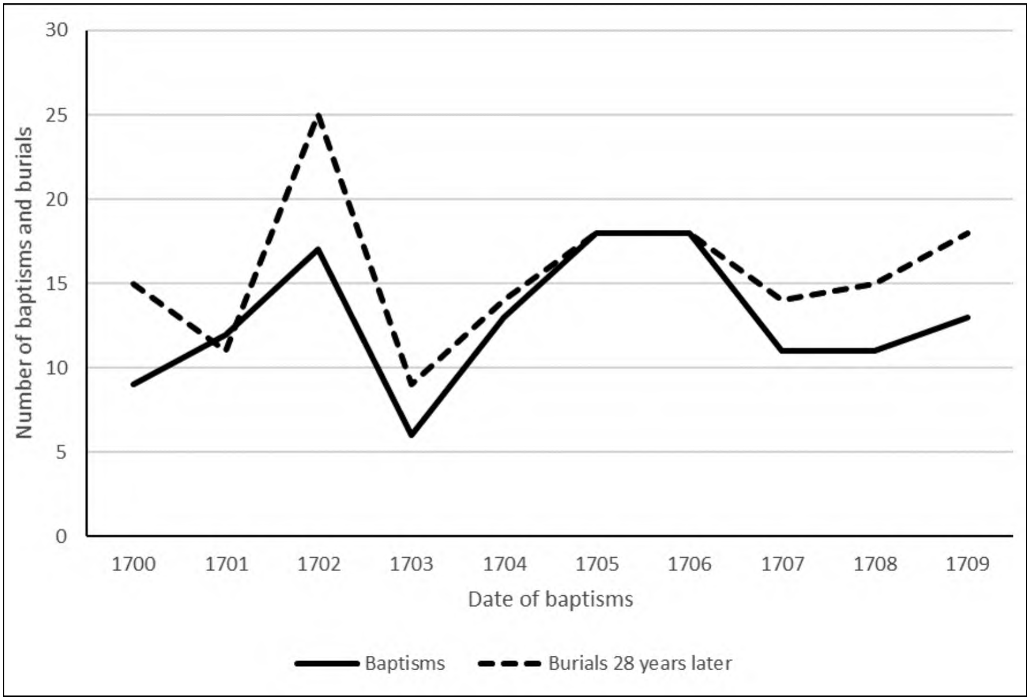
for Giggleswick parish (Giggleswick, Settle, Langcliffe, Rathmell, Stainforth manors) and a population estimate at that time of the order of 300 would be indicated (assuming one third of the population was under the age of 16).[13] Manor Court records for Giggleswick and neighbouring Settle show the number of tenants to be about 85 and 120 respectively in the later period of Elizabeth's reign, so for an estimated mean household size of 4.5 people that means a population for Giggleswick manor of about 380 and for Settle manor 540 (very few tenants held land in both manors). For Giggleswick parish, the numbers in the three small outlying townships have to be added, making a rough total of about 1,400. In comparison York had a population of 8,000–11,000 in the period 1520–1603. The population of Giggleswick civil parish (assumed coterminous with the manor) was 1270 in the 2011 census. The small size of the parish underlines the problem of getting good statistical data for distribution of ages at death and life expectancies.

Horton in Ribblesdale

The problem of applying the approach to a small parish is revealed using data from Horton

13 H. Speight, *The Craven and North-West Yorkshire Highlands* (London, 1892).

Figure 5 **Comparison of the numbers of baptisms in Horton in Ribblesdale in the years 1700–1709 with burials 34 years later**



Source: Parish registers of Horton in Ribblesdale, North Yorkshire.

in Ribblesdale. The parish register for Horton runs from 1556 to 1812 for baptisms and from 1557 for burials with a gap between 1580 and 1584.[14] The baptism cohort 1568–1577 allows calculation to be made with burials in the years 1585–1617. There are about 20 baptisms a year for these 10 years, standard deviation 4.1 (coefficient of variation 20 per cent, or 0.2). The peak is at a 20-year life expectancy (cross-correlation coefficient 0.53 with two or three mismatches) with a lesser peak at 26 years (cross-correlation coefficient 0.44 with four or five mismatches). For the baptism cohort, 1585–1594 the best fit is for a 28 year life expectancy (cross-correlation coefficient 0.50).

One might argue that Horton, being a small scattered community high up in Ribblesdale, suffered harsher living conditions and a lower life expectancy than Giggleswick. The deaths in 1597 (see Figure 1) have the effect of reducing the correlation coefficient by departing substantially from the baptisms line but showing an increase in burials 20 years later along with the slight increase in baptisms, keeping the temporal pattern shape the same. For baptisms in the later period of 1700–1709, the clear peak is at an estimated life expectancy of 34 years (cross-correlation coefficient 0.76 with no other candidates) (Figures 4 and 5).

14 North Yorkshire County Record Office, Northallerton, PR/HHR 1 Horton Parish Registers 1556–1671.

Comparison with other parishes

The Cambridge Group for the History of Population and Social Structure has provided data for 404 English parishes and lists of baptisms, marriages and burials are available from 1538 in some parishes and for later start dates in others.[15] There were about 10,000 parishes in England, so the 404 parishes are around 4 per cent of the total number of parishes, although probably a large fraction of the total population.[16] Five parishes have been chosen for comparison here, with register start dates between 1538 and 1541. The typical processing time is about 20 minutes for each parish (up to 1603).

The first two parishes were Odiham in Hampshire and Colyton in Devon. Odiham was of a similar size to Giggleswick. Three cohorts of baptisms (as for Gigglewick) were assessed, plus one more for 1588-1597 (Table 4). There were about 33 burials per year (1558–1650) so the population was about 1,200. The estimated life expectancies in Table 4 appear clearly identifiable. By contrast, in Colyton, peak values of the cross-correlation coefficient were lower and occurred at ages which seem, in general, to be rather low.

Comparison was also made with three smaller parishes: Oswaldkirk in North Yorkshire, Southill in Bedfordshire and Shepshed in Leicestershire. Annual numbers of baptisms and burials were small; the average number of baptisms per year in Oswaldkirk was 9.3 with a standard deviation of 3.1, which is 33 per cent of the mean, compared with Giggleswick at

Table 4 Peak values of the cross-correlation coefficient for 10-year baptism cohorts: Odiham, Hampshire, and Colyton, Devon, 1538-1597

Baptism cohort	Odiham, Hampshire		Colyton, Devon	
	Estimated life expectancy at birth (years)	Cross-correlation coefficient	Estimated life expectancy at birth (years)	Cross-correlation coefficient
1538–1547			22	0.56
1548–1557			na	na
1558–1567	43	0.60	26	0.45
1568–1577	35	0.67	31	0.52
1578–1587	36	0.60	22	0.51
1588–1597	33	0.71		

Note: For Colyton, 1548–1557 the maximum value of the cross-correlation coefficient was only 0.27 and occurred at an estimate life expectancy at birth of only 19 years.

Source: Parish registers of Odiham, Hampshire and Colyton, Devon.

15 For details of the parishes, see R. Schofield and A. Hinde, *Parish Register Aggregate Analyses*, 2nd edn (Alton, 2020). Although ecclesiastical registration started in 1538, only 27 parishes have good quality baptism registers which start in that year; however 163 parishes have a good quality series of annual totals of baptisms which begins by 1560.

16 The 404 parishes had a population of around 775,000 in 1811. This was 8 per cent of the then population of England, which was 9,538,827 (Census of Great Britain 1811, *Abstract of the Answers and Returns made Pursuant to an Act, Passed in the Fifty-First Year of His Majesty King George III. Intituled 'An Act for Taking an Account of the Population of Great Britain, and of the Increase or Diminution Thereof'. Preliminary Observations. Enumeration Abstract*. British Parliamentary Papers 1812 XI [C. 316], p. 427).

Table 5 Peak values of the cross-correlation coefficient for 10-year baptism cohorts: Southill, Bedfordshire and Colyton, Devon, 1541–1590

Baptism cohort	Southill, Bedfordshire		Colyton, Devon	
	Estimated life expectancy at birth year	Cross-correlation coefficient	Estimated life expectancy at birth (years)	Cross-correlation coefficient
1541–1550	20	0.61	21	0.57
1551–1560	21	0.79	na	na
1561–1570	22	0.58	22	0.53
1571–1580	[27]	[0.39]	21	0.71
1581–1590	23	0.67	21	0.64

Note: For Southill, 1571–1580 the maximum value of the cross-correlation coefficient was low, hence the estimate of life expectancy should be viewed with caution.

Source: Parish registers of Southill, Bedfordshire and Colyton, Devon.

25 per cent. In Oswaldkirk, six cohorts of ten years could be analysed in the period 1538 to 1597. There were multiple peaks with similar cross-correlation coefficients, much more so than for Giggleswick, Horton or Odiham. Southill had an estimated population of about 120 at the time of Domesday and has a current population of the order of 1,000.[17] The annual burial rate of about 16 per year indicates a sixteenth-century population of about 350. Each cohort gives a clear-cut peak in the cross-correlation coefficient and the estimated life expectancies are consistent, if rather low (Table 5). In Shepshed, the annual burial rate was, on average, about 13 per year from 1541 to 1660, which indicates a population of about 300. These days, Shepshed is a dormitory town of about 14,000 people. The cross-correlation coefficient peaks were clear cut and again show a consistently low life expectancy (Table 5).

A further 66 parishes with sixteenth-century registers from the group of 404 have been analysed to test the method for regional variation. For northern England (Yorkshire East Riding, Yorkshire North Riding, Yorkshire West Riding, Cumberland, Northumberland and Lancashire) 19 parishes with 79 cohorts of 10 years have been checked resulting in a mean life expectancy of 32.4 years with a standard deviation 7.7 years. Suffolk was chosen as being a more prosperous and favoured county (23 parishes, 106 cohorts). The life expectancy was 33.9 years with a standard deviation 6.3 years. It is statistically not possible to conclude that life expectancy in northern England is lower than that in southern or western counties, despite the supposition that this could be the case. Finally, 24 parishes, 104 cohorts, in the western counties of Hereford, Gloucestershire and Somerset yield a mean life expectancy of 34.3 years with a standard deviation of 6.3 years.

17 http://bedsarchives.bedford.gov.uk/CommunityArchives/Southill/ [accessed 20 November 2020]

Conclusions

Despite the current lack of or impossibility of analytical mathematical proof, cross-correlation in theory allows an estimate of life expectancy to be made, which might prove to be an adequate method for estimating populations of individual parishes for the time period when reliable data from independent sources are lacking. Although in small parishes the method produces variable results, the consistency in the results in larger parishes is encouraging. The technique is worth exploring further because it is very quick and easy to carry out and can be helpful in cases where the parish register is not very detailed. Simulation attempts to prove the validity of cross-correlation were unsatisfactory, probably because of a non-stationary population and changing age distribution at death. Simulation is also flawed by the assumption of the same distribution of ages at death for all baptism years. This is not realistic and leads to the ineffective cross-correlation of simulated data.

Clearly, migration will have an impact on the values of the cross-correlation coefficient, but if many parishes in a region are subject to calculation of population sizes then the effects of population migration can be mitigated. In order to further assess the cross-correlation technique it would be useful to compare its results with those of the method of family reconstitution for other parishes. However, such a comparison is not definitive since family reconstitution is subject to uncertainty. More importantly, family reconstitution data are not available for the mid sixteenth century. It is a matter of concern that life expectancies reported so far might suffer a problem of under-reporting or under-estimating of infant mortality since the calculation of the life expectancy is so strongly dependent on numbers of children dying within the first few years of life. The apparent low life expectancy in some parishes also needs further consideration.

Despite these reservations it seems reasonable to conclude that life expectancy at birth in the parishes of Giggleswick and Horton in Ribblesdale, in the time of Elizabeth I, was in the high 20s to the low 30s. Such values are comparable to the 27 to 33 years range suggested for other places in the kingdom at this time, averaged over many parishes. However, individual parishes can show much variation.

Acknowledgements

Dr A. Hinde provided the data obtained by the Cambridge Group for the History of Population and Social Structure in the form of spreadsheets listing baptisms, marriages and burials for 404 parishes in England without which this study could not have been extended. He has been most helpful in the preparation of this paper. Gill Newton provided helpful information, particularly that life expectancy data for individual parishes studied by the Cambridge Group were no longer available. Prof. Frank Woodhams made helpful proposals for increasing the speed of calculation and clarity of the text. Michael Pearson has been most helpful in encouraging and discussing the project. The author can provide example Excel spreadsheets if required, contact mandmslater@mypostoffice.co.uk.