4-4-2017

An Attempt to Integrate Previous Localized Estimates of Human Inbreeding for the Whole of Britain

John E. Pattison
University of South Australia, jepattison@dodo.com.au

Recommended Citation

Pattison, John E., "An Attempt to Integrate Previous Localized Estimates of Human Inbreeding for the Whole of Britain" (2017).
Human Biology Open Access Pre-Prints. 103.
http://digitalcommons.wayne.edu/humbiol_preprints/103

An Attempt to Integrate Previous Localized Estimates of Human Inbreeding for the Whole of Britain

John E. Pattison[1]

[1](*University of South Australia*)

Corresponding author: Dr John E Pattison, University of South Australia, Adelaide, South Australia, Australia, 5085.

Issue: 88.4

Running header: Human Inbreeding for the Whole of Britain

Key words: mating relationships, consanguineous marriages, inbreeding coefficient, logistic function, historic Britain.

Abstract

There have been a number of previous estimates of human inbreeding for Britons of British descent in Britain; each generally for different social classes, geographical regions, and/or time periods. In this study an attempt was made to collect all relevant published studies and to combine the results of these disparate studies into an integrated whole for all of Britain. This was achieved by combining weighted means of the percentage of consanguineous marriages (*f%*) reported in these earlier studies: weighted according to the number of records each author examined, the proportion of social classes or geographical regions covered by the records, and the 'merit' of their individual research methodologies. The percentage occurrences of the various consanguineous marriages, from 1st to 3rd cousins, were partitioned into a number of time periods, which allowed the weighted mean percentage inbreeding coefficients (F%) to be obtained as a function of time over the period from 1820 to 1960. The resulting temporal scatter distribution of the weighted F% values closely followed a sigmoidal curve, with a non-linear correlation coefficient of $\eta = 0.974$, which fitted well to a generalized logistic function. After about 1900 the value of the weighted F% was essentially constant at about 0.038±0.004, whereas it decreased rapidly from about 0.256±0.011 between 1820 and 1900. The upper-bound value of weighted F%, before 1820, from the fitted logistic function is 0.276. Note that this corresponds to a value of the conventional mean inbreeding coefficient $F = 0.00276$. As the first known attempt to integrate the earlier disparate values of unweighted F% for Britons of British descent for all of Britain, the results of this analysis are promising and should be useful as reference values in other related studies.

Introduction

The purpose of this study was to examine whether it is possible to integrate previously published results of human inbreeding, for Britons of British descent, in Britain, to obtain estimates of the mean percentage inbreeding coefficient, F%, for all of Britain over the past two centuries. This study is somewhat similar to, but more extensive than that reported by Smith (2001), who only looked at 1st and 2nd cousin relationships and did not attempt to reconcile or combine the results of the disparate studies cited in his report.

In general, the equation for the percentage inbreeding coefficient F% is:

$$\mathrm{F\%} = \sum_{n=0}^{\infty} \frac{f\%_n}{2^{n+2}} \qquad \text{.......... (1)}$$

where $f\%_n$ is the percentage frequency of occurrence of mating relationships (often the frequency of consanguineous marriages) of type n in the population studied. Note that F% is not the usual inbreeding coefficient F, but a factor of 100 times higher. The values for n for the various types of mating relationships are given in Table 1, and rarely include 0 and 1 or exceed 6 in practice.

TABLE 1. Mating relationships

Mating relationship	Inbreeding	
	Index n	Factor 2^{n+2}
Parent/offspring	0	4
Aunt/uncle-nephew/niece	1	8
First cousins	2	16
First cousins once removed	3	32
Second cousins	4	64
Second cousins once removed	5	128
Third cousins	6	256
Third cousins once removed	7	512
Fourth cousins	8	1024

An interest in consanguineous marriages and percentage inbreeding values, in addition to casual curiosity (Bramwell, 1939), arises from academic curiosity concerning marital mobility, demographics, population structure, etc. (Brennan, 1981; Coleman, 1980; Day 2010; Day & Smith 2013; Robinson, 1983; Smith, 2001), and medical curiosity concerning whether any genetic detrimental effects have occurred in the resulting offspring (Bell, 1940; Bundey et al., 1990; Darwin, 1875; Mitchell, 1866; Pearson, 1908a,b,c). Inbreeding values have also been used to model inbreeding over the past millennium based on the genealogical paradox (Pattison, 2001, 2003, 2004, 2007), which apparently corrects for the effective breeding population in the modelling method used. These few studies represent the little systematic information obtained to date on past levels of inbreeding in Britain (Day & Smith). Because of this limited range of data sources they were used in this study even if they were not considered optimum. In this study interest is in breeding not marriage per se. The occasional births due to unrecorded incest, adultery, prostitution, rape, etc. that occur in all societies to some extent "is sufficient to spread genealogical branches both geographically and socially" (Wachter 1980).

The reliability of the previously published percentage frequencies of consanguineous marriages, and the bias affecting them, is also unclear. For example, Smith commented, amongst other problems, that people interviewed often had poor recollection of past generations relying on their memory and oral tradition. Sheets (1981) commented that the people interviewed revealed remote male ancestors more frequently than remote female ancestors, and that some people naturally remembered, or wished to reveal, more than others. This is especially likely if a parent/offspring or uncle/niece or aunt/nephew relationship in the family causes embarrassment. Pearson (1908b) also reported that the respondents did not appear to understand what was required from them. Both Mitchell and Darwin were acutely aware of these problems in their early studies. Even the recent detailed study of Day in which she reconstructed extensive pedigrees of all the people living in two Wiltshire villages over the period studied using all available diverse data sources had its limitations, as discussed by Day, and is subject to some error. Underlying all of these problems is the additional problem of paternal discrepancy, where the assumed (or matrimonial) father is not the biological father (Pattison, 2011). It is generally acknowledged that the method of isonymy overestimates and the method of dispensations underestimate the occurrence of consanguineous marriages, while the method of pedigrees (including birth briefs) can do either (Smith). It is reasonable to assume that some of the published estimates of inbreeding will be lower, while others will be higher, than what the real, but forever unknown, situation was for all of Britain, as such studies were confined to relatively small geographical regions each with their own demographic peculiarities. It is intended by the method of analysis employed in this study, that some of these positive and negative data fluctuations or 'errors' will cancel: where the error in the calculated average will be the error in the sample of the data values being averaged reduced by the square root of the number of such data values being averaged (Hinkle et al., 1988). This is a common technique in science to reduce fluctuations or 'noise' from multiple attempts to measure the same quantity, such as in Bioinformatics (e.g. Hasan et al. 2010). Hence, an indication of how inbreeding has varied, for Britons of British descent, over the past two centuries for all of Britain should be obtained.

Materials and Methods

Data sources. The main type of records used to obtain estimates of *f%* included studies that undertook written and oral surveys, birth briefs, published and reconstructed pedigrees and dispensations by the Catholic Church for the marriage of close cousins. An extensive search was made to locate all appropriate data sources of inbreeding in Britain. The sources used were similar to those used by Smith, with additional results from other studies, and included Bell, Bramwell, Brennan, Bundey et al., Coleman, Day, Darwin, Pearson (1908c), Robinson, and Smith, and the unpublished 1951 results of Pugh and Carter as reported by Coleman. The results reported by Day & Smith are a subset of those reported by Day and are not repeated in this study. Other results not used include those obtained from ABO and Rh blood groups (e.g. Imaizumi, 1974), and by the method of isonymy (e.g., Robinson): The exceptions being Bramwell and Darwin who employed the method of isonymy before it was formalized by Crow & Mange (1965) and also incorporated other methods into their studies in an attempt to correct for the weakness of the isonymy method. They also provided the results in a form that could be used in this study: later studies using isonymy do not. Robinson used both the methods of dispensations and of isonymy, but only the dispensation results were used in this study.

The problem with the method of isonymy is that it requires a number of criteria to be satisfied (Crow, 1980), which are rarely satisfied by communities in historical Britain, and it is found that the method of isonymy consequently produces results higher than other methods for the same population (Rogers, 1987; Crow 1989). Day obtained a result about 50% higher using isonymy than from the reconstructed pedigrees for the same population in Wiltshire and the same time period. While Coleman obtained a result about 20% higher using isonymy than obtained from his marriage survey of the city of Reading. Worse still, Robinson found that the inbreeding coefficients calculated from isonymous marriages were about ten times greater than those calculated from dispensations in his study of the population of the Hebridean Island of Eriskoy. Clearly different regions of Britain satisfy the criteria for the method of isonymy to be applicable differently. But more importantly for the present study, the isonymy method only provides information about the overall rate of consanguinity, rather than what type of cousin relationship was involved and when.

The early studies of Mitchell were crude and the results considered too unreliable (acknowledged by Mitchell himself), and not used in this study. Some other studies were not used because their results were not in a form suitable for this study (e.g., Morton et al., 1976; Sheets), or the results were too early for the time period covered in this study (e.g., Smith et al., 1993). The 11 published sources of data used were all that were relevant to Britons of British descent, and they had good coverage of both rural and urban areas, and of the lower, the middle, and the upper classes in Britain. From these published sources, 35 different marriage situations were provided yielding a total of 106 values for *f%* as potential input data for the present analysis. These results included one parent/offspring and six uncle/niece-aunt/nephew relationships. However, these two types of relationships were judged to be unreliable, because they were too few in number and made a relatively large contribution to the weighted F% result, because their inbreeding factors were small (see Equation 1 and Table 1). Therefore, the study was restricted to extend from 1st to 3rd cousin relationships, which is often the case in these types of studies as people were rarely able to identify relatives beyond 3rd cousins, and where the social (and genetic) implications are small. This yielded a total of 99 *f%* marriage values as shown in Table 2. Zero-value results were included for some relationships if a serious attempt had been made by the author to find the value, and the lack of a value was not merely attributable to the author's lack of interest in that relationship. Following Smith, values designated "other" or "unknown" by Pearson (1908c), Bell and Coleman, which may have included marriages between children of half brothers or sisters, cousins removed in various degrees, and marriages of persons who were cousins in more than one line, were included as 2nd cousins once removed in this study (2½ cousins) because these relationships may be closer in blood than 3rd cousins. The relationships designated "more distant" or "remote cousins" were taken to be 3rd cousins. These approximations were made because there were few such cases and because the inbreeding factors were high, thus hardly contributing to the weighted F% values. Nonetheless, it was deemed better to take them into account than to ignore them. Similarly, Day recorded a small number of 'once removed' and more complex double cousin relationships in her study but included them in the closest appropriate 1st or 2nd cousin relationships. Unfortunately these cousin relationships could not be used separately as 'once removed' data in this study as it was not indicated in which of her five time periods they occurred. A generation length of 28 years was used in this study because it has been shown to be more appropriate than 30 years (Pattison, 2007; Day). However, this was not really an issue, as discussed by Smith) because the allocated ranges of generational years are only rough estimates. The year value associated with each 1st cousin marriage in Table 2 is the center of the range of possible years for that particular 1st cousin marriage.

TABLE 2. Source data with number of records, merit value, f% value, year, weighting factor for each type of cousin relationship, and unweighted F% value

					% Cousin Relationships															
Data Source*	Locatior or Class	Records Number	Data Type	Method Ranking	1st (n = 2)			1½ (n = 3)			2nd (n = 4)			2½ (n = 5)			3rd (n = 6)			Unweighted F%
					f%	Yr	Wt	*f%*	Yr	Wt	*f%*	Yr	Wt	*f%*	Yr	Wt	*f%*	Yr	Wt	
Bundey et al. (1990)	Urban	2,432	Hosp Survey	3	0.165	1978	0.83				0.0411	1950	0.78				0.165	1922	0.71	0.0116
Coleman (1980)	Urban	946	Marr Survey	3	0.106	1972	0.82	0.2114	1958	0.80	0.1057	1944	0.77	0.423	1930	0.74				0.0182
Robinson (1983)	Rural	22	Dispensation	2	0	1970	0.18	0	1956	0.21	18.182	1942	0.24	0	1928	0.27	0	1914	0.31	0.2841
Brennan (1981)	Rural	130	Pedigrees	4	0	1965	0.19	0	1951	0.22	0	1937	0.25							0.0000
Robinson (1983)	Rural	41	Dispensation	2	0	1936	0.25	0	1922	0.29	0	1908	0.32	0	1894	0.37	0	1880	0.41	0.0000
Pugh & Carter (1951)	Urban	5,217	Hosp Records	3	0.441	1935	0.75													0.0276
Smith (2001)	Mixed	626	Birth Briefs	4	0.319	1934	0.43				0	1906	0.41							0.0199
Brennan (1981)	Rural	325	Pedigrees	4	0.308	1930	0.26	0.3077	1916	0.30	3.6923	1902	0.34							0.0865
Bell (1940)	Urban	10,236	Hosp Records	3	0.401	1925	0.72	0.0098	1911	0.68	0.1368	1897	0.64	0.01	1883	0.60	0.01	1869	0.56	0.0276
Bell (1940)	Urban	5,676	Hosp Records	3	0.881	1925	0.72							0.3	1883	0.60				0.0574
Bell (1940)	Rural	489	GP Records	3	1.227	1925	0.28							1.841	1883	0.40				0.0911
Day (2010)	Rural	76	Pedigrees	5	1.316	1907	0.33				0	1879	0.41				0	1851	0.49	0.0823
Smith (2001)	Mixed	1,252	Birth Briefs	4	1.118	1906	0.41													0.0699
Bell (1940)	Urban	21,336	Hosp Records	3	0.595	1905	0.67							0.206	1863	0.55				0.0388
Robinson (1983)	Rural	42	Dispensation	2	0	1903	0.34	2.381	1889	0.38	14.286	1875	0.42	2.381	1861	0.46	11.91	1847	0.50	0.3627
Pearson (1908)	Urban	700	Hosp Records	3	0.857	1895	0.64	0.1429	1881	0.60	0.1429	1867	0.55				0.143	1839	0.48	0.0608
Bell (1940)	Urban	49,315	Hosp Records	3	0.606	1895	0.64	0.0041	1881	0.60	0.1075	1867	0.55	0.035	1853	0.52	0.03	1839	0.48	0.0401
Bell (1940)	Urban	852	Hosp Records	3	0.822	1890	0.62							0	1848	0.51				0.0514
Brennan (1981)	Rural	437	Pedigrees	4	1.144	1890	0.38	1.6018	1876	0.42	1.1442	1862	0.46							0.1394
Day (2010)	Rural	157	Pedigrees	5	2.548	1887	0.39				0	1859	0.47				3.822	1831	0.53	0.1742
Pearson (1908)	Middle	1,600	Marr Survey	3	4.688	1880	0.23				1.6875	1852	0.20	1.125	1838	0.19	0.25	1824	0.18	0.3291
Bell (1940)	Urban	16,803	Hosp Records	3	0.744	1880	0.59							0.167	1838	0.49				0.0478
Robinson (1983)	Rural	71	Dispensation	2	0	1875	0.42	0	1861	0.46	2.817	1847	0.50	0	1833	0.53	0	1819	0.56	0.0440
Darwin (1875)	Urban	55,501	Civil Records	1	1.701	1872	0.57													0.1063
Darwin (1875)	Rural	13,391	Civil Records	1	2.248	1872	0.43													0.1405
Darwin (1875)	Mixed	18,528	Marr Notices	1	3.571	1871	0.16													0.2232
Bell (1940)	Urban	3,400	Hosp Records	3	0.471	1865	0.55							0.118	1823	0.45				0.0303
Day (2010)	Rural	169	Pedigrees	5	3.550	1862	0.46				1.775	1834	0.53				4.734	1806	0.59	0.2681
Brennan (1981)	Rural	559	Pedigrees	4	1.61	1850	0.49	0.5367	1836	0.52	0.5367	1822	0.55							0.1258
Day (2010)	Rural	243	Pedigrees	5	2.881	1837	0.52				3.704	1809	0.58				1.235	1781	0.63	0.2427
Darwin (1875)	Mixed	3,663	Marr Survey	2	3.413	1835	0.13													0.2133
Bramwell (1939)	Urban	40,300	Par records	1	1.5	1823	0.45				2.4	1795	0.40				4.8	1767	0.35	0.1500
Bramwell (1939)	Rural	114,376	Par records	1	2.83	1823	0.55				4.08	1795	0.61				8.06	1767	0.65	0.2721
Day (2010)	Rural	234	Pedigrees	5	0.427	1812	0.58				1.282	1784	0.63				0.855	1756	0.67	0.0501
Darwin (1875)	Upper	11,538	Pedigrees	1	3.9	1790	0.03													0.2438

* Sources are listed in descending order of the year associated with their 1st cousin *f%* values, from 1978 to 1790.

The 35 unweighted F% values calculated for each marriage situation are shown in the last column in Table 2, and plotted in Figure 1, where the urban and rural results are distinguished by different data markers. Neither the urban nor the rural data points show any obvious relationship with time, which was confirmed with a non-linear correlation coefficient (using seven equal time periods) of η = 0.517. The linear trend lines are shown for the plotted urban and rural points taken separately, with linear correlation coefficients of -0.672 and -0.210, respectively. The trends of both sets of data points decrease with time, with a greater rate of decline in the urban sub-population than in the rural sub-population.

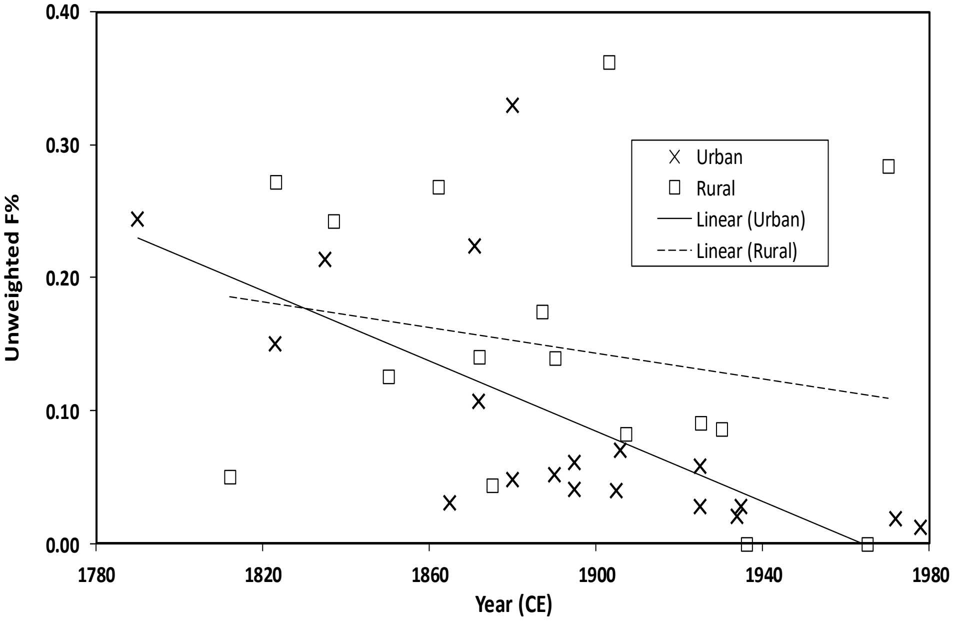

FIGURE 1. Unweighted F% values plotted against estimated marriage year of 1^{st} cousins, distinguishing between the data points that were urban and rural dwellers, each with their corresponding linear trend lines.

To undertake the analysis in this study, the number of records consulted, the relative sizes of the sub-populations examined, either rural/urban or upper/middle/lower-class, and the merit ranking for the different methods employed to collect the source data, appeared to be appropriate weighting factors. Estimates of the percentages of the British population that were upper, middle, or lower class, and rural or urban, in past centuries were obtained from a number of sources, including Coleman & Salt (1992), Eversley, et al. (1966), Harrison (1984) and Heyck (1992a,b), and are plotted in Figures 2 and 3, respectively. Polynomial trend lines were fitted to each set of data points using Microsoft Excel. In both cases, the weighting factors were obtained from the appropriate trend lines for the years of interest, and the values used are shown in Table 2. The results for Smith (2001) are given in Table 2 as ‘mixed’ class as the birth briefs are compiled by members of the Society of Genealogy, London, who may be either middle or lower class: a 2:1 mix was assumed in determining the relevant weighting factor. Similarly for Darwin, two entries are given as ‘mixed’ class where a 2:1 mix was assumed for the middle and upper classes, respectively. It is noted that Darwin initially refers to readers of the Pall Mall Gazette as upper class, but later in his paper refers to them as middle class. Upper-middle class appears more appropriate for professional people such as medical doctors, lawyers and the like (Davidoff & Hall 1987).

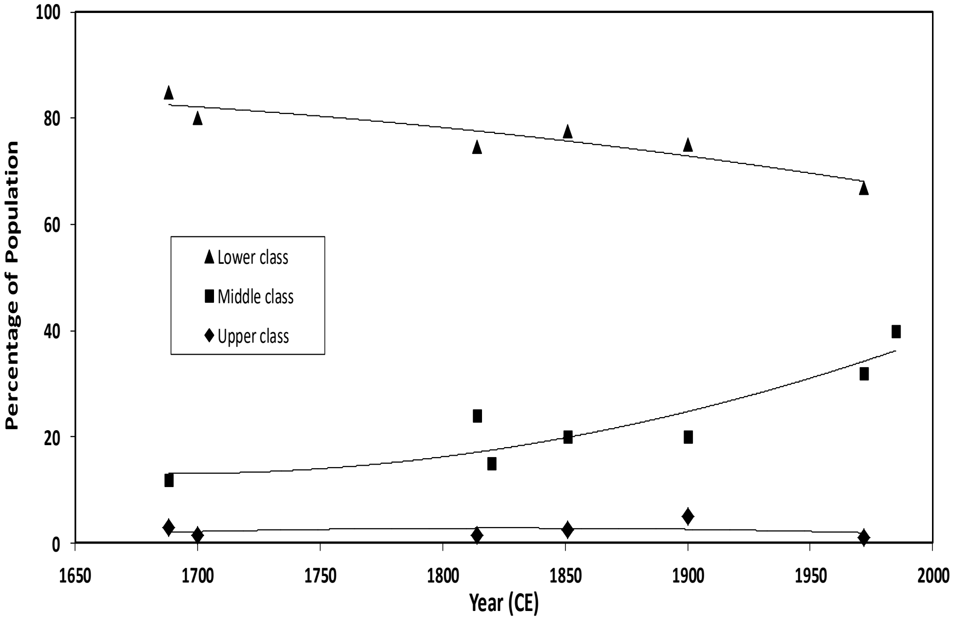

FIGURE 2. Percentage of British population of lower, middle and upper classes, with polynomial trend lines for each class.

The assignment of the merit ranking was based on the method used to collect the source data in each study according to the following scheme: 1 for the early use of isonymy (incorporating other sources) by Darwin and Bramwell; 2 for the use of dispensations by Robinson; 3 for various types of surveys, hospital records, questionnaires, etc., by Bell, Bundey et al., Coleman, Pearson, and Pugh & Carter (1951); 4 for published pedigrees including birth briefs by Brennan and Smith; and 5 for reconstructed pedigrees of Day.

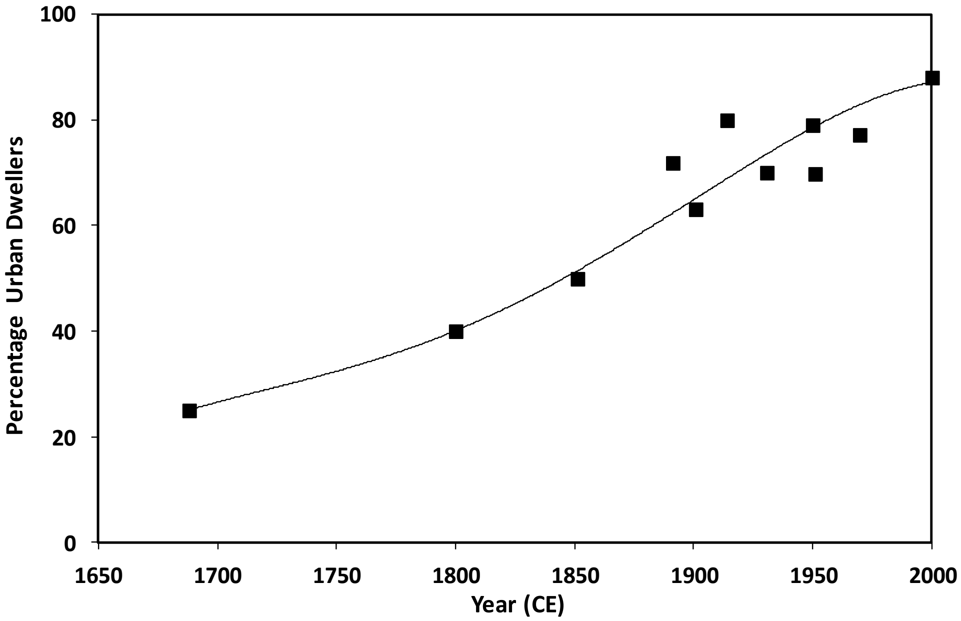

FIGURE 3. Percentage of British population that were urban dwellers with a polynomial trend line.

Note that each of the 99 values of *f%* in Table 2 has attached to it a year, a weighting for location or class, the number of records, and a merit ranking, forming one unit of datum.

Analytical methods. Into how many time periods should the data in Table 2 be partitioned, and further, how many values of *f%* are required in each time period (and marriage relationship), for a 'reliable' average? If the data in Table 2 are treated as a single time period, there are 99 *f%* values to average. If the data in Table 2 are partitioned into two time periods, there are about 50 *f%* values in each time period. If partitioned into three time periods there are about 33 *f%* values, for four time periods about 25 *f%* values, and for five time periods about 20 *f%* values. As the *f%* values were further distributed among the five marriage relationships it was considered that five time periods was the maximum number to provide 'reasonable' averages. These five partitioning possibilities are shown in Table 3. To further partition the *f%* values into time periods it was considered that as far as possible, each time period should have about the same number of *f%* values with no splitting of runs of the same year, i.e. all values of *f%* associated with the same year would be placed into the same time period. Two methods were used for this purpose. The first method kept the various *f%* values aligned in the same row by author, as displayed in Table 2, and partitioned on the years associated with the 1st cousins only. While the second method did not keep the various *f%* values aligned in the same row by author, and partitioning was undertaken on the years associated with each of the individual *f*% values whatever the cousin relationship. Table 3 shows the year periods and numbers of *f%* values in each year period, for each of these two methods of partitioning, aligned (a) and non-aligned (b), used in this analysis. Note that if both methods (a) and (b) have the same number of *f%* values in corresponding rows in Table 3, the actual *f%* values may be different.

TABLE 3. Temporal partitioning of the source data

	(a) Aligned by 1st cousins		(b) Non-aligned by 1st cousins	
Number of Periods	Years of Periods	Number of *f%* values	Years of Periods	Number of *f%* values
1	1790 to 1978	99	1756 to 1978	99
2	1903 to 1978	46	1879 to 1978	49
	1790 to 1902	53	1756 to 1878	50
3	1925 to 1978	35	1897 to 1978	33
	1880 to 1924	34	1853 to 1896	33
	1790 to 1879	30	1756 to 1852	33
4	1930 to 1978	26	1914 to 1978	24
	1903 to 1929	20	1879 to 1913	25
	1880 to 1902	23	1839 to 1878	25
	1790 to 1879	30	1756 to 1838	25
5	1936 to 1978	20	1925 to 1978	20
	1906 to 1935	19	1887 to 1924	20
	1890 to 1905	21	1865 to 1886	19
	1865 to 1889	19	1835 to 1864	20
	1790 to 1864	20	1756 to 1834	20

Using Table 2 as a template, the weighted means of the percentage rates of occurrence for each type of marriage relationship were calculated, weighted by the three factors simultaneously, and the weighted F% value was calculated using equation 1. For both methods of partitioning, the years associated with these values for the weighted F% were the weighted mean years for the 1st cousin marriages. Each partitioning method produced 15 values for the weighted F% with their associated weighted mean 1st cousin marriage years. This procedure was repeated in turn for the two partitioning methods, producing 30 values for the weighted mean F% with their associated weighted mean 1st cousin marriage years. All of these calculations were undertaken on Microsoft Excel spreadsheets.

Results

The 15 aligned and 15 non-aligned weighted F% values are plotted in Figure 4 using different markers for the aligned and non-aligned points. The sigmoidal nature of the distribution of the plotted weighted F% points is clear, and has a non-linear correlation coefficient of $\eta = 0.974$. This value is clearly better than $\eta = 0.517$ obtained for the unweighted F% results. A generalized logistic function in the form:

$$\mathrm{F\%}(t) = \left[A + \frac{K-A}{\{1+Q.exp[B.(t-C)]\}^{1/v}}\right] \qquad \text{..........} \ (2)$$

was fitted by the method of least-squares to the data points in Figure 4, where K is the upper bound value and A is the lower bound value of F%(t), Q, B , C and v are constants that determine the location and rate of the transition between K and A, and t is time in years. The generalized reduced gradient (GRG2) non-linear optimization code included in Microsoft Excel was used to fit this function.

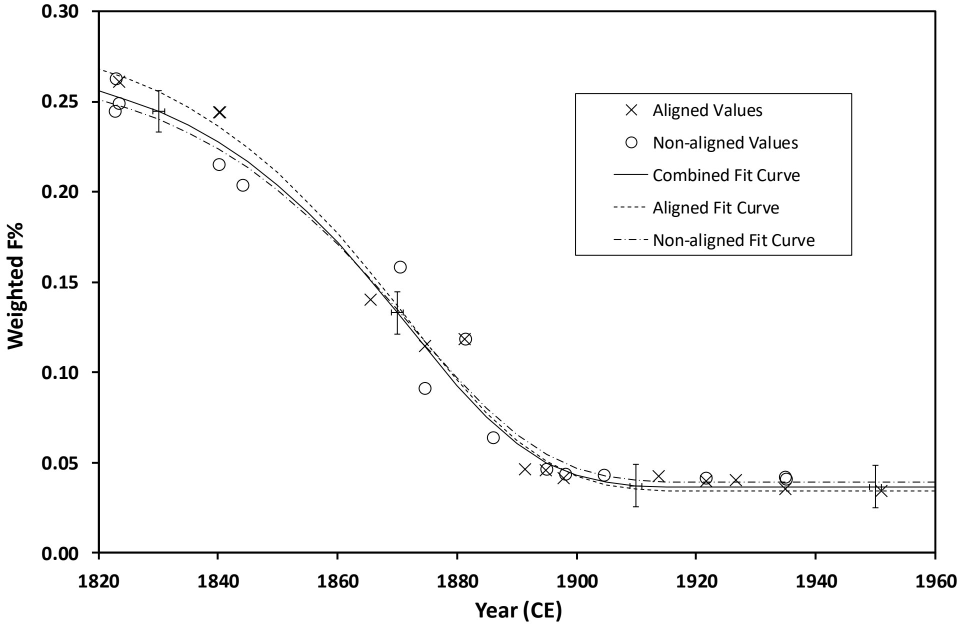

Figure 4. Combined values of weighted F% against year, distinguishing between the data points that were obtained by the aligned and non-aligned methods with generalized logistic curves fitted to all of the points combined, and to the aligned points and to the non-aligned points separately.

A generalized logistic function was first fitted to all 30 plotted weighted F% values and the resulting curve is also shown in Figure 4 as the continuous curve. The standard deviation of the spread of all 30 of the plotted weighted F% values about the fitted curve is ±0.012, which is shown as the vertical 'error' bars about the fitted curve. The horizontal 'error' bars have no significance. The weighted F% values obtained with the two partitioning methods, distinguished by different data markers in Figure 4, show no clear difference in their distributions about this fitted curve. For the combined results, at 1820 the weighted F% = 0.256±0.011, but is still varying slowly. The upper bound value, *K*, before 1820, from the fitted combined logistic function is 0.276. The post-transition period, after about 1900, where the fitted curve is essentially constant, the weighted F% = 0.038±0.004.

The procedure used to examine the combined results was repeated separately for the aligned and non-aligned results, each with 15 data points. Taken separately, the non-linear correlation coefficients η for the aligned and non-aligned weighted F% values are 0.977 and 0.971, respectively. From these two η values, the aligned partitioning method appears to give the more consistent distribution of values independently of any fitted curve. Generalized logistic functions were fitted separately to the weighted F% values for the aligned and non-aligned results and are shown as dashed curves in Figure 4. The standard deviations of the spread of the weighted F% values for the aligned and non-aligned weighted F% values are ±0.010 and ±0.013, respectively about their own fitted curves. These two values of standard deviation further indicates that the aligned method of partitioning the *f%* values produces the slightly smaller distribution of weighted F% values, this time about a fitted curve. However, as seen in Figure 4, the difference between these two curves is small, and they are at all times within ±1 standard deviation about the combined curve which was produced by the 30 plotted F% values. Since there was no obvious preference for either of the two partitioning methods used, it was considered that the logistic curve fitted to the combined aligned and non-aligned results, with its slightly larger spread of data points, was the best obtainable at this time.

Discussion

The change in the near random distribution of the 35 unweighted F% values, with a slight negative trend, in Figure 1 to the closely sigmoidal distribution of the 30 weighted F% values in Figure 4 is dramatic. A comparison of the logistic curves in Figure 4 with the scatter of the unweighted F% values plotted in Figure 1 shows that the fitted curve is not unreasonable. Unfortunately no other results for all of Britain are available to check the veracity of the results obtained in this study. It is of limited usefulness to compare the results of weighted mean F% obtained in this study for all Britain with the results of unweighted F% obtained for small regions of Britain.

The results reported in Day, and partly repeated in Day & Smith, suggest that a peak in both 1st and 2nd cousin *f*% values occurred in the mid-1800s. However, looking at the values of both 1st and 2nd cousin *f*% values given in Table 2, and of the values of unweighted F% calculated from them in Figure 1, of the five studies that produced a range of values of *f*% more or less over the 1800s, including Brennan, Bell, Day, Darwin and Robinson, only Day obtained such an apparent peak. The main reason for the claim for a peak appears to be the single value of unweighted F% in Figure 1 at (1812, 0.0501), which appears to be an outlier point from the general trend. This value may be correct for the two villages studied by Day, or may be an artifact due to incomplete or incorrect data collection. There are three other outliers, namely for Pearson at (1880, 0.329), and for Robinson at (1903, 0.362) and (1970, 0.284), that may similarly be correct for the populations and periods studied or may also be artifacts. Note that the two high Robinson results were obtained from dispensation records and would have been expected to be low rather than high. It is also noted that Darwin's *f*% results are often criticized for being high, being between 2 and 4%, but a number of other studies, including Day, Bramwell and Pearson, have produced values of *f*% in the same range. Importantly, these apparent outliers do not appear to have caused any problems in the analysis undertaken in this study.

As mentioned earlier, each author commented on the difficulties and limitations encountered in their studies, so there is no reason to assume that any of the *f*% values are significantly better or worse that the other values, other than by the method used to obtain the *f*% values, which is taken into account in the present study by the different merit rankings. It is expected that there will be local variation in the consanguinity rates according to cultural practice and population structure (Smith). Values of the percentage inbreeding coefficient F% obtained for small isolates are an unreliable guide to the average value for the whole country in which they are located. As an example, consider Switzerland. In Lasker (1985) there are three reports of inbreeding coefficients F determined by pedigrees, for the isolates Saas, Kippel and Torbel, and a report of an inbreeding coefficient, also determined by pedigrees, for all Switzerland. The average value for the three isolates is F = 0.00644 (F% = 0.644), while the value for all Switzerland is F = 0.00062 (F% = 0.062), a factor close to 10 less.

The downwards trend in unweighted F% shown in Figure 1 by the trend lines is not totally unexpected. There are numerous anecdotal reports that intermarriage among relatives had occurred in villages for centuries due, amongst other reasons, to the restricted movement to and from manor farms and villages during and after the medieval period when most people were tied to the land. For instance, it is reported amongst other anecdotes in Conniff (2003), that according to Professor Robin Fox of Rutgers University "it is likely that 80% of all marriages in history may have been between second cousins or closer". More substantial evidence comes from studies modelling the 'genealogy paradox', which is described in many books and journals (e.g. Bramwell; Jacquard 1974; Shoumatoff 1985; and Wachter). There have been a number of different approaches to modelling the pedigree collapse involved in the solution to the 'paradox' (Chang 1999; Derrida et al. 1999, 2000a,b; Pattison 2001, 2003, 2004, 2007; Wachter) and each has found that there must have been an increasing trend in inbreeding as one moves back in time.

The population of the British country side had been slowly declining since the Middle Ages, but a rapid decline occurred from the mid-1700s and through the 1800s, see Figure 3, in both England due to enclosures (Thompson 1991) and Scotland due to clearances (Richards, 2008), and the introduction of machinery to replace agricultural workers. Over the same period there was a decline in village cottage industries such as spinning and handloom weaving. Between 1793 and 1815 there was also the disruption of the population caused by the largest mobilization of naval and military forces in British history in response to the French Revolutionary and Napoleonic wars, and the depression and increased emigration in its aftermath (Colley 2005). The improvement in roads and public transport during the late 1700s and the early 1800s made movement easier (Harrison 1995; Pooley & Turnbull 1996). People were on the move looking for employment in the new factories, houses and shops of the growing middle class, see Figure 2, and new places to live, and in the process separating from their cousins and being increasingly exposed to unrelated potential marriage partners. The improved public health during this period, with lower infant mortality rates and longer life spans, led to larger families and possibly an increase in the number of cousins one had. But this may have been compensated by the cousins being further apart geographically and socially. It was clearly a turbulent time in Britain and it is not clear from studies to date what the trend was over all Britain with regards cousin marriages during the 19th century as studies of small regions do not show a consistent trend: some regions and villages were more affected by the above changes than others. This study has attempted to resolve this uncertainty.

Conclusion

As the first attempt to integrate the various disparate values of F% for Britons of British descent, for all of Britain, the results of this analysis are promising, and should prove useful in other related studies. The difference between Figures 1 and 4 is dramatic, and the results of this study show a clear decreasing trend in cousin marriages over the whole of the 19th century. The use of the three weighting factors and the two partitioning methods used in this study appear to be appropriate. The exact values for the weighting factors or the particular partitioning method used do not appear to be critical. Incorporating the results of further localized studies, similar to the detailed study by Day, providing more *f%* values for more areas of Britain would be beneficial to the analysis undertaken in this study.

Received 27 May 2016; accepted 10 August 2016.

Literature Cited

Bell, J. 1940. A determination of the consanguinity rate in the general hospital population of England and Wales. *Ann. Eugen.* 10:370-391.

Bramwell, B. S. 1939. Frequency of cousin marriages. *Genealogist's Mag.* 8:305-316.

Brennan, E. R. 1981. Kinship, demographic, social, and geographic characteristics of mate choice in Sanday, Orkney Islands, Scotland. *Am. J. Phys. Anthrop.* 55:129-138.

Bundey, S., H. Alam, A. Kaur et al. 1990. Race, consanguinity and social features in Birmingham babies: a basis for prospective study. *J. Epidem. Comm. Health.* 44:130-135.

Chang, J. T. 1999. Recent common ancestors of all present-day individuals. *Adv. Appl. Probab.* 31:1002-1026.

Coleman, D. 1980. A note on the frequency of consanguineous marriages in Reading, England in 1972/73. *Hum. Hered.* 30:278-285.

Coleman, D., and J. Salt. 1992. *The British Population: Patterns, Trends and Processes.* Oxford, UK: Oxford University Press.

Colley, L. 2005. *Britons: Forging the Nation, 1707 – 1837.* New Haven: Yale University Press.

Conniff, R. 2003. *Go Ahead, Kiss Your Cousin*. http://discovermagazine.com/2003/aug/feastkiss (viewed 26/5/2016).

Crow, J. F. 1980. The estimation of inbreeding from isonymy. *Hum. Biol.* 52:1-4.

Crow, J. F. 1989. Update to the estimation of inbreeding from isonymy. *Hum. Biol.* 61:949-954.

Crow, J. F., and J. F. Mange. 1965. Measurement of inbreeding from the frequency of marriages between persons of the same surname. *Eugenics Q.* 12:199–203

Darwin, G. H. 1875. Marriages between first cousins in England & their effects. *J. Stat. Soc. Lond.* 38:153-182.

Davidoff, L., and C. Hall. 1987. *Family Fortunes: Men and Women of the English Middle Class 1780 – 1850.* London, UK: Hutchinson.

Day, C. L. 2010. *Marriage Patterns in Two Wiltshire Parishes 1754-1914: Geographical Mobility, Consanguinity and Illegitimacy*, PhD thesis, Australian National University, Canberra, Australia.

Day, C., and M. Smith. 2013. Cousin marriage in south-western England in the nineteenth century. *J. Biosoc. Sci.* 45:405-414.

Derrida, B., S. C. Manrubia, and D. H. Zanette. 1999. Statistical properties of genealogical trees. *Phys. Rev. Lett.* 82:1987-1990.

Derrida, B., S. C. Manrubia, and D. H. Zanette. 2000a. On the genealogy of a population of biparental individuals. *J. Theor. Biol.* 203:303-315.

Derrida, B., S. C. Manrubia, and D. H. Zanette. 2000b. Distributions of repetitions of ancestors in genealogical trees. *Physica A.* 281:1-16.

Eversley, D. E. C., P. Laslett, and E. A. Wrigley. 1966. *An Introduction to English Historical Demography: From the Sixteenth to the Nineteenth Century*. London, UK: Weidenfeld & Nicolson.

Harrison, G. A. 1995. *The Human Biology of the English Village.* Oxford, UK: Oxford University Press.

Harrison, J. F. C. 1984. *The Common People: A History from the Norman Conquest to the Present.* London, UK: Fontana Books.

Hasan, A. R., J. E. Pattison, and A. Hariz. 2010. Preprocessing of Affymetrix gene chip microarray data. *Curr. Bioinformatics*. 5:270-279.

Heyck, T. W. 1992a. *The Peoples of the British Isles: A New History*, vol. 2, From 1688 to 1870. Belmont, CA: Wadsworth.

Heyck, T. W. 1992b. *The Peoples of the British Isles: A New History*, vol. 3, From 1870 to the Present. Belmont, CA: Wadsworth.

Hinkle, D. E., W. Wiersma, and S. G. Jurs. 1988. *Applied Statistics for the Behavioral Sciences.* 2nd ed. Boston, USA: Houghton Miffling.

Imaizumi, Y. 1974. Genetic structure in the United Kingdom. *Hum. Hered.* 24:151-159.

Jacquard, A. 1974. *The Genetic Structure of Populations.* New York: Springer.

Lasker, G. W. 1985. *Surnames and Genetic Structure.* Cambridge, UK: Cambridge University Press.

Mitchell, A. 1866. Blood-relationship in marriage considered in its influence upon the offspring. *Memoirs of the Anthropological Society of London* 2:402-456.

Morton, N. E., C. Smith, R. Hill et al. 1976. Population structure of Barra (Outer Hebrides). *Ann. Hum. Genet.* 39:339-352.

Pattison, J. E. 2001. New method of estimating inbreeding in large, semi-isolated populations with application to historic Britain. *J. Comp. Hum. Biol. (Homo)* 52:117-134.

Pattison, J. E. 2003. The effect of the bubonic plague epidemic on inbreeding in 14th C Britain. *Am. J. Hum. Biol.* 15:101-111.

Pattison, J. E. 2004. A comparison of inbreeding rates in India, Japan, Europe and China over historic times. *Homo: J. Comp. Hum. Biol.* 55:113-128.

Pattison, J. E. 2007. Estimating inbreeding in large semi-isolated populations: Effects of varying generation length and of migration. *Am. J. Hum. Biol.* 19:495-510.

Pattison, J. E. 2011. Integration versus apartheid in post-Roman Britain: A response to Thomas *et al.* (2008). *Hum. Biol.* 83:715-733.

Pearson, K. 1908a. The proportion of cousin marriages in the general population. *Brit. Med. J.* i:1076.

Pearson, K. 1908b. Cousin marriages. *Brit. Med. J.* i:1207.

Pearson, K. 1908c. Cousin marriages. *Brit. Med. J.* i:1395.

Pooley, C. G., and J. Turnbull. 1996. Migration trends in British rural areas from the 18th to the 20th centuries. *Int. J. Popul. Geogr.* 2:215-237.

Richards, E. 2008. The Highland Clearances: People, Landlords and Rural Turmoil. Edinburgh, UK: Birlinn.

Robinson, A. P. 1983. Inbreeding as measured by dispensations and isonymy on a small Hebridean Island, Eriskay. *Hum. Biol.* 55:289-295.

Rogers, L. A. 1987. Concordance of isonymy and pedigree measures of inbreeding: The effects of sample composition. *Hum. Biol.* 59: 753-767.

Sheets, J. W. 1981. Population structure of depreciated communities: I. The 1977 genetic demographies of Colonsay and Jura islands, the Scottish Inner Hebrides. *Soc. Biol.* 27:114-129.

Shoumatoff, A. 1985. *The Mountain of Names: A History of the Human Family*. New York: Simon & Schuster.

Smith, M. T. 2001. Estimates of cousin marriage and mean inbreeding in the United Kingdom from 'birth briefs'. *J. Biosoc. Sci.* 33:55-66.

Smith, M. T., R. M. Asquith-Charlton, L. M. Blodwell, et al. 1993. Estimating inbreeding from the Faculty Office Registers, 1534-1540. *Ann. Hum. Biol.* 20:357-368.

Thompson, E. P. 1991. *The Making of the English Working Class*. Harmond-worth, UK: Penguin.

Wachter, K. W. 1980. Ancestors at the Norman Conquest. In: Dyke, B., Morrill, W. (eds), *Genealogical Demography*. New York: Academic Press. p 85-93.