Sources and methods

This item considers a range of sources and methods commonly used in local population history. These vary in sophistication and complexity, but are intended to be of benefit to the broad LPS readership, and are accompanied by worked examples. Each item is written by an experienced population history practitioner, and will usually address both the possibilities and the pitfalls of the respective sources and methods under discussion. The LPS Board are happy to enter into correspondence on this item, which should be addressed in the first instance to the LPS General Office.

Infant mortality

Chris Galley

Infancy covers the first year of life (ages 0–1) and the infant mortality rate (IMR) measures the proportion of infants who do not survive their first year. The first formal definition of the IMR was published by William Farr in the Registrar General's *Annual Report for 1875* and by the end of the nineteenth century it had become adopted as a primary demographic indicator.[1] In 1938 even Neville Chamberlain, the then Prime Minister, was conversant with it, noting that 'sometimes we take the infant mortality rate as a sort of pointer to show how health is improving'.[2] Many contemporary publications, such as UNICEF's *The state of the world's children*, continue to use the IMR as a surrogate measure of well-being and infant mortality has been, and still is, the focus of much research aimed at lowering rates.[3] The literature on infant mortality is vast with both contemporary and historical demographers having expended much effort in seeking to explain social, spatial and temporal variations in IMRs. This article does not seek to engage in any of the debates concerning how and why IMRs in the past may have changed; instead it aims to explain how the rate may be calculated using historical population sources. In many publications, including from 1875 the majority produced by the General Register Office, the IMR is simply reported. When it is not, the rate can be calculated relatively easily, but care needs to be taken in order to ensure that the correct formula is used.[4]

1 HMSO, *Thirty-eighth annual report of the Registrar General of births, deaths and marriages in England* (London, 1878), xl.

2 Quoted by Richard Titmuss, *Population and poverty* (London, 1938), 77.

3 The latest volume, UNICEF, *The state of the world's children 2009* (UNICEF, 2009) discusses maternal and newborn health. As with previous issues, the IMR is used to assess demographic progress.

4 For examples of how the use of incorrect formulas have led to a poor understanding of the processes affecting infant mortality and a short discussion of how the IMR became adopted see C. Galley, 'George Newman—A life in public health', in E. Garrett *et al*. eds, *Infant mortality. A continuing social problem* (London, 2006), 23–6.

The IMR is one of many age-specific mortality rates and represents, within a given population, the probability of a live birth surviving to age one. Conventionally, an age-specific mortality rate is calculated by dividing the number of deaths at a given age by the population of that age at the mid-year point.[5] This implies that in order to calculate an IMR it is necessary to divide the number of infant deaths by the population of infants:

$$\text{IMR} = \frac{\text{number of infant deaths}}{\text{mid-year population of infants}}$$

Unfortunately, unlike most other age groups, which experience a similar risk of dying throughout the year, infants experience far higher risks close to birth (for example, in 1906, 49 per cent of infant deaths in England and Wales were aged 0-3 months, 21 per cent aged 3-6 months and 30 per cent aged 6-12 months).[6] This means that a mid-year count of infants will not necessarily represent the true population at risk. Instead, a better way to represent the population at risk is to use the number of live births within the year. The conventional way of calculating the annual IMR is therefore to divide the annual number of infant deaths by the annual number of live births. Thus:

$$\text{IMR} = \frac{\text{infant deaths}}{\text{live births}}$$

By convention the IMR is usually expressed as the number of infant deaths per 1,000 live births.

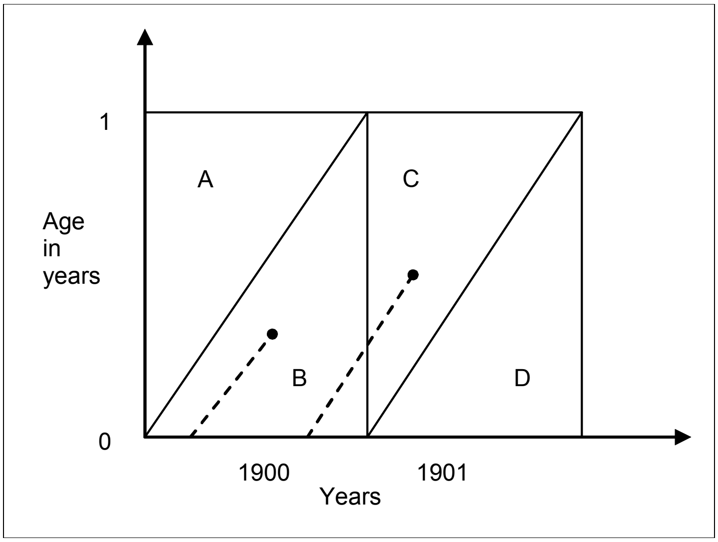
The above formula is usually used to calculate annual IMRs, but it can also be used for longer periods (five years—quinquennia—for instance). It is less useful for periods shorter than one year because in many historical populations the seasonality of both births and infant deaths could be substantial, and such fluctuations may cause the births not to be representative of the population at risk. For example, in 1911 the hot summer caused many infants to die from diarrhoea during the third quarter of the year with infants aged over three months being at greatest risk.[7] If a third quarter IMR is calculated for this year, then the majority of infant deaths in this period will have been born in previous quarters. This example illustrates a wider problem of using the above formula to calculate the IMR: within any year, some infants, especially those born towards the end of the year, will die in the following year and hence the numerator and denominator in the infant mortality calculations do not refer to exactly the same populations. Figure 1, a Lexis chart, illustrates this problem. The dashed lines represent the life courses of individual infants. In the first example the infant is born towards the beginning of 1900, but dies mid-year. The second infant is born towards the end of 1900 and dies in 1901. If it is possible to trace all the births in 1900 throughout their first year then the proportion that did not survive infancy can be determined. This method

5 A. Hinde, *Demographic methods* (London, 1998), 9–12, discusses age-specific mortality rates. Strictly speaking the mid-year point is 2 July, although 30 June is often used.

6 HMSO, *Sixty-ninth annual report of the Registrar General of births, deaths, and marriages in England and Wales (1906*) (London, 1908), 296.

7 HMSO, *Seventy-fourth annual report of the Registrar General of births, deaths, and marriages in England and Wales (1911*) (London, 1913), 71.

Figure 1 Cohort and period IMRs



requires some form of nominal record linkage procedure such as family reconstitution to be carried out.[8] In Figure 1, if the deaths in the parallelogram B-C can be related to the births in 1900 then this measure is known as the cohort IMR. To get a better indication of the risks to infants for periods of less than one year it is always better to use a cohort rate, but in many instances limitations in the data may not allow this measure to be calculated. Figure 1 also illustrates the conventional method of calculating the IMR described above—in this instance, the infant deaths in the rectangle A-B are related to the births in 1900 and this measure is known as the period IMR.

In practice there is likely to be little difference between period and cohort IMRs and in the vast majority of cases the period rate will be used because it is much easier to calculate. For example, as part of a larger study of York's population I made calculations of IMRs (cohort rates) throughout the city between 1570 and 1800. This took me the equivalent of one year's full-time work. Later when I decided to extend the series using published data, it took less than an afternoon to produce a series of IMRs (period rates) from 1837, the start of Civil Registration, until 1950. Unless otherwise mentioned, all reported IMRs are likely to be period rates and in most cases these should adequately reflect the life chances of the infant population. In exceptional circumstances, such as when the birth rate changes abruptly, a different formula may be deemed necessary: in some cases it may be appropriate to use the average of the births in two consecutive years to represent the population at risk. Corrections were also needed during the First World War when the number of births fell as many men left home to fight in the war.[9] Such minor adjustments are not usually necessary and most IMRs are calculated simply by dividing infant deaths by the number of live births.

8 Family reconstitution will the subject of a future Sources and Methods item.

9 W. Martin, 'The estimation of infant mortality', *Journal of the Royal Statistical Society*, 96 (1933), 482.

Table 1 **The effects of under- and over-registration on the accuracy of the infant mortality rate (infant mortality rate = 100 if registration is perfect)**

		Misreported births (%)										
		+25	**+20**	**+15**	**+10**	**+5**	**0**	**-5**	**-10**	**-15**	**-20**	**-25**
Misreported deaths (%)	**+25**	***100***	96	92	88	84	80	76	72	68	64	60
	+20	104	***100***	96	92	88	83	79	75	71	67	63
	+15	109	104	***100***	96	91	87	83	78	74	70	65
	+10	114	109	105	***100***	95	91	86	82	77	73	68
	+5	119	114	110	105	***100***	95	90	86	81	76	71
	0	125	120	116	114	105	***100***	95	90	85	80	75
	-5	132	126	122	117	111	106	***100***	95	89	84	79
	-10	139	133	129	124	118	112	106	***100***	94	89	83
	-15	147	141	138	131	125	119	113	106	***100***	94	88
	-20	156	150	144	138	131	125	119	113	106	***100***	94
	-25	167	160	153	147	140	133	127	120	113	107	***100***

Any infant mortality calculation is, of course, dependent on the quality of the source material and for IMRs to be accurate it is necessary that: (1) all births and deaths are reported; and (2) accurate ages at death are provided or can be calculated easily. In any registration system, especially an historical one, perfection is unlikely, but both national and local estimates of IMRs have been made for England stretching back to the sixteenth century.[10] Before discussing the methods used to derive these estimates it is necessary to be aware of the problems that may result if the IMR is calculated using inaccurate data. Table 1 shows the effects of both under- and over-registration on the accuracy of calculated IMRs. If registration is perfect, then the IMR within this population will be 100 (the central cell of the table). Likewise, if the proportion of births and infant deaths is either under- or over-represented by the same amount, then the calculated IMR will also remain at 100 (along the diagonal of the table). In practice, in some historical populations under-registration may occur because certain groups such as nonconformists or Catholics may not register their vital events in Anglican parish registers.[11] Provided that *all* births and infant deaths from such groups are excluded from the register, the IMR will be unaffected, although the calculated IMR will only be representative of the whole population if rates within both included and excluded groups are similar. Table 1 illustrates how serious problems will arise if different degrees of under- or over-recording takes place. This problem can be illustrated by a simple example. In the parish of Ashton-under-Lyne, Lancashire, as a consequence of population pressure, work began on a new church, St Peter's. The building was consecrated in 1824 and a new register started. The ages of infant burials were given in months and the period IMR between 1824 and 1836 was

10 E. A. Wrigley *et al.*, *English population history from family reconstitution 1580-1837* (Cambridge, 1997), 224.

11 M. Slack, 'Non-conformist and Anglican registration in the Halifax area 1740-99', *Local Population Studies*, 38 (1987), 44–5, showed that in Halifax during the 1790s nonconformist registers recorded 47 per cent of baptisms in the town, but only 16 per cent of burials.

calculated to be 841 infant deaths per 1,000 live births.[12] The explanation for this astonishingly high IMR is of course nothing to do with social conditions within the parish; instead it is related to registration practices. Between 1831 and 1836, annual numbers of baptisms were 24, 24, 29, 25 and 35 while annual numbers of burials were 66, 76, 84, 115, 132 and 148. Presumably with plenty of space in this new churchyard, burial plots were easy to obtain and some individuals buried their infants there even though they may have been baptised elsewhere.

After the commencement of Civil Registration in 1837 the calculation of IMRs is generally straightforward, although reliance on published data means, paradoxically, that it becomes difficult if not impossible to verify the accuracy of the data used. Before 1837 it is necessary to use ecclesiastical registers and in general these recorded baptisms and burials instead of births and deaths.[13] In most calculations of infant mortality using these sources, baptisms and burials have to be used as surrogates for births and deaths, and the one year period of observation will necessarily have to be based on the interval between baptism and burial. Many birth-baptism intervals were relatively short, although generally intervals began to lengthen from the mid eighteenth century.[14] In most cases this should not cause too many problems since the probability of dying declines considerably as infants grow older. The classic means of extracting age-specific demographic data from parish registers is via a technique known as family reconstitution which was devised by the French demographer Louis Henry and adapted for English registers by Tony Wrigley.[15] This technique involves the linking together of all events recorded in a parish register that relate to a particular marriage. Once sufficient families have been reconstituted inter-generational links are made and a wide variety of representative demographic rates can then be calculated, although results often have to be combined into cohorts of 25 years or more to ensure a sufficiently high population at risk. This technique yields robust demographic data, although it is imperative to ensure that the register is complete and contains sufficient detail over a long period of time so that correct links can be established.

Reconstitution represents a considerable investment of time and it is unlikely that the student of infant mortality will wish to undertake this procedure just to calculate IMRs. In order to speed up the linking process a simplified form of data linkage can be adopted. In this case it is not necessary to link baptisms and burials directly to a marriage; instead for each baptism the burial register is searched over the following year to determine if the infant died within

12 'Registers of St Peter's, Ashton-under-Lyne, St Peter's Blackley and St Lawrence's Chorley', *Lancashire Parish Register* Society, 127 (1990), 1-36.

13 For a short period during the Commonwealth period (1653 onwards) many registers recorded births alongside baptisms and a few deaths alongside burials. Quaker and some nonconformist registers consistently recorded births and deaths.

14 There is an extensive literature on the delay between birth and baptisms in Anglican registers: see *Local Population Studies*, 72 (2004), 83–7; 71 (2003), 81–7; 70 (2003), 49–56; 65 (2000), 9–28; 61 (1998), 57–9; 57 (1994), 72–5; 24 (1980); 11 (1973), 11–22.

15 E.A. Wrigley, 'Family reconstitution' in E. A. Wrigley ed., *An introduction to historical demography* (London, 1966), 96–159.

one year. In practice this procedure is easier to carry out if the method is reversed. Then events that specifically refer to adults can be automatically excluded from the linking process, which is just confined to those burials listed as 'son or daughter of'. This method works best if the register has been printed, since the baptism register can be photocopied and as the burial register is searched links can then be noted on the copy. Once the linking process has been completed rates can be calculated in the usual way. This method has a number of advantages over reconstitution. All baptisms are in observation—including illegitimates and migrant ones that may be omitted from a reconstitution study. Any short gap in the register, which would preclude a full reconstitution from taking place, will only result in a similar small gap in the series of calculated IMRs. Its major advantage over reconstitution is that it is much quicker to carry out and thus it enables a far greater range of registers to be included in any study. This method does, however, have one major disadvantage. Within any ecclesiastical registration system some infants will die before they can be baptised, and while obviously these infants will not appear in the baptism register there is a possibility that they may also be absent from the burial register since the unbaptised were not supposed to be interred in consecrated ground. However, many unbaptised infant burials were recorded in parish registers, and if they are consistently identified or age at death is given, then 'dummy' births for these infants can be created and cohort IMRs calculated. Only a small minority of parish registers recorded unbaptised infant burials, but those that did are likely to be among the most accurate of registers since if parish clerks took the trouble to record information that was not required it is likely that similar levels of care were taken to ensure that they accurately recorded the true numbers of baptisms and burials occurring within the parish. In practice once the linking procedure is underway deficiencies within the register often become immediately apparent, since one of the first signs of poor registration is that links between baptisms and burials become harder to establish.[16]

As awareness and concern about issues relating to infants increased, the various components of infant mortality began to be examined and this led to the introduction and use of additional measures. The most important ones are:

1) neonatal mortality—deaths in the first 28 days (or occasionally the first month);
2) post-neonatal mortality—infant deaths aged over 28 days (or one month);
3) stillbirth—the expulsion of a dead foetus that has reached the age that it is capable of independent survival (which is usually considered to be 28 weeks). The stillbirth rate is therefore the number of stillbirths divided by the total of live births and stillbirths;
4) perinatal mortality—sum of stillbirths and infants in the first week of life.

Other terms, such as 'early neonatal' (deaths in the first week), are sometimes employed and, as with all measures relating to infant mortality, it is important to refer back to the precise

16 If the simple linking procedure is undertaken on a register that does not specifically identify unbaptised infant burials then the resulting IMR will be under-recorded, although with further work it may be possible to provide correction factors. See Table 1 and the discussion in C. Galley, *The demography of early modern towns* (Liverpool, 1998), 92–100.

definition given in the text. There are two final terms that are often encountered in studies of infant mortality, endogenous and exogenous mortality. These result from the so-called biometric analysis of infant mortality. This technique, devised by the French demographer Bourgeois-Pichat, examined the distribution of infant deaths and attempted to differentiate, in the absence of any cause of death data, exogenous deaths (those associated with the external environment) from endogenous deaths (those associated with disorders inherited from the mother and birth injuries). Bourgeois-Pichat postulated that the cumulative IMR, and hence exogenous mortality, should be proportional to the function $\log^3(n+1)$ where n = age at death in days. Thus, by drawing a graph of this relationship it should be possible to split endogenous from exogenous mortality, since where the line of cumulative IMR crosses the y axis will give the endogenous IMR. In 1977, Wrigley adapted this technique to investigate the plausibility of IMRs produced via reconstitution and thereby establish the accuracy of parochial registration.[17] In many instances, when IMRs calculated from parish registers are subjected to this examination, the lines of cumulative infant mortality are straight and levels of endogenous mortality are well above zero. However, this technique is not subject to independent verification since detailed cause of death data for infants is not provided in parish registers and as a greater amount of work has been undertaken on this technique a significant number of exceptions have been found to this so-called 'rule'.[18]

Once the technicalities of measuring the IMR have been overcome, the challenge remains to seek to understand how and why rates varied between different groups and places. In England the national IMR was *circa* 175 per 1,000 live births in 1581, it increased to *circa* 200 by the late seventeenth century and then decreased to *circa* 150 by the early nineteenth century.[19] Throughout the nineteenth century the national rate in England and Wales remained relatively stable. It was 151 in 1901 and then steadily declined, reaching 5 in 2001. These national averages do, however, mask considerable spatial variations across the country. During the eighteenth century the IMR was as low as *circa* 80 in some isolated rural parishes and perhaps over 350 in parts of eighteenth-century London.[20] These variations decreased during the late eighteenth and early nineteenth centuries, but even by the first decade of the twentieth century IMRs ranged from 47 to 204 within individual registration districts. Infants are among the most vulnerable members of all societies; they are often subject to higher risks of dying than other age groups and demographic change keenly affects this group. Moreover, with sources relevant to this subject being plentiful in archives throughout the country real progress will be made in understanding the processes of change if a greater number of local studies are undertaken.

17 J. Bourgeois-Pichat, 'La measure de la mortalité infantile', *Population*, 6 (1951), 233–48, 459–80; E. A. Wrigley, 'Births and baptisms: The use of Anglican baptism registers as a source of information about the numbers of births in England before the beginning of Civil Registration', *Population Studies*, 31 (1977), 281–312.

18 See the discussion in C. Galley and N. Shelton, 'Bridging the gap: determining long-term changes in infant mortality in pre-registration England and Wales', *Population Studies*, 55 (2001), 65–77.

19 Wrigley *et al*., *English population history*, 224.

20 R. Smith and J. Oeppen, 'Place and status as determinants of infant mortality in England c. 1550–1837', in E. Garrett *et al*. eds., *Infant mortality. A continuing social problem* (London, 2006), 63–8.