

## Babies' First Names: points to note about the lists

This document provides some points to note about the lists, including some information about how National Records of Scotland's (NRS's) computer program identified each baby's first forename (and about how it would be unable to do so correctly in a few unusual cases).

There is a separate list for each year for which the information is available (at the time of writing, this was for each year from 1974 onwards). The list for the latest year is scheduled to be added within three months of the end of the year – for example, the list for 2014 was added in March 2015.

**Please note:** the publication which appears in December each year gives figures for the latest year which are based only on births which were registered in the first eleven months of that year.

National Records of Scotland's statistics of births, and the lists of names, are produced from the records of the first registration of each birth, and so do not count any names that may be given when some births are later re-registered.

The total number of babies for all the names that are listed may be slightly less than the total number of births that were registered in the year concerned. The reason for this is that the lists exclude a very small percentage of cases for which the first forename was blank, due to no forenames having been recorded when a birth was first registered. That may have happened for a number of reasons: for example, perhaps the parent(s) had not yet have decided what to call the baby, or perhaps the baby had died soon after birth and the parent(s) did not want to give it any forenames.

Variants which are based on the same name are counted separately – for example, in these statistics, 'Ben' and 'Benjamin' are different names, likewise 'Agnes' and 'Senga', and so are different spellings like 'Stephen' and 'Steven'.

Accents have been ignored, so (for example) 'Chloe', 'Chloé', 'Chloë', 'Chloè' and 'Chlöe' are all counted as the same name: 'Chloe'.

Names were identified automatically by using a computer function which extracts (from the text in the 'forenames' field) sequences of characters which are 'delimited' by spaces (or by the start or the end of the field). This means that the computer function will count a sequence of non-blank characters which contains a hyphen (e.g. 'Mary-Frances') or an apostrophe (e.g. 'O'Brien') as a single name, and that it would count 'J' as the first forename if the forenames were recorded as 'J Arthur'. However, the computer program will count as two separate names any entry that consists of two sequences of non-blank characters which have a space between them: for example, it will regard 'St Clair' as two separate names, and count 'St' as the first forename; similarly, it will regard 'Da Silva' as two separate names, and count 'Da' as the first forename. It follows that the lists may include some entries which were only part of a baby's first forename, and that the lists will not give the full first forename in such cases. It is simply not feasible for National Records of Scotland (NRS) to scrutinise carefully all the babies' names that are registered in a year, in order to identify any forenames that consist of two (or more) separate words, with the aim of counting them correctly for the purpose of these statistics.

In the statistical database from which the lists were produced, all the names have been stored in wholly upper-case format (e.g. 'McKenzie' and 'Mckenzie' are both held as 'MCKENZIE'). When NRS produced the lists, it used a computer function to convert the 'upper case' version of each name into so-called 'proper case' format. The function will almost always give the correct result (e.g. it will convert 'MARY-FRANCES' to 'Mary-Frances'), but there are some unusual names which it may be unable to return to their original form. For example, all names which are held in the database as 'MCKENZIE' will be converted to 'Mckenzie': the function cannot convert some of them to 'McKenzie' and others to 'Mckenzie'. As a result, the lists will occasionally have a lower-case letter where it should be upper-case (e.g. a first forename of 'JK' would appear in these lists as 'Jk'). However, please note that the administrative system's record of every birth registration (from which any further copies of birth certificates will be produced) has the names exactly as they were originally recorded (i.e. with upper-case letters where the original name has upper-case letters).

Finally, it should be noted that the program automatically 'corrected' a few apparent 'typos' in the data that it extracted from NRS's computer system (for example, it changed any occurrences of the digits 0 and 1 to the letters O and I).

The lists for 1974 onwards were produced using the first name that was identified as having been recorded in the 'forenames' part of the entry for the initial registration of the birth in NRS's Vital Events statistical database of birth records. As the statistical computer record allows 30 characters for forenames, there should be no truncation of any first forename.

Because 1974 is the first year for which there are records in NRS's statistical database, any lists for 1855 to 1973 that may later be produced will have to use the first name that is identified in the 'forenames' part of the entry, for the initial registration of each birth, in the index to the Register of Births. Because the index entry records hold only the first 15 characters of the forenames, any first forename which was longer than that would be truncated to 15 characters.