Preparation of Data and Estimates for World Population Prospects 2021

Preliminary Draft Technical Documentation

Prepared by Peter D. Johnson for

United Nations Population Division

April 1, 2020

This document is still a preliminary draft. It will evolve over time as experience is gained on the performance of different methods and tools. It is being shared to allow other experts on the various parts of the process so that the UNPD can learn from their experiences.

Do not cite or quote without permission.

Contents

A. Introduction....5

B. Population....7

B.1. Introduction....7

B.2. Under- and over-enumeration....12

B.2.a. Introduction....12

B.2.b. Post-Enumeration Surveys and other matching studies....12

B.2.c. Other sources to assess completeness....16

B.2.d. Demographic methods to assess completeness of census data....17

B.3. Geographic and Definitional Coverage Issues....18

B.4. Developing a total population time series....21

B.5. Age misstatement of the population....22

B.5.a. Age Heaping....23

B.5.b. Age overstatement or understatement....28

B.6. Comparing population estimates by age and sex....29

C. Mortality....30

C.1. Overview....30

C.2. Fatal discontinuities....30

C.3. Sources of mortality data....31

C.3.a. Vital registration data....31

C.3.b. Census data....31

C.3.c. Survey data....32

C.4. Adult mortality....32

C.4.a. Direct estimates....32

C.4.b. Adjustment for age-misstatement....35

C.4.c. Indirect Estimates of Adult Mortality....36

C.4.d. Old age mortality....37

C.4.e. Validation and Review....37

C.5. Infant and Child mortality....37

C.5.a. Direct estimates....38

C.5.b. Indirect estimates....38

C.5.c. Consolidating infant and child mortality estimates....38

C.6. Full life table series 39
D. Fertility 41
D.1. Overview 41
D.2. Vital registration data 41
D.3. Census and survey data 41
D.3.a. Direct fertility data from censuses and surveys 41
D.3.b. Indirect methods based on data from censuses and surveys 42
D.4. Adjustments to estimates of fertility 43
D.4.a. Brass-type P/F ratio methods 43
D.4.b. Arriaga method 43
D.4.c. Relational Gompertz model 43
D.4.d. Additional methods for evaluating the completeness of birth registration 44
D.5. Review of results and produce final estimates 44
D.5.a. Estimate trends in the TFR 44
D.5.b. Estimate the trends in 5-year ASFRs 44
D.5.c. Estimate single year ASFRS 44
E. Migration 47
E.1. Overview 47
E.2. Data sources 47
E.2.a. Direct migration records 47
E.2.b. Refugees and asylum-seekers 47
E.2.c. Foreign-born or non-citizen populations 47
E.2.d. Residual migration estimates 48
E.3. Migration models 48
E.4. In-migration and out-migration vs. net migration 48
E.5. Short-term projections of migration 49
F. Produce Reconciled Estimates 1950-2020 50
F.1. Estimate 1950 base population 50
F.2. Population reconstruction 50
F.3. Combine 1950 base population with estimates of fertility, mortality, and migration 51
F.4. Compare estimates to unadjusted and adjusted/smoothed census or survey data 51
F.5. Determine if changes are needed 53

G. Specifications of the general analysis process, including documentation 54
H. REFERENCES 55
I. Appendix: List of Diagrams 63
I.1. Population 63
I.2. Adult Mortality 63
I.3. Child Mortality 63
I.4. Fertility 63

A. Introduction

Demographers and statisticians have been studying population for centuries now. The advent of regular censuses (or population registers), vital registration, and surveys has provided a wealth of data, especially in the twentieth century. All data collection systems have problems, so demographers have come up with different methods for estimating the underlying demographic variables. The problem for the demographer is to look at all the data and estimates and try to figure out what is really going on. I picture the problem as the "true" demographic variables are like an object in a box with some windows in it. Each window has a different perspective on the objects and some of the windows may be dirty or cloudy and give distorted views of the object inside. The demographer's job is to use all the information to try to get the best reconstruction of what is in the box.

One of the most basic foundations of demographic analysis it the demographic balancing equation:

P(t+1) = P(t) + B(t) - D(t) + NM(t)

Where:

P(t) = Population at time t

B(t) = Births in the year starting at time t

D(t) = Deaths in the year starting at time t

NM(t) = Net numbers of migrants in the year starting at time t

In the past, the process played out somewhat independently for the different components of change (fertility, mortality, and migration), and these could be put together to get revised populations by age and sex. With the advent of Bayesian approaches to prediction, the idea is to get both "best estimates" of the different parts of the process and measures of uncertainty, and then look at what the best estimates are that consider all parts interacting together with a resulting uncertainty.

So the first step for each component, somewhat independently, is to do the analysis to get the best estimate using their full range of data sources and methods. These alternate estimates can be used to estimate the uncertainty. The "somewhat independently" is because in order to estimate mortality or fertility rates one often needs population denominators (or vice versa), and in some cases indirect estimation methods rely on the enumerated population itself (e.g., intercensal survival, reverse survival).

For each component of the demographic model, we need to specify a list of sources from most reliable to least reliable in order to focus on the best estimates, but also see how well they are supported by other sources or deviate eventually in some systematic ways.

It would be useful to create a timeline for each country to document demographically relevant dates including:

1. Census dates
2. Survey dates
3. Dates of geographic changes
4. Date of Independence

5. Dates of significant demographic events
    a. Wars
    b. Civil strife
    c. Natural disasters
6. Dates of significant baby booms or busts
7. Dates of significant refugee flows

[Note: UNPD has been compiling part of this information in the context of its DemoData SQL database, and will consolidate further such information from multiple authoritative sources for this purpose.]

References will be made to workbooks from the **PAS** collection from the U.S. Census Bureau (2019b) as well as the unpublished **NewPAS** collection (U.S. Census Bureau, no date). Some of these tools are also implemented in the Demographic Analysis and Population Projection System (DAPPS) tool (U.S. Census Bureau, 2019a).

References will also be made to R functions from the **DemoTools** package (Riffe, 2019). It should be noted that the **DemoTools** package is under active development, so some of the function names and functionality may change over time.

B. Population

B.1. Introduction

The estimation and projection of the distribution of the world population by geographic areas (countries[1]) depends on having accurate information for each of these areas over time. The goal is to have each person alive at a given time to be counted once and only once in all the areas of study. Since the countries and areas do not all have census counts for the same date, it is possible that mobile populations may in fact be double counted, or missed altogether. An effort must be made to minimize that double counting or missing of selected populations, so the goal is that the populations that are modeled in the WPP are all *de-facto*. From the recommendations for the 2020 round of censuses (UNSD, 2017, p. 186) we have a simple definition of the terms (emphasis added):

> In the broadest sense, the total may comprise either all usual residents of the country or all persons present in the country at the time of the census. The total of all usual residents is generally referred to as the ***de-jure*** population and the total of all persons present as the ***de-facto*** population.

These concepts can have much more impact on the distribution by geography within the country, but for some countries the difference can be significant at the national level.

In countries with good vital registration (VR), the concept of "usual resident" population is increasingly preferred. In the European Union (EU) region with more open borders and population mobility, this concept is emerging as more fit for purpose, because it "offers better information for planning and policy purposes on the demand for services, households, families and internal migration" (UNSD, 2017, p. 178).

The "usual resident" is expected to be based in most instances on the presence in a particular place for most of the last 12 months (and usually, for at least six months and one day, see UNSD, 2017, p. 40 para 2.50 for further details). The approach is consistent with the Conference of European Statisticians Recommendations for the 2020 round of censuses, and Recommendations on Statistics of International Migration, Revision 1, Statistical Papers No. 58, Rev. 1, (United Nations publication, Sales No. E.98.XVII.14).

The UN also recommend that countries clearly state the enumeration status (with estimates if possible) of certain groups (UNSD, 2017, p. 187):

> 4.85. The groups to be considered are:
>
> > (a) Nomads;
> >
> > (b) Persons living in areas to which access is difficult;
> >
> > (c) Military, naval and diplomatic personnel and their families located outside the country;

[1] The term "country" as used in this report also refers, as appropriate, to territories or areas.

(d) Merchant seafarers and fishers resident in the country but at sea at the time of the census (including those who have no place of residence other than their quarters aboard ship);

(e) Civilian residents temporarily in another country as seasonal workers;

(f) Civilian residents who cross a border daily to work in another country;

(g) Civilian residents other than those in groups (c), (e) or (f) who are working in another country;

(h) Civilian residents other than those in groups (c), (d), (e), (f) or (g) who are temporarily absent from the country;

(i) Foreign military, naval and diplomatic personnel and their families located in the country;

(j) Civilian foreigners temporarily in the country as seasonal workers;

(k) Civilian foreigners who cross a frontier daily to work in the country;

(l) Civilian foreigners other than those in groups (i), (j) or (k) who are working in the country;

(m) Civilian foreigners other than those in groups (i), (j), (k) or (l) who are in the country temporarily;

(n) Refugees in camps;

(o) Transients on ships in harbour at the time of the census.

4.86. In the case of groups (h) and (m), it is recommended that an indication be given of the criteria used in determining that presence in, or absence from, the country is temporary.

The analyst should determine whether any of these categories are significant in the country being studied. Information on the enumeration of these groups is collected also through the UN Demographic Yearbook Questionnaires (UNSD, 2020) that are sent annually to countries.

The starting point of the WPP estimates will continue to be the base population by age and sex for 1950. At this point, however, it appears that there will be one or possibly two big differences from the past:

1. The population will be by single years of age[2]
2. The population may refer to January 1 rather than midyear. The use of January 1 populations works well with events and rates by civil calendar year.

Ideally, this population would be based on data from a census (or population register) in the year 1950. Unfortunately, only about 42 countries had a census in 1950. Therefore the 1950 population must be reconstructed using one or more earlier census(es) "projected" forward, or one or more later census(es)

[2] When the existing data is of insufficient quality, population by single age will be estimated from the population by five-year age group.

"projected" backwards to 1950. Alternatively, the population age distribution may be created using a stable or quasi-stable model.

The population age/sex distribution is a combination of three components:

1. The total population
2. The sex ratio of the pop (or total pop by sex)
3. The age distribution by sex

These parts can be estimated independently or together, but in the end they need to paint a consistent picture over age, sex, and time.

Some early attempts at census-taking involved very basic head count. Each area chief would assemble the village or area and literally count heads. These estimates would be passed up to the national level to get a total, with some geographic detail but no information on the age and sex distribution. These practices were common in the 1950s, 1960s with administrative censuses conducted in various African countries during the colonial period, and pre-independence. Lebanon did such listing for food distribution in 1942, but the real individual level census listing was in 1932 and the lists are still used today to issue administrative papers.

In the UNPD inventory of primary data collection operations (Data Catalog[3]), these more recent instances have been distinguished through the type of Data Process identified as "Household listing" which have been conducted in preparation for a full census (which didn't occur) or as master sample frame (Lebanon):

- Afghanistan 2003-2005 Household Listing
- Eritrea 2000 Quick Population Count
- Iraq 2009 Household Listing
- Lebanon 2004 Census of Buildings Dwellings and Establishments

Observed or estimated populations at later dates can be used as checks on the 1950 population as well as the reasonableness of the demographic components of population change up to the data on the population itself. Censuses from before 1950, if available, can be projected forward to 1950.

For all populations, the priority list of sources of population data by age and sex is generally:

1. Population register figures by single years of age
    a. From UNSD DYB based on data provided by national statistical authorities
    b. From country national statistical authorities
    c. From other sources
2. Census figures by single years of age
    a. From UNSD DYB based on data provided by national statistical authorities
    b. From country national statistical authorities
    c. From other sources
3. Population register figures by 5-year age groups
    a. From UNSD DYB based on data provided by national statistical authorities
    b. From country national statistical authorities

[3] https://population.un.org/DataArchiveWeb/index.html#/home

   c. From other sources
4. Census figures by 5-year age groups
   a. From UNSD DYB based on data provided by national statistical authorities
   b. From country national statistical authorities
   c. From other sources
5. Survey data on population by single years of age (usually as percents)
6. Survey data on population by 5-year age groups (usually as percents)

Note that for items 5 and 6, an estimate of the total population, by sex will be needed. January 1 population data from the Human Mortality Database (HMD) project can also be considered or used, but attention should be paid to the particular country to make sure that the population definition is consistent with that desired by the UNPD. For example, the U.S. population in the HMD (2016) excludes the armed forces overseas (AFO), Alaska, and Hawaii, and is not adjusted for underenumeration.

For each set of data, the proper reference date should be verified and documented. The reference date for a census is an important concept for demographers, but that concept is not always communicated well to the census workers and respondents. In many cases, the census data are released, but the reference date is not clearly indicated. In some cases there will be a discussion of when the field work was to be done (or the forms mailed out to be returned by mail). In some cases, the original census date was changed due to weather conditions, political upheaval, or funding problems. It is incumbent upon the analyst to review the available information and verify, as well as possible, the proper reference date for the census. Some records or databases (such as the DYB) may simply record the first day of the census enumeration period. Others may use the midpoint of that period. Translating that information to a consistent reference time can be challenging, but is recommended.

The definition and interpretation of the reference date of population must be made clear. For example, when HMD refers to "Population size on January 1st" it means the population at the beginning of that day (e.g., Jan 1 means at time 00:00 of that day). However, in Haiti 1971 census enumerators guide (IPUMS, 2019):

> Only those persons who are permanent residents (present or absent) on the official date of the census are to be enumerated. Thus, babies born after this date are not included, but persons who die after this date are included.
>
> Thus if the official date of the census is the 16th of August, 1971 at midnight, then a baby born the 17th of August is not enumerated, even if the household where this baby lives is not enumerated until the 30th of August. On the other hand, an elderly person who dies after the 16th will be enumerated.

I would interpret this reference date as actually August 17 at 00:00 of that day. Standards should also be set as to how to convert the census date to a year plus fraction. For example, Excel dates are interpreted as being at the end of the date given. Finally, a decision must be made about the denominator (365, or 366 if leap year, or 365.25 always?) to use when converting from days to fractions of a year, and how to deal with dates in leap years.

Once the data by age and sex have been identified, the first step in all protocols is to redistribute the unknowns by age (or by sex and age) in a *pro rata* manner. UNPD has been computing and storing such tabulations in its DemoData SQL as "Unk. Redist,", but care should be taken that the data do show the same totals as the original data. In recent years more effort has been made to impute age during the processing phase, so this has become less of an issue. If new data are being analyzed, an effort should be made to store and properly document the resulting redistributed population in a central location for further use.

There are really two problems with enumerated population counts by age and sex:

1. Undercounting or overcounting people
2. Having the wrong age reported by the respondents (or imputed)

The first problem can stem from many issues, including:

- Whole households that are missed
- People with multiple residences
- People who try to avoid contact with the gov't
- People who shouldn't be counted (according to the census rules)
- People who should be counted but are not

The second problem is the accurate recording of people's ages as of the census reference date. These can stem from:

- Rounding their ages (or year of birth) to end with the digit 0 or 5
- People overstating their age (e.g., in part to gain prestige or pensions)
- People understating their age (e.g., to avoid military service)
- Under- or over-stating ages of women depending on societal norms about child-bearing (under-stating ages of childless women or over-stating ages of women with higher than normal numbers of children)
- Ages being reported by respondents without full knowledge (e.g., relatives or neighbors)
- Problems with census forms or procedures that end up recording the age incorrectly. For example, in the 1950 US census "a few of the cards were punched one column to the right of the proper position in at least some columns. The result is that numbers reported in certain rare categories—very young widowers and divorcés, and male Indians 10–14 or 20–24—were greatly exaggerated. " (Coale and Stephan, 1962).

The demographer's task is to correct both of these problems. A Post-Enumeration Survey (PES) can address both issues to some extent, but the results tend to be very sensitive to the size of the samples used, the quality of the fieldwork and supervision, and various logistic and implementation challenges. Demographic analysis can also be used, trying to reconstruct the population based on registered births and deaths, estimates of net international migration, and possibly, earlier censuses.

B.2. Under- and over-enumeration

B.2.a. Introduction

Apart from age-misreporting, another problem with reported census data is completeness of enumeration including differentials by age and sex. The best way to do this is with a Post-Enumeration Survey (PES) that independently enumerates selected districts, and then those results are compared to the census (UNSD,_2010). The completeness of enumeration can also be estimated by "demographic analysis", or DA, and by comparisons of the census data to administrative data (U.S. Bureau of the Census_1985). The triangulation of as many independent sources of estimates as possible with population counts can serve to validate and estimate under- and/or over-enumeration in official figures (census/register/survey).

Is there evidence of correlation between age misreporting and under- or over- count in census? They both may be related to income and education. Lower SES populations are more likely to live in more crowded housing possibly leading to undercounts, and with less education they may be more likely to misreport age. High SES populations are more likely to have multiple residences, so the possibility of double counting, but in some developing countries access, participation and response rates may be more limited – especially among gated communities. Further research on this is needed.

B.2.b. Post-Enumeration Surveys and other matching studies

A post-enumeration survey is an independent survey taken soon after the census that covers a sample of the population and tries to match the respondents to the respondents in the census. The purposes of the PES are to estimate the coverage of the census (what proportion of the population was enumerated in the census), by age and sex, if possible, and to study the differences in the characteristics of the respondents in the two data collection operations. These characteristics include age. In theory, the PES survey should be undertaken as soon as possible following the census, and be conducted by the best interviewers and with more stringent instructions and methods for assessing responses like age.

A key part of the PES is the matching of respondents in the census and PES. The coverage of the census is usually estimated by looking at the population counted in the census but not the PES and vice versa in order to estimate the total missing from the census (Siegel and Swanson, 2004, p. 73). However, because the PES is a sample, the estimates have sample errors that give a range of possible values. Thus, there are competing desires for increasing the sample size to increase precision vs. keeping it smaller to decrease the cost. One of the main drawbacks of a PES is that it can be very expensive. When looking to a PES to estimate content error, Ewbanks (1981) says:

> Because of the high costs of the reinterviewing and matching, it is important to weigh these costs against the potential benefits. The values of the process will depend on whether the reinterview is designed solely to estimate the errors in the original interviews or whether it is designed to collect more detailed information to supplement the original survey.

Other matching studies can be done to check on the census data, including comparing registered births to the census (U.S. Bureau of the Census, 1953, see below) and to regular surveys, like the current Population Survey (CPS) in the U.S. (Siegel and Swanson, 2004).

Ewbanks (1981, p. 20) presents results of the relative net undercount by age and sex from seven post-enumeration studies (reproduced as Figure B-1 with color-coding). The results show a tendency for "severe undercounting of males aged 15-29 and females aged 15-24, the same groups whose ages might be misreported because of reliance on marital status and parity in estimating of ages. In most of the populations ... there was relative overcounting (that is, less undercounting) of both sexes over age 35." I was surprised that these results did not show consistent, high net undercount for the population under 5. It may be that the methods used in a PES are not as likely to detect undercount in this age group.

Figure B-1. Relative Omission Rates by Age and Sex from Ewbanks (1981)

TABLE 1. Estimated Proportions of Different Age-Sex Groups Omitted from Census Relative to the Overall Undercount: Seven Populations

	Korea	Korea	Liberia	Malaysia	Paraguay	U.S. 1960	
Sex and age	**1960**	**1970**	**1974**	**1970**	**1972**	**Whites**	**NonWhites**
Females							
0		6.6	3.6				
1-4		-0.8	5.9				
0-4	1.5	0.7	5.4	0.4	0.7	-0.4	-1.8
5-9	-1.5	-1.7	1.1	-0.4	-2.6	0.0	-3.6
10-14	-0.8	-1.8	-0.7	-0.2	1.1	-0.1	-4.2
15-19	3.3	3.7	3.8	0.6	4.0	0.8	2.2
20-24	4.1	4.6	-0.5	-0.3	7.4	0.8	1.6
25-29	0.2	0.9	-2.8	-0.1	1.2	-0.2	0.7
30-34	-1.9	-1.9	-2.5	-1.4	-1.6	-1.0	-2.4
35-39	-2.3	-2.2	-3.4	-1.8	-4.3	-1.8	-2.1
40-44	-2.9	-2.2	-5.3	-1.9	-3.1	-1.8	-1.8
45-49	-3.4	-2.0	-1.6	-1.1	-5.3	-0.9	0.3
50-54	-2.0	-1.3	-2.5	0.0	-2.7	2.6	11.0
55-59	-2.5	-2.3	-3.0	0.7	-2.3	0.0	2.1
60-64	-0.7	-0.4	3.8	1.1	-4.6	2.6	6.5
65+	0.0	4.8	-2.6	6.4	-0.3	0.5	4.5
Overall coverage	95.6	95.0	88.8	96.3	91.6	98.4	91.9
Males							
0		3.0	3.6				
1-4		-0.5	0.9				
0-4	0.6	0.3	1.3	-0.6	0.9	-0.8	-3.6
5-9	-1.6	-2.3	0.4	-1.1	-2.6	-0.4	-5.8
10-14	-2.2	-1.8	0.6	-0.7	-2.6	-0.3	-6.4
15-19	1.2	3.5	6.1	1.2	3.9	1.0	1.8
20-24	5.7	6.6	2.8	2.2	7.1	1.5	7.4
25-29	4.2	4.7	0.4	0.6	6.5	1.4	9.9
30-34	0.7	-1.1	0.1	0.8	0.9	0.3	8.0
35-39	-2.9	-1.6	0.7	-1.0	-1.5	-0.3	4.0
40-44	-4.1	-3.1	-3.5	-1.5	-3.2	-0.9	2.1
45-49	-2.6	-3.3	-3.6	-1.1	-2.6	-1.2	0.7
50-54	-4.1	-3.4	-5.2	-1.4	-6.1	0.8	7.7
55-59	-1.9	-2.4	-7.3	-0.4	-4.2	-2.4	-5.6
60-64	-3.3	-4.3	-3.7	0.4	-0.3	0.2	-1.3
65+	-0.9	8.9	-4.4	4.4	-3.9	1.0	-10.2
Overall coverage	94.7	95.0	89.2	95.6	90.6	97.2	89.1

Note: The table entries are 1 - [μ(x)/μ], where μ = overall coverage for the sex and μ(x) = coverage at age x. A positive value therefore indicates a relative undercount for the age group.

Source: Ewbanks, 1981, p. 20

[Action item: Compile a more complete set of PES estimates of net coverage errors by age to look at spatial and temporal patterns. This could also include separately identified estimates based on demographic analysis. – UNPD status as of mid-march 2020: an initial data compilation has been completed, and need to be further consolidated.]

PES's have been done in various forms in the US since 1950, but often the sample error is so high that it is hard to determine definitively whether there was an undercount or an overcount.

Age-specific under- or over-enumeration can significantly distort the age distribution of the population. A consistent pattern across many countries is the undercounting of young children. The Infant Enumeration Study 1950 (U.S. Census Bureau, 1953) looked at infants enumerated in the 1950 census to the birth registration data, and followed up on registered births that were not matched by a questionnaire sent to the parents recorded on the birth certificate. Estimates of the undercount for infants was estimated at four percent. What they found was (USBC, 1953, p. 2):

> In about 82 percent of the 16,045 cases in which infants were classified as definitely or probably missed in the census, the parents were also missed.

The reasons for non-reporting of the household were:

> (a) Absent or moved during the enumeration period;
> (b) ... enumerators overlooked some obscure buildings, failed to enumerate all dwelling units in a structure, or listed some occupied units as "vacant";
> (c) ... failure of relatives or nonrelatives to report the parents and the infant were staying with them, probably because the enumerator did not ask specifically about people who were living temporarily on the dwelling unit.
>
> In the 18 percent of cases where the parents were enumerated but the infant was missed, the reasons given ... included:
>
> (a) A neighbor gave incomplete information;
> (b) The family did not think infants were to be reported;
> (c) The infant died between April 1 and the time of the census enumeration.

In spite of the evidence of miscounting, the official estimates for the U.S. do not include adjustments for undercount.

Young adult enumeration has been found to be a problem in many countries (Ewbanks, 1981). This can result from the mobility of people in this age group due to attending school, finding a job, or getting married. After the 1960 U.S. census, administrative records checks of college students resulted in an estimated 2.5-2.7% undercount (Siegel and Swanson, 2004, p. 75). Note that more recently there have been issues of possible double counting of college students, both at their college location and their "home" address.

Comparisons of the 1960 census to Social Security records resulted in estimates of 5.1-5.7% undercount (Siegel and Swanson, 2004, p. 75).

Some countries (like Canada) do use the results of the PES to adjust their population estimates. It is important to read the metadata (which may include footnotes in tables) to check on whether estimates are adjusted or not. In the HMD estimates for Canada (2019a):

> Population estimates obtained from census enumerations are subject to undercount errors, which vary from census to census (typically 1–3% for recent censuses). Starting in 1971, Statistics Canada has been correcting the July 1st population estimates to adjust for the net undercount in the census. No such adjustments were applied to population estimates for the earlier years.

This does raise questions about whether pre-1971 populations should be adjusted to be consistent with later estimates.

PESs taken in India in 1981, 1991, and 2001 show relatively consistent patterns of net omissions by age, with high levels for the population under 5 and for young adults (males 15-39 and females 15-29 in 2001), and some indications of a rise in the oldest age groups (Gerland, 2013). These figures also show a slow rise in overall net omission rates over time from a minimum of 0.68% in 1961 to 2.33% in 2001.

B.2.c. Other sources to assess completeness

Other sources of data can be used to assess the quality of enumeration. These are mostly administrative sources that are collecting data for other purposes. These can include:

- Vital statistics from birth (and death) registration or from health authorities
- Education statistics (numbers of students by age): particularly relevant for primary and secondary age groups where/when school enrollment is quasi-universal
- Voter rolls for adult voting age groups
- Tax records for adults
- Immunization records, especially for infants under age 1 (e.g., DTP3)
- Records of issuance of IDs to residents
- Old-age pension and medical records

As noted above, registered births and deaths and records of migration flows can be combined to produce demographic estimates of expected populations (either total or by age and sex).

The use of other administrative data tends to be used to provide alternative estimates for particular age ranges of the population. As noted above, these records and data are collected for other purposes, so the first task is to locate these types of data. Some of these data may be found in statistical yearbooks produced by national governments. This may point to the part of the government that collects this data so that additional information and detail may be obtained from the particular government agency website. Other data may be aggregated by international organizations such as WHO, UNICEF, UNESCO, etc.

The second task is to determine what age group they relate to so that comparisons can be made. This can be part of the record for some of these sources (e.g., educational data or old-age pension and medical records), while others will not be as straightforward (e.g., national IDs, voter rolls and tax records). These will often provide lower bound estimates of the population in particular age groups since the administrative records may not cover everyone in the age group (e.g., non-voters, children not attending school, older people not covered by pensions or medical services).

In some cases administrative data can over-estimate a population. This can happen when looking at pension or old age medical records systems where individual records are not removed when the person dies. For example, Coale and Kisker (1990) found:

> For example, in January 1979 the [Medicare] rolls included 39 non-white females at ages 114 through 121, and three years later 33 were listed between the ages of 117 and 124. The deaths of only six persons out of 39 in three years is not credible...

In Japan (Wikipedia, 2019):

> Sogen Kato was a Japanese man thought to have been Tokyo's oldest man until July 2010, when his mummified corpse was found in his bedroom. It was concluded he had likely died in November 1978, aged 79, and his family had never announced his death in an attempt to preserve his longevity record... One of Kato's relatives was found guilty of fraud; his relatives claimed ¥9,500,000 (US$117,939; £72,030) of pension meant for Kato.

But in recent years, the scaling-up in many countries of national ID systems relying on biometric technology contributes increasingly to insure more timely, higher quality and more up-to-date registration data, especially for adults (The World Bank, 2020).

B.2.d. Demographic methods to assess completeness of census data

Various methods have been developed to assess the completeness of census data. These can generally be lumped under the heading of demographic analysis or DA. These include the following:

1. Analysis of census-survival ratios
2. Independent estimates based on estimated births, deaths, and net migrants
3. Analysis of sex ratios by age
4. Stable or quasi-stable models

The U.S. also has for many years done what is called "Demographic Analysis" (Robinson, 2010) which involves reconstructing the population by age and sex based on (adjusted) births by sex and reported deaths by age and sex, as well as estimates of net international migration by age and sex. Since births were not considered complete before about 1935 this only gives an estimate of the population under age 15 in 1950 which increases to under 75 in 2010. Starting in 2000, this was supplemented by data on the older population based on Medicare records. In 2010, the estimates for ages 65 to 74 could be done using either the true DA method or based on Medicare.

Note that even though the estimates of undercount for the U.S. census in 1940 was over 5 percent and was still about 3 percent in 1960 and 1970, the HMD does not adjust the population for undercount (Wilmoth, et al., 2019) "... no correction to the published population estimates." More importantly, the differential undercount is still high for certain age groups, especially under 5, where the estimated undercount in 2010 was estimated at 4.6 percent (O'Hare, 2019).

For the youngest ages, procedures like the **PAS BASEPOP.xls** workbook that projects the smoothed female population back for 10 years (based on preliminary estimates of mortality) and combines that information with preliminary estimates of fertility to get revised estimates of the population under age A new version in **NewPAS** called **BPE** does not smooth the population first, allowing the user to use the reported population or to smooth the population using a different method separate from this procedure. These workbook estimates are done using 5-year age groups (plus the under 1 population), but it would be useful to expand this to get the estimated population by single years of age.

The reported sex ratios by age can sometimes be compared to expected sex ratios (e.g., based on estimated of model life tables), to point to possible missing populations. For example, if the sex ratios are too low in the young adult ages, it could point to a deficit of males, possibly due to higher levels of mobility. However, it should be determined whether some of that could be the result of migration out of the country. Unexpectedly high sex ratios at older ages may point to an undercount of females in this age group.

For the oldest age groups, and for countries lacking reliable registration of deaths at older ages, it can be useful to construct stationary or stable populations based on estimates of mortality at older ages (perhaps using mortality models like the Kannisto model (Thatcher, et al., 1998)). The **PAS** workbook **OPAG** was mostly designed for expanding the population by 5-year age groups up to age 80+ by assuming that the population ages 45+ is stable. This could be modified to extend to higher ages as well. If the input populations are smoothed, then the sequence of 5-year age groups may point to the original open-ended age group being too high (if the population in the first split 5-year age group seems too high) or too low, but it is up to the analyst to decide if this is due to age misreporting or coverage errors.

The **NewPAS** workbook **Pop100h.xlsm** expands on this type of analysis, by fitting a series of stationary and stable populations to input populations assuming the input open-ended age group, but also 5 and 10 years younger. Although designed simply to extend the population up to age 100+, the same type of analysis of sudden shifts in the models can effectively correct some age misreporting or point to relative under- or over-counts in the older age groups.

The Generalized Growth Balance (GGB) method (Hill, 1987) is primarily a method for adjusting reported deaths by age for under-reporting. However, the method only estimates the completeness of death reporting relative to one of the censuses, and also estimates the relative completeness of the other census to the first. LAMBdA (2019, p. 275) found that these relative levels of completeness were generally consistent with other estimates of census completeness. (Moultrie, et al., 2013) derives this correction factor called delta by applying the Synthetic Extinct Generations method, and the use of an iterative approach to find the value that produces coverage estimates that are constant by age between selected lower and upper age limits.

B.3. Geographic and Definitional Coverage Issues

Over the course of time from 1950 to the present (and beyond) there may have been changes to the territory for a country as well as definitional changes (e.g., *de-facto* vs. *de-jure*).

For changes in geographic coverage, we look at the currently defined boundaries, but we need to know the key dates when territory changed status. For example, Alaska and Hawaii were formerly US territories, but became states in 1959. The HMD (2016) indicates:

> Until 1959, data on population and deaths refer to territory of the United States excluding Alaska and Hawaii. Since 1959, geographical coverage includes the 50 states and the District of Columbia.

For UNPD, the goal would be to include Alaska and Hawaii from 1950 on. This somewhat limits the usefulness of the HMD data for the US prior to 1959. However, the HMD protocol (Wilmoth, et al., 2019, Appendix D) covers how changes in coverage of data (e.g. due to territorial change) can be accomplished in parts of their procedures. This involves estimating adjustment factors by age for the years where it is needed. The factors for the USA allow adjustments of estimated populations (e.g., using the extinct cohort method) for the exclusion of Alaska and Hawaii before 1959. This could also be used to produce adjusted estimates from 1950-1958 that include Alaska and Hawaii.

Sometimes the issue with the geographic coverage may not be definitional, but related to situations where a country does not have access to part of its territory due to civil unrest, conflict, or territorial claims and accession by other parties. This would include such areas as Crimea (claimed by Ukraine and Russia), Abkhazia and South Ossetia for Georgia, West Bank for Israel, etc. These instances require special handling thru the application of alternative estimation methods or the use of additional time-dependent adjustment factors.

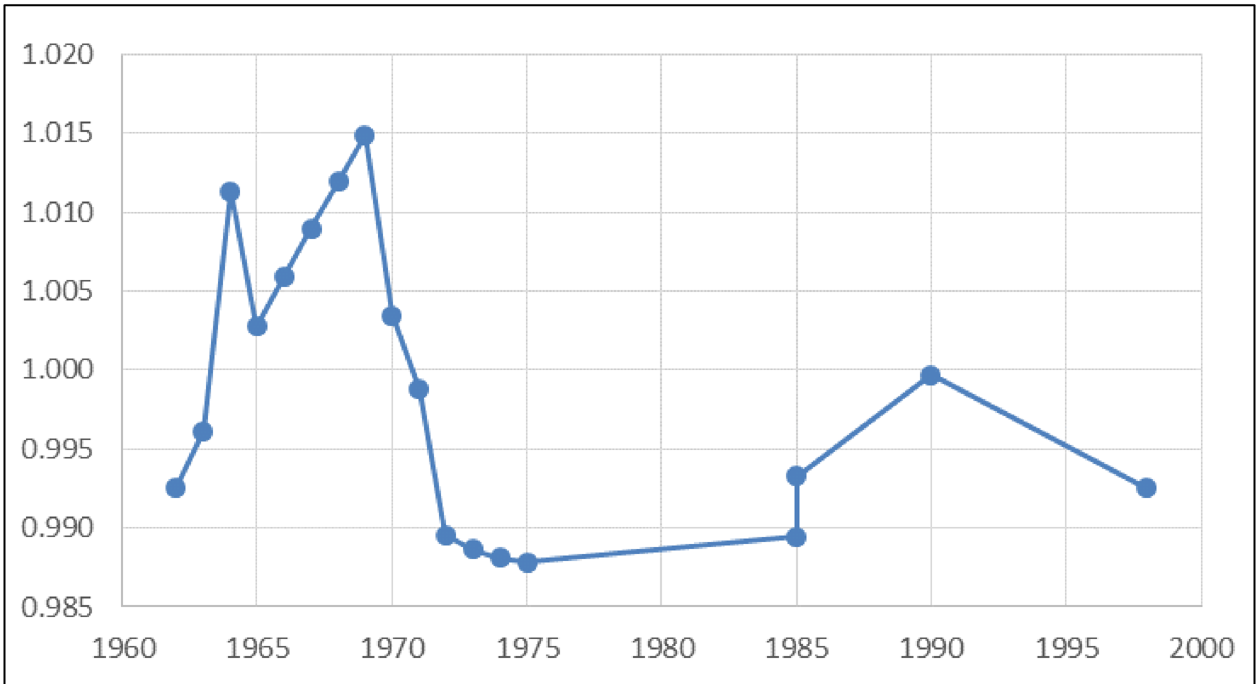
The breakup of the USSR and Yugoslavia can mostly be handled when there are subnational data from the larger entity in censuses that can be used for earlier dates. If this type of data is not available as far in the past as desired, then adjustment factors (by age and sex if possible) can be constructed, similar to the territorial adjustments used in HMD.

Some of the contemporary countries are the result of the unification or federation of one or more areas (e.g., Germany, Viet Nam, Malaysia, UR of Tanzania, Yemen). One of the problems this creates is that there may not be the same data (e.g. census data) for the same years. One solution is to create separate estimates for the subareas during the time they were separate and then aggregate the results.

Definitional changes in statistical concept of the population enumerated or tabulated (e.g. between *de-facto* and *de-jure*) can be difficult to adjust for unless the country collects and publishes enough data to compute both for the same date. In theory, the goal for WPP is to have the *de-facto* population for all countries, but that can be difficult for countries that have only collected *de-jure* population counts. In some cases a careful reading of some of the census reports may provide clues about populations of certain groups that may help in estimating the *de-facto* population.

For Indonesia, UNPD DemoData SQL database has total population series for both *de-jure* and *de-facto* populations from UNSD DYB. These overlap for some years and the ratio of *de-facto* to *de-jure* varies over time, with a minimum of 0.988 and a maximum of 1.015. The graph of these ratios does not show a smooth trend, but perhaps the analyst can determine whether the big change from 1969 to about 1972 has some explanation.

Figure B-2. Ratio of De Facto Population to De Jure Population for Indonesia



In the US, the treatment of the armed forces overseas (AFO) is an issue. As noted above, the estimates in the HMD exclude this population (and deaths) from 1940-1969.

The intercensal estimates that the HMD uses as a base that do not include the AFO or their overseas deaths, instead treat the AFO like another country with migration flows to and from the US. In the global context of UNPD this would mean this population (and its deaths) are not included in the global total.

Another issue is the treatment of foreign workers. If long-term foreign workers are present in a country such as the oil-rich countries in the Middle East, they should be counted their and not their home countries. However, some countries, like the Philippines, do count them, as long as they have been absent less the five years. The latest estimate of overseas Filipino workers (OFW) is about 2.3 million in 2017 (Philippines Statistics Authority, 2019). This is consistent with the *de-jure* concept used in the Philippine censuses. Again, in theory, these workers should not be counted in the Philippines, but rather in the country where they are working. For its 2010 census, Mongolia included the Mongolians living abroad in the final official census population. Not all Mongolians living abroad were counted, however, because the enumeration for Mongolians living abroad was conducted on a voluntary basis (people had to contact the consulate/embassy and request a web-link where they could fill out the census questionnaire).

For many countries (like Indonesia) there may be a fairly complete time series of total population (possibly by sex) in addition to census results. The census results often come out at different times. For example many countries will release a preliminary count from a census based on processing summary sheets. An intermediate count may result from processing a sample of census questionnaires. This will often be superseded by the final count after the full set of census questionnaires have been processed.

This, in turn, may be replaced by an adjusted count based on a PES and/or detailed analysis of the population. In addition, the detailed breakdown by age and sex, needed for cohort-based evaluation, often comes only years later, and sometimes initially based on a sample and later based on the full census.

The analyst should be careful to look at these different totals to make sure that the best estimate is used for each country. When a time series of populations has been published after a census, the country may not actively go back to adjust those estimates to make them consistent with the results of a new census. The **NewPAS** workbook **CONEXP**.**xls**, which implements the "constant exponential adjustment" or "Das Gupta method 6" (Johnson and Way, 2003) can be used to adjust such intermediate estimates. This basically assumes that the difference between the previous time series and the census is distributed exponentially throughout the intercensal period. This will prevent discontinuities in the time series.

If the time series is not adjusted, the analyst can develop a set of adjustment factors based on PES and/or demographic analysis (with interpolation between censuses) to obtain an adjusted time series.

[Action item: Consolidate for as many countries/censuses as possible information on census coverage to be used first for each census by itself, but also to provide a statistical distribution by region/periods we can use for other countries without such information. – UNPD status as of mid-march 2020: an initial data compilation has been completed, and need to be further consolidated.]

B.4. Developing a total population time series

It is useful to have a consistent time series of population that has been adjusted for *de-facto*/*de-jure* consistency (and other definitional issues), geographical changes, and adjustments for under- or over-enumeration (see Diagrams PopT-1 to PopT-4) . Some of these estimates may be available from the DYB database, but these are not necessarily updated after a new census is available. The first step is to see if the country has produced such estimates. These may be published on the statistical office website or in statistical yearbooks. Regardless of the source of the estimates, attention must be paid to the reference dates of the estimates and to any notes about the population type (e.g., *de-facto* of *de-jure*), geographical coverage, and whether the estimates have been adjusted in any way. Interpolation of between these estimates to a census date can help ascertain whether they are consistent with the census.

From alternative estimates, the series that is most consistent should be chosen. This is our preliminary population time series or PP(t).

The next step is to develop estimates of adjustment factors for the following, as needed:

1. *De-facto*/*de-jure* adjustment factors, adjDFDJ(t) (see Diagram PopT-1)
2. Geographic changes, adGeo(t) (see Diagram PopT-2)
3. Completeness of census enumeration, adjComp(t) (see Diagram PopT-3)

If the factor above is not an issue in the country, then the factor can be set to one. The final (or at least preliminary) estimate of population (subject to final revision based on the population reconciliation with all demographic components), FP(t), would then be:

FP(t) = PP(t) * adjDFDJ(t) * adjGeo(t) * adjComp(t)

B.5. Age misstatement of the population

Age misstatement means that the age reported in a data collection instrument (e.g., census, survey, or death certificate) is not reported correctly. It seems that in most cases the phenomena of age misstatement is a result of lack of accurate knowledge of age by the respondent. This can be because the respondent does not have enough education to accurately report the age or the society does not, in general, keep track of dates and ages. In other cases, a census enumerator may simply estimate the age based on the appearance of the person. In censuses and in most household surveys, the age of all the member of a household are typically obtained thru proxy, i.e., the person of reference or head of household. In such instances, the risk for age misreporting by another member of the household or even a neighbor increases.

One factor that can affect age heaping is the form of the question asked in the census. If the question is simply "age" then the person might not understand the concept of age at last birthday and/or (depending on when they are answering), the concept of the reference date of the census. The use of the date of birth (DoB) can help in areas where respondents are familiar with the calendar and likely to know this information. If the question only asks for the year of birth (YoB) then the age is ambiguous, and this is true even if the question asks month and year (depending on the reference date of the census). When both age and date of birth are asked, then the processing procedure should have decision rules of how to resolve inconsistencies between the two inputs. Sometimes, if only DoB is asked, there ends up being heaping on birth years that end in 0 (or 5), which will cause ages other than 0 and 1 to have heaping if the census year is not in a year ending in 0 or 5. For example, the Central African Republic census of 1988 shows heaping on ages 0, 5, and 8, with digit 0 being the highest, but 8 being higher than 5. In Gabon 1993, the heaping is highest for males on age 3 and second highest for females.

Therefore, it would be good to record the form of the age question on the census or survey, but if the question was not just age or date of birth (DoB), then some information of how these responses were reconciled would be useful. Having this information in a central repository would be useful.

[Action item for further research: UNPD has a comprehensive collection of census questionnaires, and some of them have been codified by UNSD for the 2000 and 2010 population census rounds. It would be worth indeed reviewing, coding and expanding the coverage and analyzing heaping with respect to this supplementary information. In this context, the analysis of the heaping patterns from household surveys like DHS and MICS compared to censuses can be compared to assess the effect of the instrument used to get age.]

Having data by single years of age is essential for assessing the quality and correcting some of the problems.

The LAMBdA team assumes that "age overstatement and age heaping are completely separate things." (LAMBdA team, 2019, personal communication). I don't think that they are completely separate things because heaping can contribute to age overstatement. Heaping occurs at various ages, and often across the whole age distribution, while age exaggeration is a problem much more acute at older ages, and reflect a much more systematic upward bias that plagues both age reporting for population and deaths.

Perhaps at age 78 someone rounds up their age to 80 and stays there for a while but then at age 82 jumps to 85 and age 87 jumps to 90, etc. This is probably related to the observation/assumption by Coale and Kisker (1990) that in analyzing U.S. data that reported age increased by about 1.1 year for every increase in actual age for the white population, and about 1.2 for the non-white population. They also found that "... age heaping was highly correlated with the proportion over 70 that was reported to be over 95 in censuses from many countries."

Age misreporting can be summarized in a matrix view of the population (LAMBdA, 2019, p. 31).

$$\Pi^o = \Theta \Pi^T \tag{11}$$

Where

Π^o = is the (nx1) vector of observed populations

Π^T = is the (nx1) vector of true populations

Θ = the (nxn) square matrix of transition probabilities from true age to observed age

LAMBdA (2019, p. 31) then multiplies both sides of the equation by the inverse of theta, allowing a solution for the true population:

$$\Theta^{-1}\Pi^o = \Pi^T \tag{12}$$

They argue that even with some input into the theta matrix, the inverse matrix is hard to solve for and go on to develop a simplified model of age misstatement. I have done some work on looking at simple models of theta, and found that the inverse contained some negative values, which should not happen. Instead, the matrix of transitions (numbers at true age, t, reporting observed age, o) should just be converted to the probability that someone reporting age o is actually age t.

[Action item: Further research on this is needed.]

B.5.a. Age Heaping

To assess the quality of data by single years of age, a number of measures have been proposed. One of the most commonly used is the Whipple index, which looks at heaping on digits 0 and 5. Noumbissi (1992) generalized this to have an index for each digit and Spoorenberg (2007) created a summary index to summarize the results. Myers (1940) developed a "blended" method that creates a measure of what percentage of ages are reported with each digit, and a summary measure which was the sum of the absolute differences between each measure and the expected 10%. This was later modified (1954) to use half this total (or the sum of the positive deviations) which can be interpreted as the minimum percent of the population that misreported their age. Bachi (1951 and 1953) came up with a similar measure, and others have followed.

In an analysis by Gabriella Rafaella (2019), single year data from the DYB was analyzed using all these measures. The correlations between the indices were extremely high between the major ones (at least 0.98). Whipple is not the best choice since it focuses only on heaping on 0 and 5. Of the other three, any will do. Rafaella's analysis indicates that Bachi was the most highly correlated with the others, so that may be the best choice. Unfortunately, the generic version of the Bachi index in **DemoTools** (Riffe, 2019) needs to be reviewed for consistency with the original version. I also found that the Bachi measures seemed to better reflect the heaping of deaths in the U.S. than the Myers values, but this needs to be verified, and also tested with data from other countries. This is also in question because the Myers analysis was not implemented correctly in the **PAS** workbook **SINGAGE.xls**. One additional change in the **PAS** version that should be made is to divide the original summary measure by two to make it consistent with the Bachi index (and later version of the Myers index).

[Action item: Review measures of heaping and make sure that both R and workbook versions are correct and consistent.]

I would therefore recommend using the Bachi (or Myers) index and deviations by age to measure age heaping. These are preferred over the Spoorenberg/Noumbissi methods because those are measured relative to the 5-year interval centered on the age of interest, so are not as well able to measure the relative strength of zero vs. 5 preference. Diagram PopA-1 outlines the process for reviewing and smoothing population data by single years of age.

Given the fact of high levels of age heaping in many countries (especially in the past), the question is whether the highly distorted population can impart any information about the different cohorts of the population. The choices are whether to smooth the very imperfect single year data or just deal with splitting the 5-year age groups. I feel that smoothing the single-year data is better able to redistribute the population closer to the correct age groups. It may also be useful, in some cases, to regroup the age groups into non-standard groups with the age(s) of maximum heaping in the center of the groups (e.g., 3-7, 8-12). However, if there is higher preference for 0 than 5, this still may result in an undulating single year curve.

[Action item: Improvements to some **DemoTools** functions to allow for regrouping ages would be helpful].

It is probably worthwhile to have a cutoff value of a heaping index, and always smooth the population if the index (e.g., Bachi) is greater than a certain level. Here are categories for two measures of problems with the age distribution (UN, 1952, UN, 1990):

- UN Age/Sex Accuracy Score
  - <20 is considered accurate
  - 20-40 is inaccurate
  - >40 highly inaccurate
- Whipple Index (WI)
  - Highly accurate (WI< 105)
  - Fairly accurate (105<=WI<=109.9)
  - Approximate (110<=WI<=124.9)
  - Rough (125<=WI<=174.9)
  - Very rough (175<=WI)

I have not seen recommended levels of the Bachi or Myers (revised) index to indicate problems, but given the assumption that the heaping is equal for 0 and 5 [or on 0 or 5 only] the corresponding limits would be approximately:

- Bachi or Myers Heaping Index (HI)
  - Highly accurate (HI< 1) [HI< 1.125]
  - Fairly accurate (1<=HI<2) [1.125<=HI<2.25]
  - Approximate (2<=HI<5) [2.25<=HI<5.625]
  - Rough (5<=HI<15) [5.625<=HI<16.875]
  - Very rough (15<=HI) [16.875<=HI]

Since the Bachi and Myers (revised) indices can be interpreted as the minimum amount of change needed to "fix" the age distribution, the two accurate categories indicate less than about 2% misreporting.

In the proposed R recipe for analysis, Tim Riffe (2018) had the following categories:

Table B-1. Categories of measures of age heaping

Category	**Whipple**	**Myers**	**5-yr roughness**	**Zero-pref sawtooth**
	-0.01	-0.01		-0.01
negligible	1.05	1.00	-0.01	1.00
light	1.20	3.00	0.10	1.10
moderate	1.50	10.00	0.50	1.50
heavy	10.00	100.00	inf	inf

Data for Italy 1871 from the HMD site (2019c) give Bachi indices of 8.5 for males and 11.4 for females, and there is very little heaping except for age 0 (males 7.7 and females 10.2). For age 5 there is about a tenth of the heaping as zero (males 0.8 and females 1.3). These would be characterized as "rough" in the categories above. In spite of this, these were used in the HMD without smoothing, but they do warn (HMD, 2017): "The data prior to 1906 should be used with extra caution due to problems of quality."

The choices of methods for smoothing single year data include:

1. Moving average
2. Loess
3. Polynomial

See below for methods that deal with smoothing 5-year age groups and graduating the resulting estimates.

The moving average in **DemoTools** (**mav**) is somewhat limited. The input parameter, n, indicates the number of ages that are averaged. The formula seems to be effectively:

$$map(x) = \frac{1}{n} \sum_{y=x+floor(-(n-1)/2)}^{x+floor((n-1)/2)} p(y) \qquad 14$$

One problem with this implementation is that if n is even, then the data averaged are not centered on age x. For example, if n=10 then the ages used are from 0 to 9. An alternative formulation (see also

Zelnick, 1961 and Gray, 1987) uses the number of age groups above and below the age of interest, so instead of identifying it with the value n the input would be identified as (n-1)/2, so for n=3 an alternate designation would be that we are using +/- 1 age groups. When n is an even number, the limits are no longer integers, so if n=10 then the lower limit is y = x-4.5 and upper limit +4.5 so I have been referring to it as the 4.5 moving average. In this case the fifth age below and above are weighted by 0.5. This can be summarized in the alternative formula:

$$map(x) = \frac{1}{2n}\left(\sum_{y=x-\text{int}((n-1)/2)-1}^{x+\text{int}((n-1)/2)} p(x) + \sum_{y=x-\text{int}((n-1)/2)}^{x+\text{int}((n-1)/2)+1} p(x) \right) \quad 14$$

Once it is determined that smoothing is in order (e.g., HI > 2), the separate deviations should be examined to determine what ages are heaped (vs. avoided). If the heaping is approximately equal on both 0 and 5, then it is probably a good idea to select a method like the moving average that includes the 5 ages immediately around the target age plus ½ the next oldest and youngest ages. If the heaping is much higher for one or the other, then the average over the ten ages around the target (with the extremes again at half the weight) would work. The goal is to try to reflect the underlying shape of the population curve, but without peaks and valleys that don't seem to have a basis in the true relative sizes of the cohorts. When smoothing populations, care should be taken that the process is not removing real cohort size changes.

Many smoothing procedures don't do well in adjusting the youngest and oldest populations, due in part to problems of relative underenumeration or very high degrees of age overstatement. Part of that may be corrected by using a different method for some younger ages (e.g., a moving average with fewer terms). In either case it is recommended that the analyst use special methods for these age groups, including some that adjust for net undercount as well as age misreporting.

[Action item: If populations by single years of age are not available from all censuses, then will the splitting of 5-year age groups cause differences in the population of single year age cohorts over time? This could be studied by comparing smoothed data by single years of age to smoothed and split 5-year data (using various methods). This may help identify the best methods for smoothing and splitting 5-year data, regardless of whether some data are available by single ages.]

If data are not available by single years of age, the first task is to assess if some smoothing might be needed (see Diagram PopA-2). The usual method for assessing this is the UN Age-Sex Accuracy Index (UN, 1952). This measures the degree to which 5-year age ratios differ from 0 and sex ratios change slowly over the age range. If you are concentrating on one sex, then you can look just at the age accuracy score which is just the sum of absolute differences of the age ratios from 1. The "zero preference sawtooth index" (Riffe, 2019 and Feeney, 2013) similarly looks at age ratios, but looks at the values within 10-year age groups. "If in the evaluated range there are at most two exceptions to this rule (0s>5s), then the ratio of the mean of these ratios is returned." In contrast to the UN score, this is particularly looking for large heaping on 0s.

Another measure that could be used is "five year roughness" that is part of **DemoTools** (Riffe, 2019). This is computed as follows:

> First we group data to 5-year age bins. Then we take first differences (d1) of these within the evaluated age range. Then we smooth first differences (d1s) using a generic smoother (ogive()). Roughness is defined as the mean of the absolute differences between mean(abs(d1 - d1s) / abs(d1s)).

[Action item: I looked at the code and the "ogive" seems to be the loess smoother, so the documentation should be clearer. Unfortunately, the current implementation requires the input of single ages, but the routine only needs 5-year age groups, and it would be nice to be able to use it where only 5-year data are available.]

If it is determined (e.g., based on the categories in Table B-1 or similar) that smoothing of the 5-year age groups before splitting is desired, there are several possible approaches, including:

1. **PAS AGESMTH** routines (also available in **DemoTools**)
    a. Carrier-Farrag (CF)
    b. Karup-King Newton (KKN)
    c. Arriaga
    d. United Nations
    e. Strong moving average
2. Additional methods in **DemoTools**
    a. Zigzag
    b. Moving average

The first three methods all start by creating 10-year totals and then splitting them by fitting a curve to the series. A quick look at some results seems to indicate that the CF and Arriaga are the closest together, followed by CF and KKN, and then KKN and Arriaga. The Arriaga method has the advantage of providing a method for estimating the youngest and oldest ages. The Strong smoothing is really just what I would term a 0.5 moving average over 10-year age groups (or mav with n=3). If this seems the most appropriate, it may be better to look at developing a stable or quasi-stable population model. The UN method seems to be somewhere between the first three methods and the Strong smoothing.

When 5-year age groups are grouped into 10-year groups before splitting starting at age 0 or 10, then any excess heaped on 0's will get redistributed to older ages, whereas many of those with rounded ages will have rounded up. This can be improved by starting the smoothing at age 5 instead.

[Action item: It might be useful to look at summary measures of heaping and see what the impact of the different methods is. Maybe look at the correlation of these measures by smoothing method. Also compare these results to smoothing of single year data when that is available.]

Once a smoothing method has been selected, a graduation method can be chosen from the **graduate** function in **DemoTools** (Riffe, 2019):

1. Sprague
2. Beers (ordinary or modified)
3. Monotonic spline
4. Pclm (Penalized Composite Link Model)

I have used the Beers (ordinary) in the past, and it works well. At the U.S. Census Bureau we used it regularly, often with the Johnson (Stover, et al., 2008) adjustment for age 0, to create the base population used in the Rural-Urban Projection (RUP) program (U.S. Census Bureau, 2014) for projections. The modified version does some smoothing and does not necessarily reproduce the totals. The Sprague method seems to work well also. The monotonic spline has the advantage of allowing varied inputs, including 10-year age groups or 0-4 and then 10-year groups. This can be helpful to force the excess population at ages ending in 0 to be redistributed in both directions (younger and older). It also guarantees there will be no negative populations, which can happen with Beers and Sprague. The **DemoTools monoCloseout** routine is a blend of Sprague at younger ages with the monotonic spline at older ages to prevent any negative values. We may want to look at the assumptions at the youngest ages: a test run of the monotonic spline using the 1950 WPP 2019 female population for Bangladesh (admittedly not real data for analysis) resulted in much lower estimates for age 0 compared to other methods (at least 8% lower). However this same data did produce negative numbers at older ages for the Sprague and both Beers versions.

These smoothed/graduated populations can then be used for:

1. Cohort comparison
2. Mortality estimation
3. Census evaluation
4. Estimation of 1950 population

[Action item: Comparison of different methods for splitting 5-year ages into single years of age to determine which works best in certain circumstances.]

[Recommended steps: Single age data should be used as much as possible and be smoothed if necessary. The original or smoothed single age data should be compared to graduated 5-year data. Series should be compared, and a set of conditions/criteria/threshold be formulated for when to accept one series over another.]

B.5.b. Age overstatement or understatement

Age over- or under-statement (other than related to age heaping) can be more difficult to correct. The LAMBdA (2019) model discussed above dealt with this for older ages (above age 45) by using a model based on conditional net overstatement probabilities and age-specific propensities to over-state age estimated based on the results of a 2002 evaluation study linking Costa Rica 2000 census records to voter register records that included age information from birth certificates. The net overall tendency to overstate is estimated based on cohort comparisons of one census to another adjusted for relative completeness of the two censuses estimated from the GGB (as noted above).

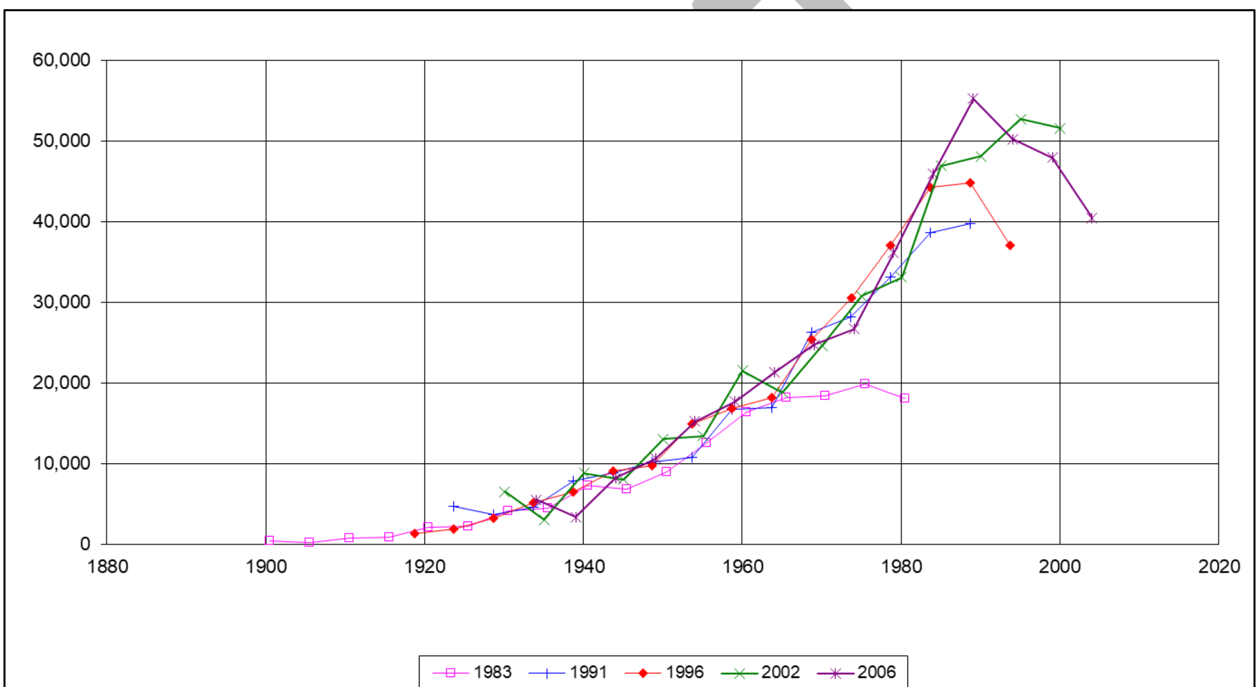
As noted above, fitting a stationary or stable population model to data at older ages (e.g. using the **NewPAS** workbook **POP100h.xls**) can help develop improved estimates at these ages, but the question is at what age is it better to replace the reported data (possibly smoothed) with models.

In countries with reliable and complete death registration, the use of the (near-) extinct generation method as used the HMD method protocol allows to obtain more accurate population estimates at older ages than typically available from censuses (Wilmoth, et al., 2019, Condran, et al.,1991, Jdanov, 2015).

B.6. Comparing population estimates by age and sex

Once the initial analysis of the population estimates is completed, the results based on different censuses (or data sources) should be compared to makes sure that they are consistent. The **PAS** workbook **GRPOP-YB.xls** (or **NewPAS** versions **GRPOP-YBNEW.xls** and **POP1COHNew.xls**) plot populations estimates for different dates by birth cohort instead of by age. This type of plot can also be applied to the reported populations to help identify cohorts that may be smaller or larger than expected.

Figure B-3. Male Population of Djibouti by Census and 5-Year Cohort



If some of the estimates have been based on smoothing (and adjusting) populations by single years of age and some have been based on smoothing and splitting data by 5-year age groups, then this can help identify if there are inconsistencies in the resulting estimates. This type of plot can also help identify cohorts that may have been affected by migration.

Another way of assessing the consistency of the estimates is to compute census survival ratios. Using populations by single years of age should make it easier to do this analysis if the intercensal interval is not 5 or 10 years. Survival ratios are expected to decrease monotonically with age, and to always be lower than 1. In addition, any abnormal deviations in some age groups are expected to follow the birth cohorts as they age if these deviations reflect real historical events (e.g., baby boom or baby burst). Otherwise such abnormal deviations reflect (greater) deficiencies in one of the censuses being compared, or some of the age groups experience substantial in/out migrations.

C. Mortality

C.1. Overview

Mortality is very important to the estimation process and in some cases it is useful in analyzing the population age structure. The move to single year data means the split of age-specific death rates (${}_nm_x$) from abridged ages (0, 1-4, 5-9, ..., 100+) to single ages (complete life tables).[4]

Mortality data comes from various sources, but primarily from vital registration, sample registration systems, surveys, and censuses. Additional information may come from health and demographic surveillance systems (HDSSs) such as Matlab in Bangladesh, and those in the INDEPTH network (http://www.indepth-network.org/), and the ALPHA network (https://alpha.lshtm.ac.uk/). The direct estimates are generally based on deaths and exposure directly reported for particular time periods (e.g., vital registration and population estimates, household deaths and enumerated population, birth or sibling histories). Indirect methods use data on the survival of children, parents, siblings, or spouses. As with population data, one key initial task is to ascertain the reference date(s) of the data, as well as the population universe of the data. For example, if certain people are excluded from a survey (e.g., married women only, workers in group quarters, or residents of certain provinces), this should be noted and an assessment be made of the possible magnitude.

Next, the territorial and population group coverage of the data should be determined, if possible. In some cases there can be problems with access to parts of the country due to war, civil strife, or weather/natural disasters. In certain countries, the registration system only includes data on the national population (e.g., Japan and Saudi Arabia) and in some it may include deaths overseas to the national population. In these cases, an attempt should be made to either obtain an appropriate population (e.g., the Japanese population) and assume the same mortality for the excluded population, or try to determine an adjustment factor for the deaths.

The data on deaths by age and sex whether from vital registration or household deaths collected in a census or survey should first be adjusted to account for any deaths with unknown sex and age. The standard method, as with population, is to prorate the deaths based on the distribution of known age and sex.

C.2. Fatal discontinuities

The major goal of mortality estimation is to get the overall general level of mortality by age and sex. We expect these estimates to change slowly from year to year (and age to age). However, there are situations that can cause "fatal discontinuities" in the terminology of the Global Burden of Disease (GBD, 2018b). These include:

- Wars
- Civil unrest
- Natural disasters (hurricanes, earthquakes, floods, etc.)

[4] The UNGROUP package from Marius Pascariu using dx and Lx from the abridged life tables does a great job at graduating in a very robust and consistent way the life tables.

The goal in mortality estimation is that the impact of these unusual events are generally added onto the underlying mortality risk. In some cases this may entail "removing" these causes from the data and then adding them back at a later stage of the process.

Another disruption of expected mortality change over time has been the impact of HIV/AIDS on mortality (and the population). Some groups (e.g., UNAIDS, U.S. Census Bureau, GBD) try to estimate the non-AIDS mortality, model AIDS mortality separately, and then add the two parts together. This was done by the UN through WPP 2017. In WPP 2019, they instead tried to model the impact by using under-five and adult mortality estimates from census, survey and VR, data on adult HIV prevalence and coverage rate for ART, and the age pattern of mortality was derived using a special life table model that accounts for the impact of HIV/AIDS (Sharrow, et al., 2014; United Nations, 2019).

C.3. Sources of mortality data

C.3.a. Vital registration data

For vital registration data, it should be determined whether the data reported are by date of occurrence or date of registration. Delayed registration can cause undercounts of deaths for certain years, but if there is a push to register deaths (even from earlier years) that may cause an overcount, depending on how the data are processed. Also, when the level of registration is low, increasing the coverage of registration may make it appear that deaths are increasing faster than they really are.

Place of occurrence vs. place of residence also matters, especially for small territories (e.g., Monaco has a lot of foreigners coming to give birth and using these vital events with the usual residents leads to a gross fertility for Monaco that is 2-3 times higher than expected). At the subnational level, especially between urban and rural areas, this also matters if/when events (i.e., births and deaths) happen in health centers/hospitals.

The HMD methods (Wilmoth, et al., 2019) provide the most rigorous methods for producing life tables for countries with excellent data. It would be convenient to use these data as is since they are by single years of age and often cover a long period of time. However, there are some areas where these estimates (or their component parts) may not meet the needs of the UNPD:

1. Some of the population universes do not cover the desired population universe
    a. Italy changed the basis of its population estimates from *de-facto* to *de-jure* in 1981.
    b. Japanese vital registration covers only persons of Japanese nationality.
2. Populations are not adjusted for underenumeration (unless done so by the country)
    a. U.S. population not adjusted in spite of estimated undercounts of up to about 4 percent and high differential completeness by age.
    b. The Canada population estimates starting in 1971 are adjusted for undercoverage of the census, but the earlier years are not.
3. The estimates do not go back to 1950 (e.g., Israel and Chile) in instances where some of the vital registration data are deficient (e.g., at older ages).

C.3.b. Census data

Censuses can contain direct data on deaths (e.g., household deaths in the last 12 months) as well as data for indirect estimates (e.g., children ever born and children surviving or orphanhood, or even intercensal population survival at older ages). An advantage of the census as a source of data is the full coverage of the country. This has advantages when the deaths under study are relatively rare (like for older children or maternal deaths). In some cases it also allows for developing estimates for subnational areas, perhaps at very low levels of geography. However, censuses are very expensive and each question asked adds to the cost. They also tend to be taken only about every 10 years in most countries, so it takes time to get updated information. Finally, for some types of questions more extensive training may be needed to properly collect the desired data.

C.3.c. Survey data

Surveys are often used to collect data on a more frequent basis than censuses. They often include more extensive questionnaires, and the interviewers usually get more training than census enumerators. The downside to surveys are that they are samples with sampling errors that need to be taken into consideration.

C.4. Adult mortality

See Diagram MortA-1 for an overview of the adult mortality estimation process. This process is viewed differently in Diagram MortA-2, which summarizes the actions to take given various combinations of mortality data (or the lack thereof).

C.4.a. Direct estimates

Direct estimates of adult mortality can be based on vital registration systems, sample vital registration systems, and retrospective reports on deaths in a household reported in a census or survey. Diagram MortA-3 summarizes the process for dealing with death registration data which is also summarized in Diagram MortA-4 as a decision table.

Direct estimates of deaths by sex and age (whether single years of age or by 5-year age groups) should be reviewed and smoothed if needed (see Diagram MortA-5) prior to computing age-specific death rates.

For countries where adult death reporting may be incomplete, there are a number of methods that have been developed to try to correct for this by comparing the deaths by age to the population by age. The first one was the Brass Growth Balance method (Brass, 1975). However, this method assumes that the population is stable, so is not useful in many situations. The Generalized Growth Balance (GGB) method (Hill, 1987) and the Synthetic Extinct Generation (SEG) method by Bennett and Horiuchi (1981 and 1984) removed the requirement for stability by using two censuses. The combined GGB-SEG method involves using the relative completeness of two censuses to adjust the SEG method for differences in completeness of two censuses. This stage is summarized in Diagram MortA-6.

Hill et al. (2009) did testing of the GGB, SEG, extended SEG (where the relative completeness needed to get the c(a) values to be consistent by age is found by iteration), and the GGB-SEG and found:

> In populations thought not to be affected by migration, the optimal strategy appears to be to apply the General Growth Balance method (fitting to the age range 5+ to 65+) to estimate census coverage change, adjust one or other of the two population age distributions for the estimated coverage change, and then apply the Synthetic Extinct Generations method to the adjusted data, also fitting to the age range 5+ to 65+. In populations thought to be experiencing

> substantial migration, applying either the GGB or SEG to the age range 30+ to 65+ reduces the effect of migration; applying both and averaging the results appears to give the smallest error, but this procedure seems very inelegant.

Murray, et al. (2010) did a similar set of simulation tests and found:

> On the basis of the three different validation datasets, we believe SEG 55–80, GGB 40–70, and GGBSEG 50–70 are the best methods that can be currently used to estimate relative completeness of death registration. The combination of the three optimal DDMs will yield much better results than the current practice of application of DDMs without optimal age trimming. Selection of optimal age trims has also substantially reduced the bias associated with migration reported in previous work.

They do go on to say:

> While usage of partial death registration is useful for estimating mortality levels among adults, the application of DDMs, even from the optimal age trims we have suggested here, should be interpreted with considerable caution; the uncertainty around relative completeness of registration is likely to be at least +/-20% of the estimated level, and perhaps considerably more.

The uncertainty of +/-20% seems very high to me. I would like to look at this in more detail.

The LAMBdA project (2019) did more extensive testing of many of these, which are summarized in their table 3.8 (p. 39) reproduced below as Figure C-1.

Figure C-1. LAMBdA Project Methods to Adjust for Completeness of Death Registration

Table 3.8: Methods to adjust for completeness of death registration: assumptions and required data.[1]

Method	Assumptions	Required Data
Brass (B)	1-2-3-4-5	B
Brass-Hill (BHill)[1]	2-3-4	A
Brass-Martin (BMartin)[2]	1-2-3-4-6	B
Bennet-Horiuchi No 1 (BH_1)[3]	1-2-3-4	A
Bennet-Horiuchi No 2 (BH_2)	1-2-3-4	A
Bennet-Horicuhi No 3 (BH_3)	1-2-3-4	A
Bennet-Horiuchi No 4 (BH_4)	1-2-3-4	A
Bennet-Horiuchi No 5 (2SBH_4)	1-2-3-4	A
Preston-Hill No 1 (PH_1)	1-2-3-4-5	B
Preston-Hill No 2 (PH_2)	1-2-3-4	A
Preston-Bennet (PB)	1-2-3-4	A
Preston-Lahiri No 1 (PL_1)[3]	1-2-3-4	A
Preston-Lahiri No 2 (PL_2)	1-2-3-4	A

This should be A per personal communication.

[1] BHill is a method we also use to retrieve both estimates of the ratio of completeness of the first relative to the second census and estimates of relative completeness of death registration.
[2] BMartin is a variant of Brass classic method that relaxes the assumption of stability and assumes instead limited forms of past mortality decline.
[3] See Chapter 9 (Section 4) for definitions of the four variants of Bennet-Horiuchi method and the two variants of Preston-Lahiri method.

KEYS FOR ASSUMPTIONS
1. Identical completeness of census counts in both census
2. Closed to migration
3. No age misreporting
4. Invariant completeness by age
5. Stability
6. Quasi stability

KEYS FOR REQUIRED DATA
A. Two censuses and intercensal deaths
B. One census and one to three years of deaths by age

They did an extensive series of simulations to see which methods did best under different conditions and concluded (p. 51):

i. In the absence of exogenous information about the difference in completeness between the two census and if the assumption of age invariant completeness holds, use BHill [GGB] method to estimate the relative completeness of two consecutive censuses;
ii. Adjust the observed rate of intercensal growth to account for defective relative completeness of censuses use the two-stage procedure (2SBH_4) [GGB-SEG] (see Section 6 of Chapter 10).

LAMBdA (2019, p.40) discusses adjustments to deaths and the impact of age misreporting:

> ... all methods assume either no age misreporting or, alternatively, age misreporting that perturbs only trivially the figures of cumulative population above adult ages. This poses a conundrum: if, as asserted before, LAC population and mortality counts are heavily affected by age overstatement, how can one expect to obtain precise estimates of relative completeness

using techniques that are highly vulnerable to age misreporting? Two conditions offer an escape from this trap. The first is that the type of age misreporting that predominates in LAC is net age overstatement. When using cumulative populations over some age x the damage done to the target quantity by age misreporting only depends on population flows across age x originating at younger ages. It is insensitive to transfers of population above age x. Furthermore, the relative volume of flows, e.g. the relative error of the target quantity, is generally light for late adulthood (less than 65 or 70) though it begins to mount after age 75 or so. Since in all cases computations only require to employ observations up to ages 70 or 75, the impact of age overstatement will be minor. The second favorable condition that circumvents the problem is that the optimal method (Bennett-Horiuchi No 4, BH 4) is also the least sensitive to age misreporting of the type encountered in LAC...

The GGB, SEG, and combined GGB-SEG methods were also used in the development of the Global Burden of Disease (GBD, 2018a and 2018b) 2017 estimates of mortality.

Note that the **TDE** workbooks **AM_GGB** and **AM_SEG** allow inputs of intercensal net migrants by age, and use these to adjust the age-specific growth rates in the methods. In the **AM_SEG** workbook, the combined GGB-SEG method is done by entering the intercept from the GGB results, and the extended approach (iterated estimate of relative census completeness) can be done using Excel solver to find the flatter line. The R package "**DDM**: Death Registration Coverage Estimation" developed by Riffe, Lime and Queiroz implements these methods, and provides various options to select the optimal age range used for each method.

Overall, Hill (2017, page 18) recommends a GGB-SEG combination using different age trims for the two applications with first an estimate of census completeness change by applying GGB to the wide age range 5 to 65 years, adjusting the populations accordingly, and then secondly applying SEG, estimating coverage for the age range 50 to 70 years. This approach should maximize the effectiveness of GGB for estimating completeness change, while minimizing SEG errors from migration.

Direct estimates of mortality can also be made using data collected on the survival of siblings. This is done in some DHS surveys (Croft, et al., 2018). This method looks at the exposure and deaths of siblings and computes the age-specific death rates by 5-year age groups, ${}_5m_x$ (which are referred to in the source as age-specific mortality rates) for various retrospective periods. These rates are also converted to ${}_5q_x$ values to estimate the probability of dying between ages 15 and 50, ${}_{35}q_{15}$.

C.4.b. Adjustment for age-misstatement

If deaths are available by single years of age, they should be examined for evidence of age heaping. Although the usual measures of age heaping were designed for analyzing population data, there is no real reason why they couldn't also be used to summarize heaping of ages at death. The problem with age at death is that the person with perhaps the best knowledge of their age is the decedent, so the question is how well do other informants (e.g., doctors or family) know (or have documentation) of the person's age. The propensity for the reported age to be too high can be based on the appearance of the body, or to the assumption that "they must be pretty old to have died."

In the LAMBdA model of age misreporting (2019) they assume that while the propensity for net age overstatement may be different for the population and deaths, the pattern of how far away the age is reported is assumed to be the same for both.

The HMD (Wilmoth, et al., 2019) method is to use deaths by single years of age as they are. If the data are only available by 5-year age groups, they are split into single years of age using a quadratic spline. This is probably a good approach to use, but if single year data are available there should be an investigation of whether there is evidence of age heaping that might affect the results. A brief investigation of the US data seemed to indicate that there was some heaping, but the pattern was not just preference for digits 0 and 5 but may have also had some heaping on year of birth in the late 1950's early 1960's.

[Action item: As noted before, the implementations of the Bachi and Myers methods in both **DemoTools** and **PAS** need to be verified. Once this is done, a separate comparison of Bachi vs. Myers applied to death data should be done, since the shape of the age distribution of deaths is quite different from the population.]

Comparing the spline estimates based on 5-year estimates to the reported data gives another view of where there may be heaping or avoidance of ages but also where the spline may not work well (e.g., at younger ages).

In the HMD protocol (Wilmoth, et al., 2019), deaths in the open-ended age group (or above a certain age >=85) are then split into Lexis triangles directly by fitting the Kannisto model to the data as described in HMD protocol (2019, Appendix C).

LAMBdA (2019), as noted above, has developed a model of net age overstatement that they use for adjusting both population and deaths. Unfortunately, I did not find the explanations sufficiently clear to be able to reproduce their methods. The LAMBdA website is still under development, and currently only the original data and life tables for the "pivot" years are available. The final full set of complete life tables has not been released yet. The website will also, at some point, include copies of the code used to do the analysis (partly in R and partly in STATA). This may help understand the process more fully.

C.4.c. Indirect Estimates of Adult Mortality

A number of methods have been developed to estimate adult mortality indirectly. These are usually based on reports in censuses or surveys of the survival status of various relatives, like parents (orphanhood methods), siblings (sibling methods), and spouses (widowhood methods). These data are often combined with a model life table to develop a time series of dated estimates of adult mortality, often summarized as ${}_{45}q_{15}$ or ${}_{35}q_{15}$ (see UNPD, 1983 and Moultrie, et al., 2013). This process is summarized in Diagrams MortA-1 and MortA-2.

Another method for estimating adult mortality is to look at census survival ratios by age (see UNPD, 1983 and 2002). If the censuses are about 5, 10, or 15 years apart, the patterns of survival ratios by sex can be examined and perhaps compared to model life tables to get implied levels of overall or adult mortality (e.g., e0, e10, e15, or 45q15). If the census populations are first smoothed to remove most of the age misreporting and heaping, then it will be easier to see the levels of mortality without extra fluctuations. If the intercensal period is not an even multiple of 5 years, then some interpolation work is needed to get the cohorts to line up, but it can be done. An alternative approach is to try projecting the earlier census to the later census date (perhaps with some interpolation) using a series of model life tables with different levels of mortality to see what levels are needed to reproduce the later census. As noted above, high levels of age misreporting can limit the usefulness of this method. Differential age misreporting by age and between the two censuses can also make the estimates of mortality unreliable.

A final problem is the impact of migration. Net out-migration will cause this method to over-estimate mortality and in-migration the reverse. If there are available estimates of net migration, they can be used to adjust the mortality estimates for the impact of migration.

C.4.d. Old age mortality

Old age mortality based on registration data or household deaths is frequently under-estimated due primarily to net age overstatement. To correct for this, a model of old-age mortality can be fitted to estimates at younger adult ages (adjusted as necessary for under reporting of deaths) to produce adjusted estimates for older ages (see Diagram MortA-8).

When mortality data are available at least up to age 90 or more, the preferred model currently seems to be the Kannisto model (Thatcher, et al., 1998) which assumes that the force of mortality follows a logistic curve with upper asymptote 1. In some cases this can still result in estimates that appear to be much too low compared to the death rates in other countries. In this case, an estimate of mortality at some high age can be introduced to "pull up" the estimates to a more realistic level. One such estimate was developed by Coale and Guo (1989) and assumes that ${}_5m_{105} = 0.66 + {}_5m_{75}$. Coale and Kisker (1990), using a different model, assumed that 1m100 = 1.0 for males and 0.8 for females. I feel that having such different asymptotes for males and females is unrealistic. I also wonder about the implicit assumption that the logistic function is symetric around the inflection point.

The **NewPAS** workbook **LTExtNewN.xls** fits the Kannisto model to observed data using weighted least squares that can include the Coale and Guo estimate.

The HMD (Wilmoth, et al., 2019) uses the Kannisto model fitted to the last 20 single year death rate estimates to distribute deaths in the open-ended age group (e.g., 80, 90, and 100).

When mortality data are only available for early open age groups (e.g., 60 or more), to complete the life table the nqx values are extrapolated by the United Nations until no survivors remain, by fitting a Makeham function through the last six nqx / (l - nqx) values available (United Nations, 1982, p. 31).

In instances where mortality data are available at least up to age 80 or more, the United Nations has been using the Coale and Guo (1989) function to extend mortality rates at older ages.

The R package “MortalityLaws” developed by Marius Pascariu implements most of these mortality functions with various fitting options, and a subset of them are used by the life table functions implemented in the DemoTools package.

C.4.e. Validation and Review

The estimates of mortality by age and sex over time should be reviewed for consistency (see Diagram MortA-7).

C.5. Infant and Child mortality

The overall process of estimating infant and child mortality is summarized in Diagram MortC-1.

C.5.a. Direct estimates

Ideally, infant and child mortality are computed based on registered deaths and births and population estimates (see Diagram MortC-2). Unfortunately, for most countries of the world this is not possible.

Estimating infant and child mortality can be a problem when vital registration data are incomplete. In this case, the data needed are both the infant and child deaths (ages 0-4) and the births or population under age 5. The sharp decline of mortality from early infancy to age 4 (and beyond) makes the problem more difficult. When using census data as a base for age-specific death rate, the population under age 5 (and especially under age 1) can be particularly distorted by undercounting. As noted before, when death registration is incomplete it can be difficult to know whether what seems like a slowdown in reductions of mortality is real or just the result of increasing completeness of death registration.

For many developing countries, the best source of infant and child mortality data is from complete birth histories in surveys like those in the Demographic and Health Survey (DHS), Multiple Indicator Cluster Survey (MICS), and older World Fertility Survey (WFS) programs. These estimates are developed by tracking from each sampled surviving mother the births and deaths over time and age of death, and constructing Lexis rectangles of deaths and exposures of the age groups early neonatal (under 7 days), late neonatal (7-30 days), neo-natal (under 30 days), post neonatal (30 days to 1 year), then single ages 1, 2, 3, and 4 (Croft, et al., 2018). The resulting ${}_nq_x$ values can then be chained together to get the infant mortality rate (IMR or ${}_1q_0$), child mortality rate (${}_4q_1$), and under-5 mortality rate (U5MR or ${}_5q_0$).

C.5.b. Indirect estimates

Since many developing countries still do not have fully functional complete vital registration, a number of indirect methods have also been developed. These are basically variations on the original Brass method (1975) that looks at the proportion dead of children ever born by mother's age, and transforming that to the cumulative probability of death by ages 1, 2, 3, 5, etc. When combined with a model life table system and the assumption that mortality has been changing linearly over time, these can be converted to estimates of the IMR or 5q0 and reference dates over time. Due to differences in the age pattern of infant and child mortality by country (and model life table), the estimates of 5q0 are considered more robust to the assumed model life table. The time since first birth (TSFB) variant of this method requires data on the time of the first birth for each mother, and tabulates the CEB and CS or CD by the time since first birth: 0-4 years, 5-9 years, etc., and has been preferred over the age of mother version for use in IGME estimation since 2014 (UN IGME, 2014) This method has been found to be more reliable in situations of changing fertility (and mortality) (Verhulst, 2017).

C.5.c. Consolidating infant and child mortality estimates

Evaluations of estimates of infant and child mortality usually rely on comparing the estimates from different sources and methods over time. The UN Interagency Group for Child Mortality Estimation (IGME) has developed procedures for collecting, re-estimating, and consolidating estimates of infant and child mortality (primarily ${}_5q_0$) to produce estimates for all countries of the world from about 1985 to 2018 (UN IGME, 2019a and 2019b). This process is summarized in Diagram MortC-3 with additional notes in Diagram MortC-4. The estimates considered and used in this process are presented on the childmortality.org website in addition to the B3 model results to allow the analyst to evaluate the process and results. Unfortunately, the IGME estimates do not always go back to 1950 for all countries, so other methods will be needed. For WPP 2021 UNPD will use a hierarchical model to estimate back series to 1950 in a more coherent way based on IGME output estimates. Regional asymptotic levels of mortality that existed back in the 1950s should be used to predict more sensible common regional levels while preserving some reasonable ranking relations between countries within regions.

IGME uses assumed model life table relationships to estimate IMR for countries without registration data. For some countries (e.g. in Western Africa/the Sahel) they have developed new models since the existing models do not seem to accurately reflect the relationship between IMR and U5MR. IGME also has produced models of the sex ratio of infant and child mortality (citation).

The GBD (2018a and 2018b) has also produced infant and child mortality estimates from 1950 to 2017. The inputs to the process are theoretically available and the process is described, but the system is much more complicated. These estimates should be reviewed and considered as well in light of existing empirical evidences.

One issue with infant and child mortality is the selection of separation factors to allow conversion between q_x and m_x or l_x to L_x. Demographers for many years used the formulas developed by Coale and Demeny (Coale, et al., 1983) for their model life tables, especially for age 0. Recent data have indicated that rather than continuing to lower levels as IMR declines, the values actually tend to rise (Andreev and Kingkade, 2015). These equations should be preferred over the Coale-Demeny for most countries, especially as IMR declines. **DemoTools** life table functions will include these alternatives for getting separation factors.

[Action item: Preliminary analysis of the timing of the changes in the U.S. seem to coincide with the Roe vs. Wade Supreme Court decision that legalized abortion, although the pattern is rather complicated. More research such as looking at the data used by Andreev and Kingkade combined with data on the prevalence of abortions may be informative.]

C.6. Full life table series

Once the data for a country have been fully explored, it is necessary to create complete life tables for every year from 1950 to 2020. For countries with good vital statistics, these may already have been computed based on deaths and population data, perhaps with some adjustments. For most developing countries, however, this is not the case. In some cases, there may be complete life tables available for selected years and the question is how to fill in the gaps.

In the **RUP** program, age-specific death rates, ${}_1m_x$, for years between inputs were interpolated by age on the log scale:

$$\ln({}_1m_x(t)) = [(t_1\text{-}t)*\ln({}_1m_x(t_0)) + (t\text{-}t_0)*\ln({}_1m_x(t_0))] /(t_1\text{-}t_0)$$

For countries without data to compute full life tables, a simple solution is to combine available infant and child mortality, ${}_5q_0$, (perhaps from IGME) with indirect estimates of adult mortality, e.g., ${}_{35}q_{15}$ or ${}_{45}q_{15}$ with a new flexible model life table system like the log-quad (Wilmoth, et al., 2012) or the SVD (Clark, 2019). In these cases, the question is whether these model-based estimates will be consistent with more empirically-based estimates.

This may be a case where mortality-standard based systems, like the Brass (1975) logit system, would result in estimates that better combine the characteristics of the mortality profile of the country. The limited Lee-Carter method (Li and al., 2004) and the UNPD tool **Li_2018_Limited_Lee-Carter-v4.xlsm** can be used for this purpose. This essentially creates a standard which is an average of 2 or more years of data as well as observed rates of change by age.

[Action item: Is there a way to leverage the SVD model to capture the country characteristics (using additional dimensions) to come up with estimates that better reflect the situation than just using just the usual two parameter inputs? In principle, it should be possible to use a similar approach as with 1982 UN model life tables with a user-defined standard (UNPD, 1982 chapter IV).]

Another issue with developing life tables where there is minimal data is how to constrain the results so that unexpected results (like crossovers by sex) don't occur. In the past, UNPD has used the coherent (non-divergent) option (Li and Lee, 2005) in the limited Lee-Carter tool. But there is a way also to accomplish this with the SVD model.

[Action item: Update the limited Lee-Carter model tools to output life tables by single calendar year. Also allow for single year mx input or build in graduation to produce complete life tables.]

For countries with significant levels of HIV/AIDS, alternative models can be used (e.g., Sharrow, et al., 2014) that combine information about prevalence of HIV and levels of antiretroviral treatment (ART) over time. But the UN plans to use an expanded SVDcomp model that would allow modeling countries with and without HIV/AIDS epidemics.

D. Fertility

D.1. Overview

Fertility has a big impact on the population, but we know that the immediate impact is only on the population under age 1. Over the last century fertility has changed significantly for most countries of the world.

D.2. Vital registration data

Universal birth registration in a country is the gold standard for collecting data on fertility. Some countries (e.g., India, Bangladesh, and China) have developed sample vital registration systems (SRVS) that use data from selected areas to estimate the numbers for subareas of the country as well as the country as a whole. The design and assumption is that the registration system in the sample areas is completely recording the births. In addition, the sample design must be developed, and periodically updated to ensure the accuracy of the national estimates.

If birth registration is not complete, then improvements in completeness may be interpreted as rising fertility (or slower decline). There can also be problems with delayed registration, where a birth is not registered until months (or years) after the birth occurred. If there is a substantial time lag then some births that die in infancy may not be reported at all, and the infant death may also go unreported.

Computing ASFRs from vital registration data also requires having good estimates of the female population by age for the denominators. See Diagram F-3 for the process of estimating fertility based on registered births.

D.3. Census and survey data

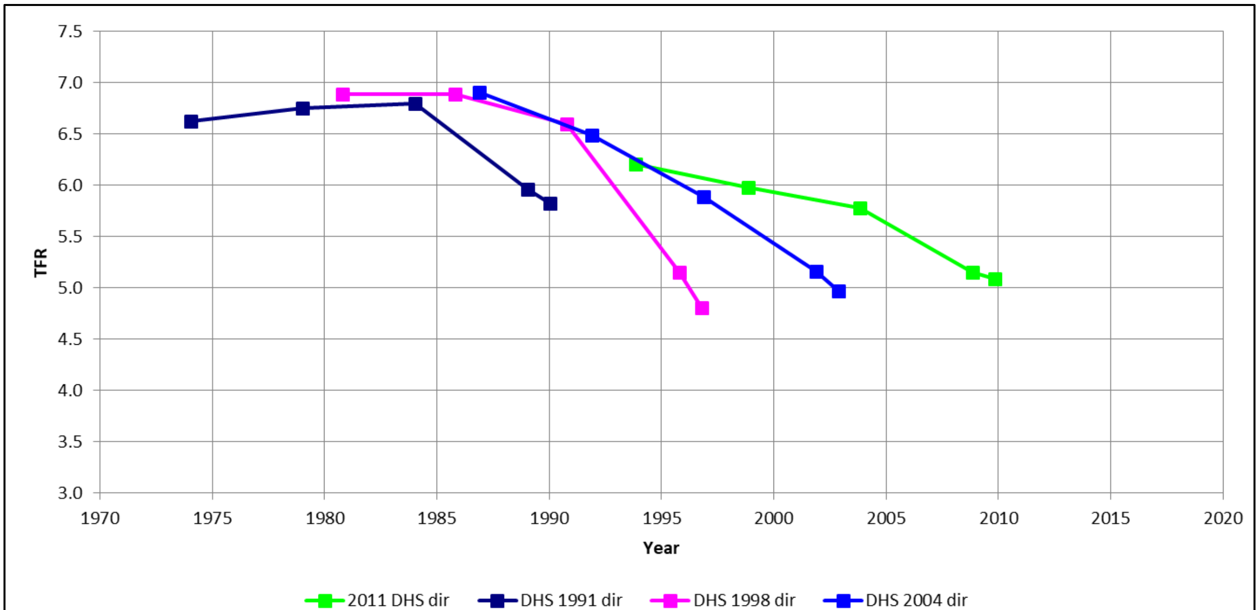
D.3.a. Direct fertility data from censuses and surveys

Birth data can also be collected in censuses and surveys. In censuses and some surveys, this is usually in the form of a question asked of women of childbearing age on the number of births in the last 12 months. This may be supplemented with questions on the number of children ever born (as well as those surviving or who died). Diagram Fert-1 summarizes the process of putting together these sources of data. Data on recent births may be inaccurate if the respondent is not clear on the meaning of the reference period. In other cases the respondent, oftentimes the household head, may omit some births in the household. In high mortality areas, newborns or infants that died may also be omitted.

In addition to these questions, surveys, starting with the World Fertility Survey (WFS), have also collected data on full birth histories of women in the reproductive ages (generally ages 15-49). These data can be used to develop estimates of age-specific fertility rates (ASFRs) for periods of time prior to the survey. Two problems arise from this method. First, if older women are not included then there will not be ASFRs for older age groups for earlier time periods. Second, the respondents often don't have a good understanding of the calendar, so births can end up unevenly distributed over the different periods before the survey. Sometimes births are misplaced to before a key reporting window (usually the 5 years before the survey) so both the respondent and the enumerator do not need to go through longer survey modules.

Figure D-1 shows TFR estimates from DHS surveys in Cameroon, and it can be seen that the "recent" estimates in each of the surveys seem to significantly underestimate the fertility compared to the estimates based on later surveys.

Figure D-1. DHS Direct TFR Estimates for Cameroon



D.3.b. Indirect methods based on data from censuses and surveys

Data on population by age can also be used to indirectly estimate data fertility. Reverse survival of the population of children and women, based on estimates of mortality, combined with estimated patterns of fertility can be used to estimate fertility for 1-15 years before the census or survey. Regression methods, such as those developed by Rele (1967), do something similar based on the assumption that the population is stable. Estimates of fertility can also be made by trying to fit a stable population model to the age distribution in conjunction with estimates of mortality. These methods are especially useful in estimating fertility for periods in the past when other data are not available and when the population may, in fact have been close to a stable population.

The cohort-completed fertility method (Feeney, 2014) can be used to get estimates of TFR for about 12 to 23 or more years before a census or survey (depending on the availability of the data by age of women). Such method is useful for getting historical estimates when there are not many sources to work with.

The own-children method (Cho, 1973) is a much more involved method that requires access to the micro-data from a census or survey. In this method women are matched with their "own" children and both are reverse-survived back in time to obtain estimates of births and the female population. The method adjusts for children not matched to mothers.

D.4. Adjustments to estimates of fertility

Data from vital registration systems, censuses and surveys can often be defective, so methods have been developed to try to adjust the estimates. In general, data on CEB should first be examined to see if the El-Badry method (1961) should be used to estimate the proportion of women belonging to the category "parity not stated" who should have been classified as childless (see Diagram Fert-1). Figure Fert-2 shows the possible steps for evaluating and adjusting the fertility estimates based on the average children ever born.

D.4.a. Brass-type P/F ratio methods

Brass (1975) developed the P/F ratio method to adjust fertility data. The basic idea is that women (especially younger women) are better able to report the number of children they have had than specify the number born in a particular time period (e.g., the year before the survey). The average CEB by age (denoted P) are compared to the cumulated ASFRs (F) to create P/F ratios that can be used to adjust the reported ASFRs. One problem with this method is that it implicitly assumes that fertility has been constant in the recent past.

When there are reported CEB data 5 or 10 years apart then they can be used to compute synthetic cohort CEBs that can be used to compare to the reported ASFRs in the period prior to the second set of data, allowing a relaxation of the stable fertility assumption, but conditional to the reliability of each of these cross-sectional data sources.

D.4.b. Arriaga method

Arriaga tried to relax the assumption of constant fertility by comparing the CEB at two time points. This can also result in biased results (Johnson and Lollock, 2007).

[Action item: Further research on this method would be helpful. I think a variation of this method could be used to allow application of the synthetic cohort P/F ratio method. By estimating CEB-consistent ASFRs, it may allow for age-specific estimates of adjustment factors, like the ΔP/ΔF ratio method.]

D.4.c. Relational Gompertz model

The relational Gompertz model was developed to provide a flexible way of modeling the fertility curve (ASFRs). The model can be used with 5-year ages or single years of age. It has been used to produce an alternative version of the P/F ratio method (Moultrie, et al., 2013). As is stated in that source:

> In this manual all methods of fertility estimation that make use of the P/F ratio method are recast here to use the relational Gompertz model.

Although these methods based on the relational Gompertz can be useful, they basically replace the input data with the fitted model, so the assumption is that all fertility profiles can be fitted using the model that has a fixed standard. A more generic version of the model would allow for the user to specify their own standard pattern, and would also give the user the choice of using the input ASFRs to be adjusted based on the analysis.

D.4.d. Additional methods for evaluating the completeness of birth registration

In addition to the methods used to evaluate and adjust fertility estimates based on data from censuses and surveys, the UN Manual X (UNPD, 1983) presents a series of methods for comparing data on births by age from the vital registration system to data on children ever born by age of mother in a census (see Diagram Fert-4). These methods require more data than many other methods, but there are potential benefits to these methods, such as the ability to measure the completeness of birth registration by age (at least up to about age 30-34). One of these methods, cohort parity comparison with vital registration data, is included in the **TDE** tools as **FE_CohortVR_1.XLSX**.

[Action item: Develop tools to implement these methods of birth registration evaluation, and compare results to other, simpler methods. A new R package called "fertestr" developed by Lima and da Silva will implement most of these direct and indirect assessment methods https://github.com/josehcms/fertestr" https://github.com/josehcms/fertestr]

D.5. Review of results and produce final estimates

Once the data have been collected and alternative estimates developed, the analyst must put together the full set of estimates of ASFRs from 1950-2020. Ideally these estimates should be by single years of age and cover the full range of ages where women give birth.

D.5.a. Estimate trends in the TFR

The analysts need to review the TFR estimates from different sources, and estimation methods as applicable, for the full range available years. From this set of estimates a sequence of TFR values should be estimated based on the data, giving differential analytical weights (e.g., a loess or spline regression that takes into account the various types of data characteristics such as the type of method used to derive the estimate and the recall period) or including only estimates that are felt to accurately reflect the true level. A Bayesian approach has been proposed by Alkema, et al. (2012), and the GBD (2018d) uses a spatiotemporal Gaussian process regression.

D.5.b. Estimate the trends in 5-year ASFRs

Similar time-series analysis of the ASFRs should be done to get annual estimates, making sure that the relationships between adjacent ages and the time trends remain reasonable. These ASFRs should also be converted to annual proportional ASFRs (PASFRs). These can then be multiplied by the estimates of TFRs to get additional 5-year ASFR series. The full set of ASFRs should then be reviewed for consistency over time, and review the consistency with CEB measures over time.

D.5.c. Estimate single year ASFRS

Although the relational Gompertz model can produce single year ASFRs, it does not necessarily reproduce the 5-year ASFRs, so the Schmertmann spline approach has been adopted as a reasonable way to subdivide the 5-year ASFRs into single ages (Schmertmann, 2012). He states that "... ages that may extend outside the range of the original data (for example, below age 15 or above age 50)." This may be useful since methods used to estimate fertility usually ignore fertility below age 15 or ages 50+.

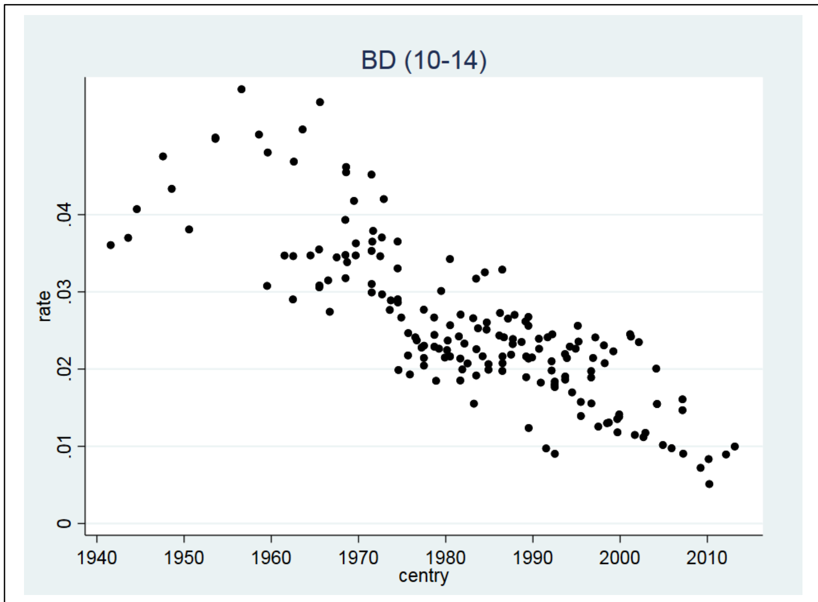
Schmertmann's model can be recalibrated using ASFRs by single age from good VR countries or computed from demographic surveys with full birth history.[5]

Pullum, et al. (2018) estimated fertility rates at ages below 15 from DHS surveys using the full birth history information for those in the youngest ages. They found that ASFRs for ages 10-14 ranged as high as 9 or 10 per 1000 women. The surveys analyzed were all after the year 2000, so the significance of this for earlier periods, with higher fertility, is not yet known. My concern with under 15 fertility is the impact it might have on P/F type fertility methods where the under-15 fertility is not included in the "F" estimates, but is included in the "P" part.

Other sources of ASFRs under age 15 include the UN Demographic Yearbook and the Human Fertility Database (2020). UNPD plans to use this info by single age (typically from age 12 onward). The computation of 10-14 fertility rates for all demographic surveys with full birth history information (WFS, DHS, MICS, etc.) provide results that seem to rather coherent (see the Figure D-2 below for Bangladesh).

[5] UNPD contracted Bruno Schoumaker to compute the ASFRs for the 10 years prior a survey for all demographic surveys (WFS, DHS, MICS, etc.) with full birth history information.

Figure D-2. Estimated of ASFR 10-14 for Bangladesh



E. Migration

E.1. Overview

Migration is the most difficult part of the demographic balancing equation to study. This is partly because migration can happen at any age and that it can happen multiple times over an individual's lifetime.

E.2. Data sources

E.2.a. Direct migration records

Some countries have systems for recording migration. For example, many countries keep records of legal immigrants. If there are not records of emigration, then the story will be incomplete. Most OECD countries, especially within the European Union, publish official annual estimates that can be used as initial information to estimate international migration. In addition of these estimates, Eurostat computed also an extra component called "statistical error" (or error of closure) to is required for the balancing equation.

Arrival/departure records are collected in many countries, but these are not intended as records of migration and the numbers are often overwhelmed by tourist movement.

Population censuses also include sometimes questions on migration (e.g. residence 5 years ago, former members of each household who had emigrated abroad, etc.) that can serve to provide estimate of international migration.

E.2.b. Refugees and asylum-seekers

Refugee movements are often sudden, and they vary considerably in how soon they return to their country of origin (if ever). Luckily, the UN High Commissioner for Refugees (UNHCR) collects and publishes this data on a regular basis.

[Action item: I have interpreted the declines of refugees and asylum seekers from a country as indicating their return to their country of origin. How are those that settle permanently in a new country treated in the UNHCR data? Most of the refugees and asylum seekers return back to their country of origin, but for those that get resettled or get integrated into the country of asylum, once they get the citizenship of that new country they also get removed from the refugee stock. In OECD countries it is typically within 5 to 10 years after arriving in that country. This means that some net changes in refugee stocks do not represent returns to the origin country, but rather a change in status in their destination country.]

Based on UNHCR annual stock data (end of year) available by country of origin and country of asylum, UNPD computes marginal totals and annual differences are computed from the stock to approximate net flows.

E.2.c. Foreign-born or non-citizen populations

Censuses often collect data on the foreign-born or non-citizen population. If the tables of these data give the countries of origin or citizenship, this can help estimate migration for those countries (as well as the country where the census was taken). Additional detail by age and sex will also be helpful in tracking the levels of migration by age and sex. Several researchers have been estimating net migration flows from such information, with Azose and Raftery building on Guy Abel's approach to get less biased estimates (Azose and Raftery, 2019 and Abel and Cohen, 2019). See version 2 (UNPD, 2019), with updated figures based on WPP 2019 which includes a comparison of the differences between WPP 2017 and WPP 2019 net migrations and the impact on the results.

E.2.d. Residual migration estimates

For many countries there are few useful data on migration, so the migration is often estimated as a residual. This amounts to simply rearranging the terms of the balancing equation to solve for NM(t). In general, the procedure is to project forward the population from one census to another based on estimated fertility and mortality in the intercensal period, and compare the results to the second census. There are several problems with this approach:

- There can be inconsistencies in the two censuses in terms of the completeness of enumeration that may not have been corrected fully
- The measures of numbers of net migrants are for Lexis parallelograms rather than squares, and they cover the whole intercensal period rather than individual years or periods.
- When the migrants are assumed to move within the period is another issue (see UN Population Branch, 1956, p. 59). When working with 5-year methods, the assumption made when the migrants move have different impact on the population. Assuming that all migrants are moving at the beginning of a 5-year period, at the end of the period, or evenly spread out over the period, has implications for the population submitted to the mortality and fertility rates. When working on an annual basis, this issue bears less importance.

The **NewPAS** workbook **ResidualMigBetaZA.xls** uses the results of a **RUP** projection to try to deconstruct the intercensal migration into single years of age and time assuming a constant pattern by age and sex over the period. The workbook then tries to simulate rerunning the **RUP** projection to see how well the model does in reproducing the later census.

E.3. Migration models

Since we sometimes only have a general idea about the total net flow of migrants, it can be useful to use models to distribute these migrants by age and sex based on information about the type of migration (e.g., labor, family, refugee). The Rogers and Castro model (1981) has often been used, but the full model has a number of parameters (see **RogersCastro.xlsx**), so some simplified versions (e.g., UNPD_**Migration Age Patterns.xls**) have been developed as well. One question is whether it is worthwhile to try to fit the model to residual migration estimates. The model does generate migration estimates by single years of age, which is useful for the new population model to be used for WPP 2021.

E.4. In-migration and out-migration vs. net migration

Since the data on migration is very sparse for most countries, the best we can usually do is compute some levels of net migration. The problem with net migration is that it does not necessarily follow a nice neat pattern by age. In countries with relatively large migration flows in both directions it may be worthwhile to model these separately. This is probably particularly true for countries like those in the Middle East with large non-citizen labor flows. If the net migration by age has positive values at some ages and negative at others, then multiplying the distribution by a constant in order to increase the overall level, say doubling it, will double both the net positive and negative numbers. It is far more likely, if the total migration is positive, that either there was an increase in the in-migrants or a decrease in out-migrants, which may result in more of an upward shift of the net migrants by age. This approach can be important in countries where the overall net migration is close to zero, but there are substantial offsetting in and out migrants at different ages that can differentially affect the age distribution.

E.5. Short-term projections of migration

Since residual migration and other estimates based on census data only allow updated estimates every 10 years or so, standards are needed for how to bring the estimates up through the year 2019. At the Census Bureau, when I was there, in cases where the levels have been fairly constant for a long period, we would create an average over a reasonable period of time 5-20 years, and assume that rate for the near term. If there has been a sudden shift, an evaluation should be made of whether this is a permanent change or a reaction to temporary conditions (in the country or in nearby countries).

F. Produce Reconciled Estimates 1950-2020

F.1. Estimate 1950 base population

Once the census population data have been adjusted/corrected and estimates have been developed for fertility, mortality, and migration, the 1950 base population must be estimated. This can be done using "reverse projection" from later censuses (or forward projections from earlier censuses), but that has limits due to the lack of specificity of the older ages (e.g., projecting back from a population with an open-ended age group of 100+ in 1960 to 1950 will result in an open-ended age group of 90+). Projecting back later censuses would have lower and lower open-ended age groups. These issues can be resolved using models of the older population discussed earlier.

If the interval from 1950 to the first census is long or the quality of the early censuses is in doubt, an alternative is to develop a stable or quasi-stable model of the population. For stable populations, all that is needed is a life table, a sex ratio at birth, and a growth rate or ASFRs. This is implemented in the **NewPAS** workbook **StablePopSingleAge.XLS**.

For a quasi-stable model, the idea is to start at a year when it is felt that the population was stable and then project the population to 1950.

F.2. Population reconstruction

One proposed way to estimate the base population is the development of a Bayesian hierarchical model of population reconstruction (Wheldon, 2016). This type of model takes initial estimates of the base population, fertility, mortality, and migration as well as estimates of uncertainty and determines the distribution of possible estimates of each of the parts. Although this method is appealing as a way to get the parts to be consistent, there are some issues:

1. The method is rather complex, despite being based on the same fundamental principle than for the deterministic reconstruction. The population reconstruction is done through simulation with sampling and resampling of the underlying statistical distributions of each demographic component and populations. Ultimately only a subset of combinations of the various input parameters lead to internally consistent estimates over time, by age and cohorts.
2. The method requires the development of priors' distribution that should be driven by observed data.
3. Need to check on the measures of "goodness of fit".
4. The applications that have been presented in papers so far have been for populations by 5-year age groups. Wheldon applied the approach to China data with 1x1 analytical framework, including using education statistics as auxiliary population count by age and sex. The approach provides excellent results reconciling the various data if we assume that the education statistics are more accurate than the censuses and surveys in terms of fertility and children enumerated during the 1 child policy period.
5. The initial estimates (priors) for each demographic component should be based as much as possible on independent data sources to evaluate the various interrelationships between them through the cohort-component framework and the demographic balancing equation over time, and cohorts. The use of estimates based on intercensal population comparisons to inform the

initial set of estimates should be considered with caution if the same population information also gets used to derive indirect fertility or mortality estimates.[6]

6. Demographers have some ideas (i.e., priors and models) about the age pattern or "shape" of the various functions: mortality, fertility, and migration. How does the method control the results to follow these?
7. It is not clear to me how the method will deal with certain types of discontinuities (e.g., excess mortality (due to wars, civil conflict, or natural disasters, or sudden refugee movement).

A version of this type of analysis is also used by the GBD (2018b). In the IHME approach, a distribution of population estimates gets generated based on the initial distribution of population estimates (with assumed errors) and prior estimates of mortality and fertility time trends estimates – but these fertility and mortality estimates are not reconciled with the populations (i.e., no posterior distribution of fertility and mortality get computed back). It is somewhat unclear in this approach how much the migration component is used not only for residual migrations, but also statistical error of closure with the balancing demographic balance relationships.

F.3. Combine 1950 base population with estimates of fertility, mortality, and migration

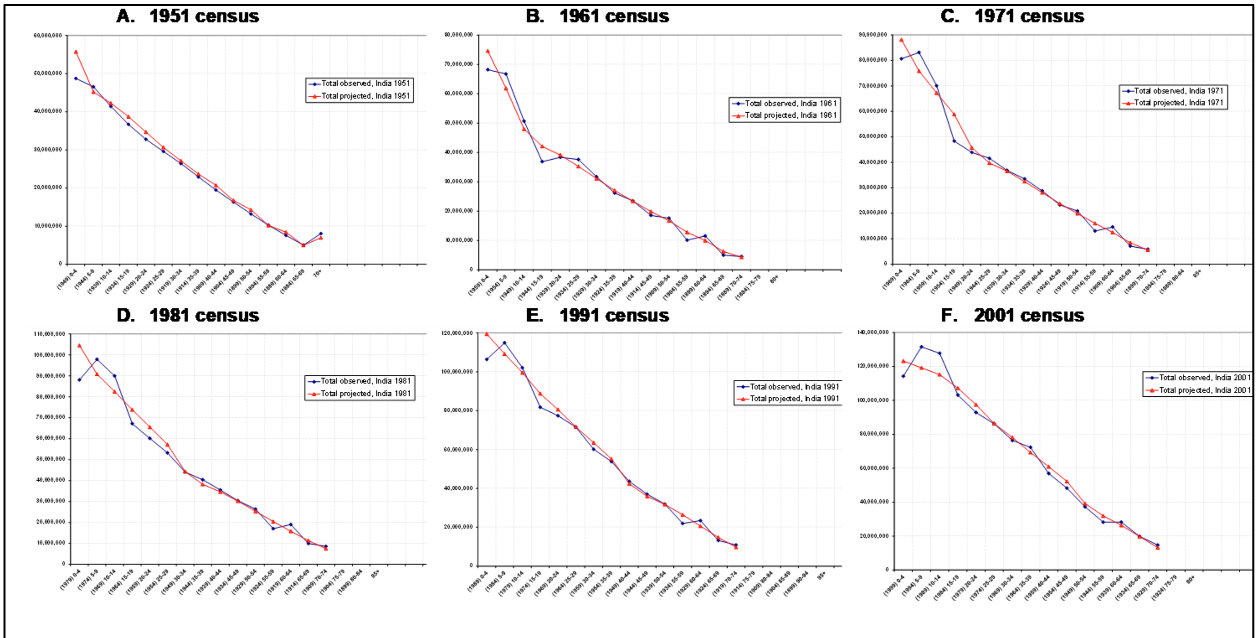
The cohort-component population projection (CCPP) method is then used to project the 1950 base population, by age and sex using the estimates of fertility, mortality, and migration. One of the first decisions that needs to be made for WPP 2021 is whether the 1950 base will be January 1 (beginning of year) or midyear. For this round the base population will be by sex and single years of age up to 100+. The CCPP method will also require mortality and migration data by single years of age, but a decision will need to be made whether abridged data could be entered for these and have a procedure for splitting built into the CCPP model. Although ASFRs by single years of age are not required by the 1x1 CCPP model, they do tend to smooth out the resulting births. If it is desired to allow 5-year ASFRs as input, they can be used as a step function to apply to the single-year female population, but splitting using the Schmertmann (2012) method could be used on demand.

F.4. Compare estimates to unadjusted and adjusted/smoothed census or survey data

The results of the projection should be compared to the unadjusted and adjusted/smoothed population data from census or survey data to make sure they are consistent. Gerland (2014) presented comparisons for India to several censuses (seeFigure F-1).

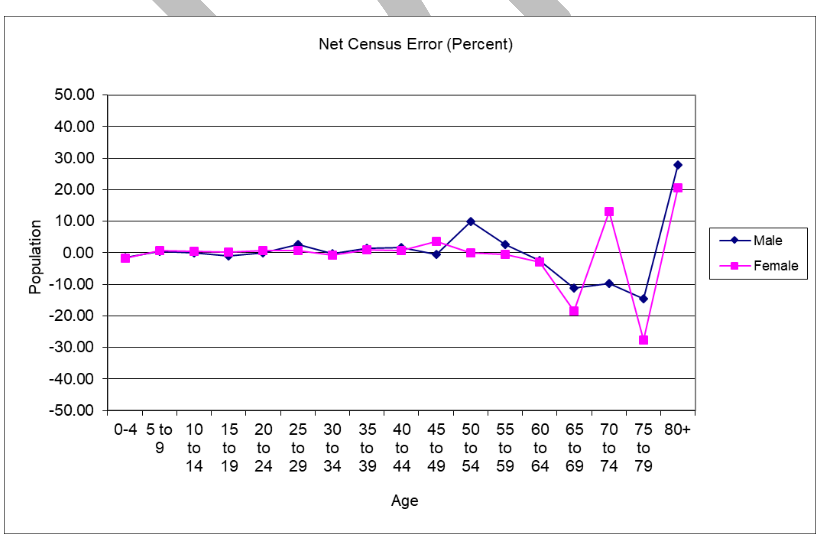
[6] This applies for instance to fertility from reverse survival or mortality from intercensal population survival. If the results given by these methods are used to derive estimates of fertility or mortality, the former need to be treated as non-informative or be used more for validation purposes. The death distribution methods (DDM), however, only use population estimates to compute completeness ratios for death registration, and therefore such information can more easily be used in this context, and can still benefit from a more integrated evaluation through multiple censuses cohort evaluation than single adjacent pairs as typically done with all DDM applications.

Figure F-1. Comparisons of Population Reconstruction to Census Data by Age for India



Comparison to the unadjusted data can also be used to compute percent differences by age and sex (similar to the **NewPAS** workbooks **RUPCEN.xls** and **RUPCENS.xls**). These workbooks treat the projection as the adjusted population and display the deviation of the census populations from the projections as "net census errors." These comparisons are often best viewed graphically (see figure below for Jordan 1994). In this case the fit is quite good for most ages, but starting at age 50-54 for males and 65-69 for females there seems to be evidence of age heaping and age overstatement.

Figure F-2. Implied Net Census Error From NewPAS Workbook RUPCEN



If the migration inputs to the CCPP model are not correct, some of the differences may be due to migration that is not fully accounted for in the model.

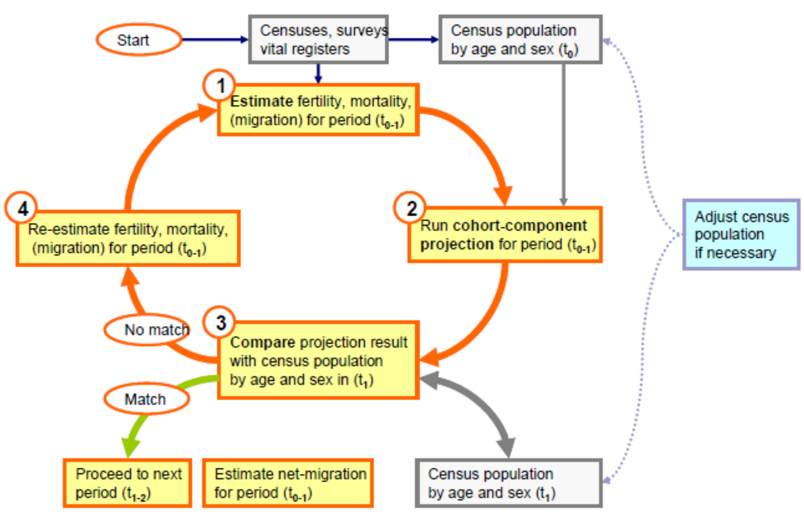
[Action item: A procedure should be developed to plot the percent differences or net census errors by age from different censuses by birth cohort.]

F.5. Determine if changes are needed

Significant differences between the population model and censuses for certain birth cohorts over time may point to changes needed in the fertility estimates, mortality estimates, or migration estimates. In the WPP 2019 Methodology report (UNPD, 2019) this iterative process was summarized in the flow diagram reproduced as Figure F-3.

Figure F-3. WPP Process to Ensure Consistency

Figure I.1. Process used to ensure intercensal consistency between demographic components and populations



NOTE: The diagram above illustrates how individual estimates of fertility, mortality and net-migration were subjected to tests of internal consistency using a cohort-component projection framework for the period between t_0 and t_1. This procedure has been applied in each new revision of the *World Population Prospects*. Past estimates are re-evaluated when new information becomes available; therefore, with every revision past demographic trends may be adjusted.

This entails circling back to the various component estimates to see where "tweaks" to the estimates might improve the consistency.

G. Specifications of the general analysis process, including documentation

One of the most difficult tasks in creating a set of demographic estimates is to accurately describe and document what sources were used and what methods were used to create the final product. Ideally, this documentation should also include indications of methods that were tried but rejected.

In the era of using spreadsheets, some people tried using linked workbooks to help show the relationships between observed data and estimates. These, however, often proved to be problems of the links getting broken (especially if files are moved). I have also tried to use a workbook (TableListJ.xls) that tracks the tables/files used in the analysis, but the system did not have a method for automatically showing what tables linked to other tables.

A possibility in this regard is to create scripts in R that show the steps and parameters that were used. This raises the question of how and where basic data sources are stored.[7] The range of demographic resources available in R is already large and increasing fast. However, there are some problems with different packages expecting data in different formats, so additional utilities may be useful. This would include programs to read data from or write to (specific) workbooks or databases.

Standards for naming files and/or variables may also be needed to help in keeping track of the data and processing.

Currently, the internal Eagle front-end and DemoData SQL database provide a lot of information, and utilities for extracting data from these would be helpful. The "Short notes" system provides a very good start in the attempt to standardize the linking of methods and sources, and something like it should be maintained.[8] Data sources and related meta-information for the World Population Prospects 2019 are available for each country or area in textual format[9] and can be downloaded in structured tabular format[10]. The textual information for each country is dynamically generated based on the respective coding of the various data sources and methods used for each demographic component.

[Action item: I need to summarize some observations about the "short notes."]

[7] UNPD store the basic data in a SQL server database (DemoData - https://population.un.org/DemoData/web/) which will be accessible thru an API (including with an R wrapper called DDSQLtools as companion to DemoTools), and will provide a web dissemination front-end thru a DataPortal (SQL database and front-end with visualization - https://population.un.org/dataportal/).

[8] For WPP2021, the data use status will be fully integrated with the empirical series in the SQL database, i.e., if series are selected as “used” they’ll be included in the statistical modelling. Series that are only “considered” will be either excluded or given a very low analytical weight.

[9] https://population.un.org/wpp2019/DataSources/

[10] https://population.un.org/wpp/Download/Metadata/Documentation/

H. REFERENCES

Abel, G.J. (2019). Update for WPP2019. https://guyabel.com/publication/bilateral-international-migration-flow-estimates/

Abel, G.J., and Cohen, J.E. (2019). Bilateral international migration flow estimates for 200 countries. *Scientific Data*, 6:82, 575-594. https://www.nature.com/articles/s41597-019-0089-3.

Azose, J.J. and Raftery, A.E. (2019). Estimation of emigration, return migration, and transit migration between all pairs of countries. Proceedings of the National Academy of Sciences of the United States of America (PNAS), 116:1, 116-122. https://www.pnas.org/content/116/1/116.

Alkema, L., Raftery, A., Gerland, P., Clark, S. J., and Pelletier, F. (2012). Estimating trends in the total fertility rate with uncertainty using imperfect data: Examples from West Africa. *Demographic Research*, 26(15), 331-362. doi: 10.4054/DemRes.2012.26.15.

Andreev, E.M., and Kingkade, W.W. (2015). Average age at death in infancy and infant mortality level: Reconsidering the Coale-Demeny formulas at current levels of low mortality. *Demographic Research*. 33(13), 363-390.

Bachi, R., (1951). The Tendency of Round Off Age Returns: Measurement and Corrections. *Bulletin of the International Statistical Institute*. 33(4), 195-222.

Bachi, R., (1953). Measurement of the Tendency of Round Off Age Returns. *Bulletin of the International statistical Institute*. 34(3), 129-137.

Bennett, N. G., and Horiuchi, S. (1981). Estimating the completeness of death registration in a closed population. *Population Index*. 47, 207-221.

Bennett, N. G., and Horiuchi, S. (1984). Mortality estimation from registered deaths in less developed countries. *Demography*. 21, 217-233.

Brass, W. (1975). *Methods for Estimating Fertility and Mortality from Limited and Defective Data*. Chapel Hill, North Carolina: Carolina Population Center, Laboratories for Population Studies.

Cho, L-J. (1973). The own-children approach to fertility estimation: an elaboration. *International Population Conference*. 2, 263-280.

Clark, S.J. (2019). A General Age-Specific Mortality Model With and Example Indexed by Child Mortality or Both Child and Adult Mortality. *Demography*. 56: 1131-1159.

Coale, A. and Guo, G. (1989). Revised regional model life tables at very low levels of mortality. *Population Index*. 55(4), 613-643.

Coale, A. and Kisker, E.E. (1990). Defects in Data on Old-Age Mortality in the United States: New Procedures for Calculating Mortality Schedules and Life Tables at the Highest Ages. *Asian and Pacific Population Forum*. Spring 1990. Pp. 1-31.

Coale, A. J. and Stephan, F. F. (1962). The Case of the Indians and the Teen-Age Widows. *Journal of the American Statistical Association*, 57(298), 338-347

Coale, A. J., Vaughan, B., and Demeny, P. G. (1983). *Regional Model Life Tables and Stable Populations*. New York: Academic Press Inc.

Condran, G. A., Himes, C.L., and Preston, S. H._(1991)._*Old-Age Mortality Patterns in Low-Mortality Countries: An Evaluation of Population and Death Data at Advanced Ages, 1950 to the Present*. Population Bulletin of the United Nations. 30.

Croft, T. N., Aileen M. J., Marshall, C., Allen, K., et al. (2018). *Guide to DHS Statistics*. Rockville, Maryland, USA: ICF.

El Badry, M.A. 1961. Failure of enumerators to make entries of zero: errors in recording childless cases in population censuses. *Journal of the American statistical Association*. 56 (296), 909-924.

Feeney, G. (2014). *The Population Census as a Time Machine*. Demography-Statistics-Information Technology, Letter No. 4, 15 January 2014. http://demographer.com/dsitl/04-population-census-as-time-machine/.

Feeney, G. (2013). Removing "Zigzag" from Age Data. http://demographer.com/white-papers/2013-removing-zigzag-from-age-data/.

Feeney, G. 1979. A technique for correcting age distributions for heaping on multiples of five. *Asian Pacific Census Forum*. 5(2) 12–14.

GBD 2017 Mortality Collaborators. (2018a). Global, regional, and national age-sex-specific mortality and life expectancy, 1950–2017: a systematic analysis for the Global Burden of Disease Study 2017. *Lancet*. 392, 1684–1735.

GBD 2017 Mortality Collaborators. (2018b). Methods appendix to "Global, regional, and national age-sex-specific mortality and life expectancy, 1950–2017: a systematic analysis for the Global Burden of Disease Study 2017." *Lancet*. 392, 1684–1735.

GBD 2017 Population and Fertility Collaborators. (2018c). "Population and fertility by age and sex for 195 countries and territories 1950–2017: a systematic analysis for the Global Burden of Disease 2017." *Lancet*. 392: 1995–2051.

GBD 2017 Population and Fertility Collaborators. (2018d). Methods appendix to "Population and fertility by age and sex for 195 countries and territories 1950–2017: a systematic analysis for the Global Burden of Disease 2017." *Lancet*. 392: 1995–2051.

Gerland, P. (2014). "UN Population Division's Methodology in Preparing Base Population for Projections: Case study for India", *Asian Population Studies*. 10(3), 274–303.

Gompertz, B. (1825). On the nature of the function expressive of the law of human mortality, and on a new mode of determining the value of life contingencies. *Philosophical Transactions of the Royal Society*. 115, 513–585.

Gray, A. (1987). The Missing Ages: Adjusting for Digit Preference. *Asian and Pacific Population Forum*. 1(2), 11-22.

Hill, K. (2017). Analytical Methods to Evaluate the Completeness and Quality of Death Registration: Current State of Knowledge. UN Population Division Technical Paper No. 2017/2. https://www.un.org/en/development/desa/population/publications/pdf/technical/TP2017-2.pdf.

Hill, K. (1987). Estimating census and death registration completeness. *Asian and Pacific Population Forum*. 1 (3). 8-13, 23-24.

Hill, K., You, D., and Choi, Y. (2009). Death distribution methods for estimating adult mortality: Sensitivity analysis with simulated data error. *Demographic Research*. 21, 235-254. http://www.demographic-research.org/Volumes/Vol21/9/.

Hill, K., and Trussell, T. J. (1977). Further developments in indirect mortality estimation. *Population Studies*. 31(2), 313-334.

Human Fertility Database. (2020). Max Planck Institute for Demographic Research (Germany) and Vienna Institute of Demography (Austria). www.humanfertility.org.

Human Mortality Database (HMD). (2019a). *About Mortality Data for Canada*. By K. Andreev and R. Bourbeau (subsequently updated by N. Ouellette and D. Dukhovnov). Last update January 17, 2019 by D. Dukhovnov (reviewed by R. Bourbeau). https://www.mortality.org/hmd/CAN/InputDB/CANcom.pdf.

Human Mortality Database (HMD). (2019b). *About Mortality Data for United States*. By M. Andreeva. Last Updated by M. Barbieri: October 2, 2019. https://www.mortality.org/hmd/USA/InputDB/USAcom.pdf

Human Mortality Database (HMD). (2019c) Data from the 1871 census of Italy. University of California, Berkeley (USA), and Max Planck Institute for Demographic Research (Germany). https://www.mortality.org/hmd/ITA/InputDB/ITApop.txt. (data downloaded on 11/26/19).

Human Mortality Database (HMD). (2017). *About Mortality Data for Italy*. By D.A. Glei. Updated and revised by G. Borges, T. Riffe, and M. Andreeva. Last Updated: December 20, 2017. https://www.mortality.org/hmd/ITA/InputDB/ITAcom.pdf.

Human Mortality Database (HMD). (2016). *Special Methods Used for Selected Populations*. Last Revised: April 21, 2016. https://www.mortality.org/Public/Docs/SpecialMethods.pdf

IPUMS International. (2019). Census questionnaire for the 1971 Census of Haiti (translated into English). https://international.ipums.org/international-action/source_documents/enum_instruct_ht1971a_tag.xml

Jdanov, D.A. (2015). Official population statistics and the Human Mortality Database estimates of populations aged 80+ in Germany and nine other European countries. *Demographic Research*. 13: 14, 335-362.

Johnson, P. and Lollock, L. (2007). *A Re-Evaluation of the Arriaga Fertility Method*. Poster presented at the Annual Meeting of the Population Association of America, New York, March 29-31, 2007.

Johnson, P.D., and Way, P.O., (2003). *The Role of Population Estimates and Projections in the Evaluation of Censuses*. International Programs Center, Population Division, U.S. Census Bureau. Paper presented at the 2003 Population Census Conference of the Association of National Census and Statistics Directors of America, Asia, and the Pacific in Kyoto, November 19-21, 2003.

The Latin American Mortality Database (LAMBdA) Team. (2019). *Methodological Document*, Version 2.1. January 21, 2019.

Li, N., and Gerland, P. (2013). *Using Census Data to Estimate Old-Age Mortality*. Paper presented at XXVII IUSSP International Population Conference. Busan, Korea.

Li, N., Lee, R., and Tuljapurkar, S. (2004). Using the Lee–Carter method to forecast mortality for populations with limited data. *International Statistical Review*, *72*(1), 19-36.

Li, N., and Lee, R. (2005). Coherent Mortality Forecasts for a Group of Populations: An Extension of the Lee-Carter Method. *Demography*. 42:3, 575-594.

Luther, N. Y., and Cho, L-J. (1988). Reconstruction of birth histories from census and household survey data. *Population Studies*. 42(3), 451-472.

Makeham, W. M. (1860). On the law of mortality and the construction of annuity tables. *Journal of the Institute of Actuaries and Assurance Magazine*. 8(6), 301–310.

Martin, L. (1980). A modification for use in destabilized populations of Brass's technique for estimating completeness of death registration. *Population Studies*. 34, 381-95.

Moultrie, T., Dorrington, R, Hill, A., Hill, K., Timaeus, I., and Zaba, B. (eds) (2013). *Tools for Demographic Estimation*. Paris: International Union for the Scientific Study of Population. http://demographicestimation.iussp.org/.

Myers, R. (1954). Accuracy of Age Reporting in the 1950 United States Census. *Journal of the American Statistical Association*, 49(268), 826-831

Myers, R. (1940). Errors and Bias in the Reporting of Ages in the Census Data. *Transactions of the Actuarial Society of America*. 41(2), 411-415.

Noumbissi, A. (1992). L'indice de Whipple modifié : une application aux données du Cameroun, de la Suède et de la Belgique. *Population*, 47(4), pp. 1038-1041.

O'Hare, W. P. (2019). *Differential Undercounts in the U.S. Census. Who is Missed?* Springer Briefs in Population Studies.

Palloni, A., and Heligman, L. (1985). Re-estimation of structural parameters to obtain estimates of mortality in developing countries. *Population Bulletin of the United Nations*. 18, 10-33.

Philippine Statistics Authority. (2019). *Total Number of OFWs Estimated at 2.3 Million (Results from the 2018 Survey on Overseas Filipinos)*. News Release Reference Number 2019-066. https://psa.gov.ph/statistics/survey/labor-and-employment/survey-overseas-filipinos.

Preston, S. (1983). An integrated system for demographic estimation from two age distributions. *Demography*. 20 (2), 213-226.

Preston, S., and Bennett, N. G. (1983). A census-based method for estimating adult mortality. *Population Studies*. 37 (1), 91-104.

Preston, S., Coale, A. J., Trussell, J. and Weinstein, M. (1980). Estimating the completeness of reporting of adult deaths in populations that are approximately stable. *Population Index*. 46 (2), 179-202.

Preston, S., and Hill, K. (1980). Estimating the completeness of death registration. *Population Studies*. 34, 349-366.

Pullum, T. W., Croft, T., and MacQuarrie, K. L. D. (2018). *Methods to Estimate Under-15 Fertility Using Demographic and Health Surveys Data*. DHS Methodological Reports No. 23. Rockville, Maryland, USA: ICF.

Rele, J. R. *Fertility analysis through extension of stable population concepts*. (1967). Population Monograph Series. 2. Berkeley, CA: Institute of International Studies, University of California.

Riffe, T. (2019). *DemoTools: Standardize, Evaluate, and Adjust Demographic Data*. https://rdrr.io/github/timriffe/DemoTools/.

Riffe, T. (2018). Recipe for smoothing population (R code). Personal communication from Patrick Gerland.

Robinson, J. G. (2010*). Coverage of population in Decennial Census 2000 based on demographic analysis: The history behind the numbers*. Decennial Census Bureau, Working Paper No. 91, available online at http://www.Census.gov/population/www/documentation/twps0091/twps0091.pdf.

Rogers, A. and Castro, L.J. (1981). *Model Migration Schedules*. Report RR-81-30. Laxenburg, Austria: International Institute for Applied Systems Analysis.

Schmertmann. C.P. (2012). *Calibrated Spline Estimation of Detailed Fertility Schedules from Abridged Data*. Max Planck Institute for Demographic Research Working Paper. 5 July 2012. http://schmert.net/calibrated-spline/docs/CS%20smoother.pdf.

Sharrow, D., Clark, S. and Raftery, A.E., (2014). Modeling age-specific mortality for countries with generalized HIV epidemics. *PLoS One*. 9(5), e96447. doi.org/10.1371/journal.pone.0096447

Siegel, J.S., and Swanson, D.A. (2004). *The Methods and Materials of Demography*. Second edition. San Diego, CA: Elsevier Academic Press.

Spoorenberg, T. (2007). Quality of age reporting: extension and application of the modified Whipple's index. *Population* (English Edition). 62(4), 729-741.

Stover, J., Johnson, P., Zaba, B., Zwahlen, M., Dabis, F., Ekpini, R.E. (2008). The Spectrum projection package: improvements in estimating mortality, ART needs, PMTCT impact and uncertainty bounds. *Sexually Transmitted Infections*. 84 Suppl 1:i24-i30. doi: 10.1136/sti.2008.029868.

Thatcher, A. R., Kannisto, V., and Vaupel, J. W. (1998). *The force of mortality at ages 80 to 120*. Monographs on Population Aging. 5, Odense.

Timæus I. M. (1991). Estimation of adult mortality from orphanhood before and since marriage. *Population Studies*. 45(3), 455-472.

Timæus I. M. (1991). Estimation of mortality from orphanhood in adulthood. *Demography*. 28(2), 213-227.

Timæus, I. M. (1992). Estimation of adult mortality from paternal orphanhood: a reassessment and a new approach. *Population Bulletin of the United Nations*. 33, 47-63.

Timæus, I. M., Zaba, B., and Ali, M. (2001). Estimation of adult mortality from data on adult siblings. Zaba, B., and Blacker, J. (eds). *Brass Tacks: Essays in Medical Demography*. London: Athlone, 43-66.

Trussell, T. J. (1975). A re-estimation of the multiplying factors for the Brass technique for determining childhood survivorship rates. *Population Studies*. 29 (1) 97-108.

UN Inter-agency Group on Child Mortality Estimation (UN IGME). (2019a). *Levels & Trends in Child Mortality*. Report 2019. UNICEF: New York.

UN Inter-agency Group on Child Mortality Estimation (UN IGME). (2019b). Interactive website. https://childmortality.org/data

UN Inter-agency Group on Child Mortality Estimation (UN IGME). (2014). *Levels & Trends in Child Mortality*. Report 2014. UNICEF: New York.

United Nations. (1952). "Accuracy tests for census age distributions tabulated in five-year and ten-year groups." in *Population Bulletin*, no. 2, ST/SOA/Ser.N/2, pp. 59-79. New York.

United Nations, Department of Economic and Social Affairs, Population Division (UNPD). (2019). *International migrant stock 2019*. https://www.un.org/en/development/desa/population/migration/data/estimates2/estimates19.asp.

United Nations, Department of Economic and Social Affairs, Population Division (UNPD). (2019). *World Population Prospects 2019: Methodology of the United Nations Population Estimates and Projections*. Working Paper No. ESA/P/WP.425. New York: United Nations.

United Nations, Department of Economic and Social Affairs, Population Division (UNPD). (2002). *Methods for Estimating Adult Mortality*. https://www.un.org/en/development/desa/population/publications/pdf/mortality/estimating-adult-mortality.pdf. ESA/P/WP.175. Sales no. E.83.XIII.2. New York: United Nations.

United Nations. Department of Economic and Social Affairs, Population Division (UNPD). (1983). *Manual X. Indirect Techniques for Demographic Estimation*. A collaboration of the Population Division of the Department of International Economic and Social Affairs of the United Nations Secretariat with the Committee on Population and Demography of the National Research Council, United States National Academy of Sciences. Population Studies, No. 81. New York: United Nations.

United Nations. Department of Economic and Social Affairs, Population Division (UNPD). (1982). Model Life Tables for Developing Countries. Population Studies, No. 77. ST/ESA/SER.A/77. Sales No. E.81.XIII.7.

United Nations. Department of Economic and Social Affairs, Population Branch. (1956). Manual III. Methods for Population Projections by Sex and Age. Population Studies No. 25. ST/SOA/Series A/25. Sales No. 56.XIII.3. United Nations: New York: United Nations.

UN Department of Economic and Social Affairs, Statistics Division (UNSD). (2020). *Demographic Yearbook System. Questionnaires*. https://unstats.un.org/unsd/demographic-social/products/dyb/#questionnaires.

UN Department of Economic and Social Affairs, Statistics Division (UNSD). (2017). *Principles and Recommendations for Population and Housing Censuses*. Revision 3. United Nations Publication ST/ESA/STAT/SER.M/67/Rev.3. United Nations: New York.

UN Department of Economic and Social Affairs, Statistics Division (UNSD). (2010). Post-Enumeration Surveys. Operational Guidelines. Technical Report. 2010 World Population and Housing Census Programme, United Nations: New York.

UN Department of Economic and Social Affairs, Statistics Division (UNSD). (1990). *Demographic Yearbook 1988*. New York.

U.S. Bureau of the Census. (1985). *Evaluating Censuses of Population and Housing*. Statistical Training Document. ISP-TR-5. Washington, DC.

U.S. Bureau of the Census. (1960). *The Post-Enumeration Survey: 1950. An evaluation of the 1950 Censuses of Population and Housing*. Technical Paper No. 4. Washington.

U.S. Bureau of the Census. (1953). *The Infant Enumeration Study 1950*. Procedural Studies of the 1950 Censuses. No. 1. Washington.

U.S. Census Bureau. (2019a). *Demographic Analysis & Population Projection System (DAPPS) Software*. https://www.census.gov/data/software/dapps.html.

U.S. Census Bureau. (2019b). *Population Analysis System*. https://www.census.gov/data/software/pas.html.

U.S. Census Bureau. (2014). *Rural Urban Projection (RUP) Software*. https://www.census.gov/data/software/rup.html.

U.S. Census Bureau. (2013). *International Data Base Population Estimates and Projections Methodology*. https://www2.census.gov/programs-surveys/international-programs/technical-documentation/methodology/idb-methodology.pdf.

U.S. Census Bureau. (no date). NewPAS unpublished workbooks.

Verhulst, A., (2017). Child mortality estimation: An assessment of summary birth history methods using microsimulation. *Demographic Research*. 34:39, 1075-1128.

Wheldon, M.C., Raftery, A.E., Clark, S.J., and Gerland, P. (2016). Bayesian population reconstruction of female populations for less developed and more developed countries. *Population Studies*. 70(1), 21-37.

Wilmoth, J.R., Andreev, K., Jdanov, D., Glei, D.A., and Riffe, T., with the assistance of Boe, C., Bubenheim, M., Philipov, D., Shkolnikov, V., Vachon, P., Winant, C., Barbieri, M. (2019). *Methods Protocol for the Human Mortality Database*. Last Revised: October 5, 2019 (Version 6). https://www.mortality.org/Public/Docs/MethodsProtocol.pdf.

Wilmoth, J.R., Zureick, S., Canudas-Romo, V., Inoue, M., and Sawyer, C. (2012): A flexible two-dimensional mortality model for use in indirect estimation, *Population Studies*. 66(1), 1-28.

The Wold Bank. (2020). Identification for Development (ID4D). https://id4d.worldbank.org/.

Zelnick, M. (1961). Age Heaping in the United States Census: 1880-1950. *Milbank Memorial Fund Quarterly*. 39(3), 540-573.

I. Appendix: List of Diagrams

Page
- Legend for symbols used.. 64

I.1. Population

- PopT-1: Time Series of Statistical Concepts Adjustment Factors 65
- PopT-2: Time Series of Geographic Adjustment Factors........ 66
- PopT-3: Time Series of Completeness Adjustment Factors 67
- PopT-4: Final Time Series Adjustment 68

- PopA-1: Analysis and Adjustment of Populations by Single Years of Age........ 69
- PopA-2: Analysis and Adjustment of Populations by Sex and 5-year Age Groups 70

I.2. Adult Mortality

- MortA-1: Adult Mortality Overview........ 71
- MortA-2: Adult Mortality Decision Table 72
- MortA-3: Registered Deaths by Age and Sex 73
- MortA-4: Adult Death Registration Decision Table 74
- MortA-5: Check Quality of Adult Deaths by Age and Sex 75
- MortA-6: Intercensal Adjustment of Deaths by Age and Sex........ 76
- MortA-7: Adult Mortality Validation and Review 77
- MortA-8: Old Age Assessment and Adjustment........ 78

I.3. Child Mortality

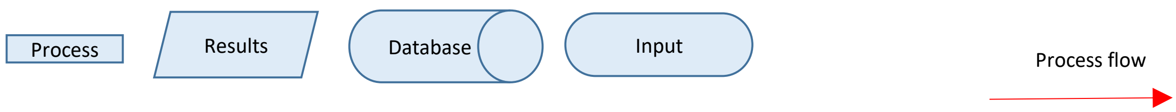
- MortC-1: Child Mortality Overview 79
- MortC-2: Child Mortality Registration Data 80
- MortC-3: IGME Adjustments and B3 Model........ 81
- MortC-4: IGME Notes........ 82

I.4. Fertility

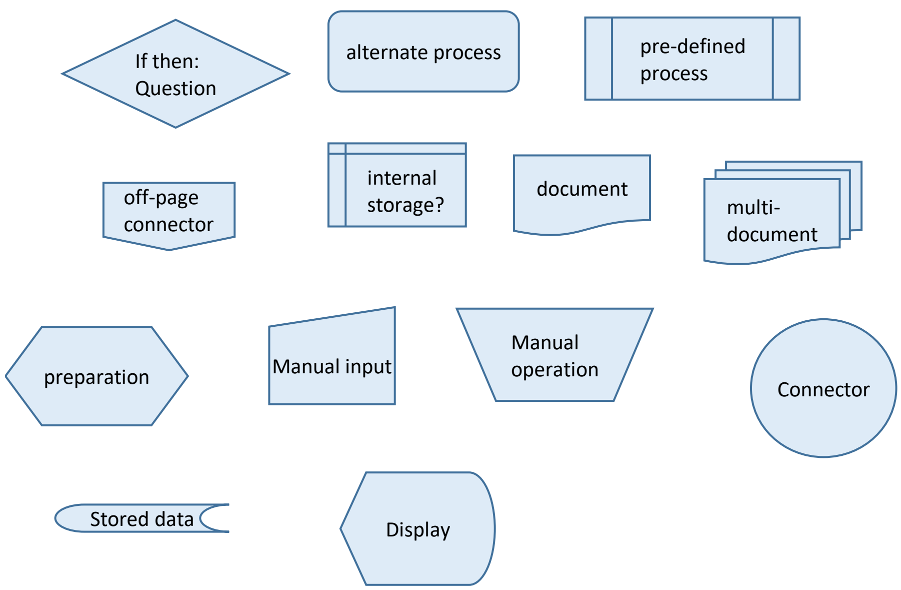
- Fert-1: Compute Fertility Estimates Based on Censuses and Surveys 83
- Fert-2: Evaluate and Adjust Fertility Estimates Based on Censuses and Surveys 84
- Fert-3: Compute Fertility Estimates Based on Registration Data 85
- Fert-4: Evaluate and Adjust Registered Fertility Compared to Data
  From Censuses and Surveys 86

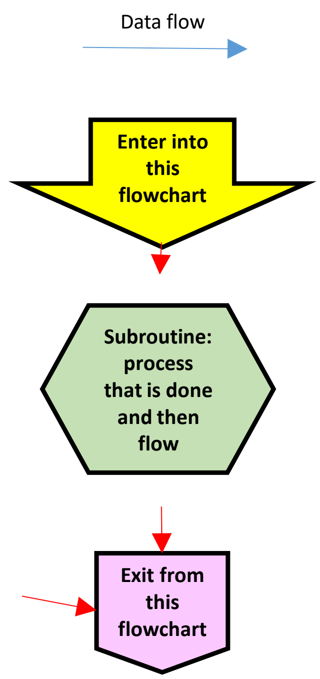
Legend for symbols used:

GBD symbols



Other symbols to consider





Grouped items

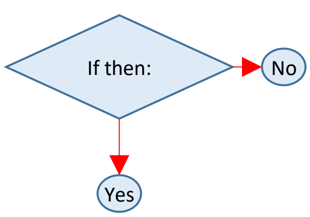


Diagram PopT-1: Time Series of Statistical Concepts Adjustment Factors

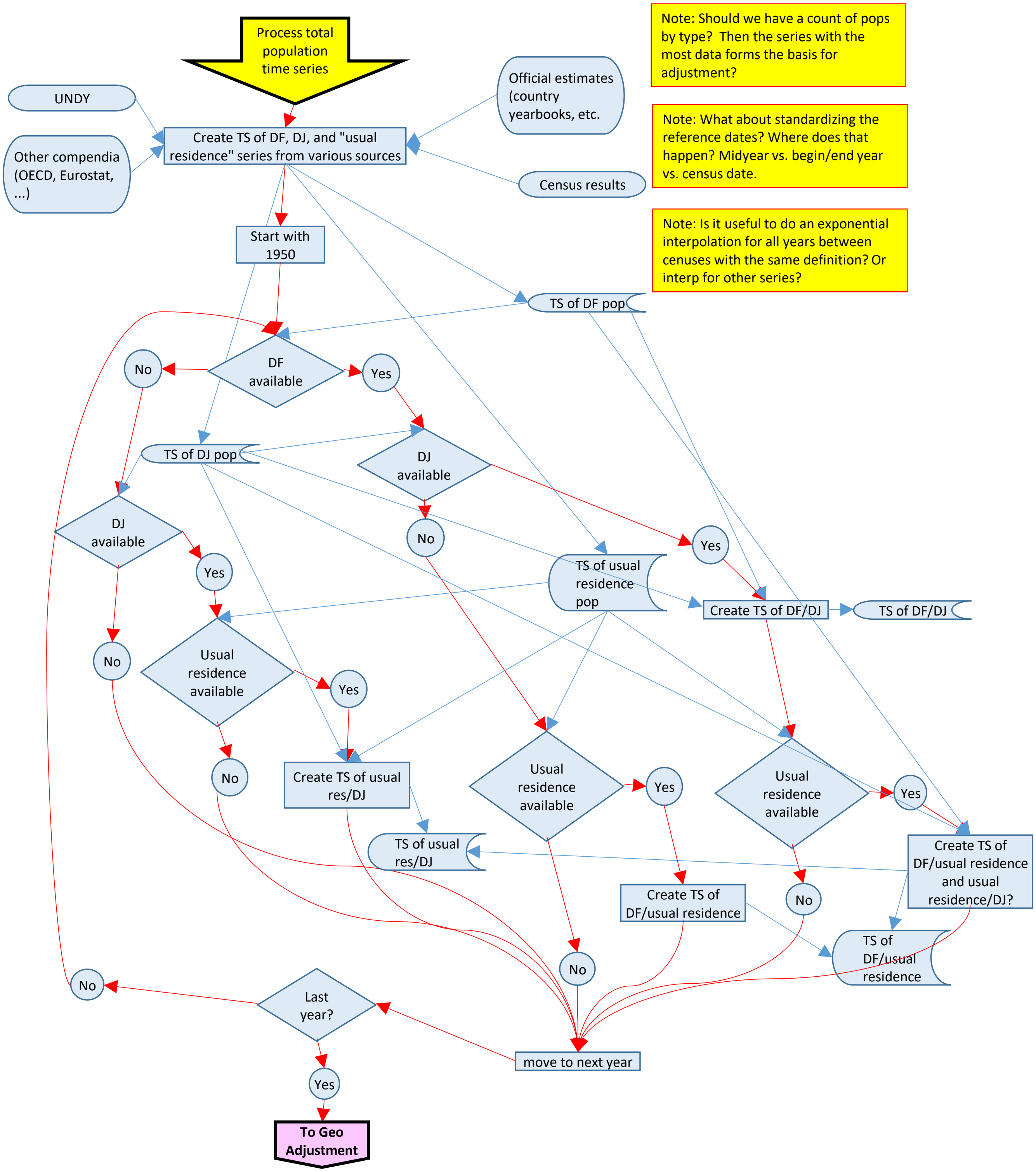


Diagram PopT-2: Time Series of Geographic Adjustment Factors

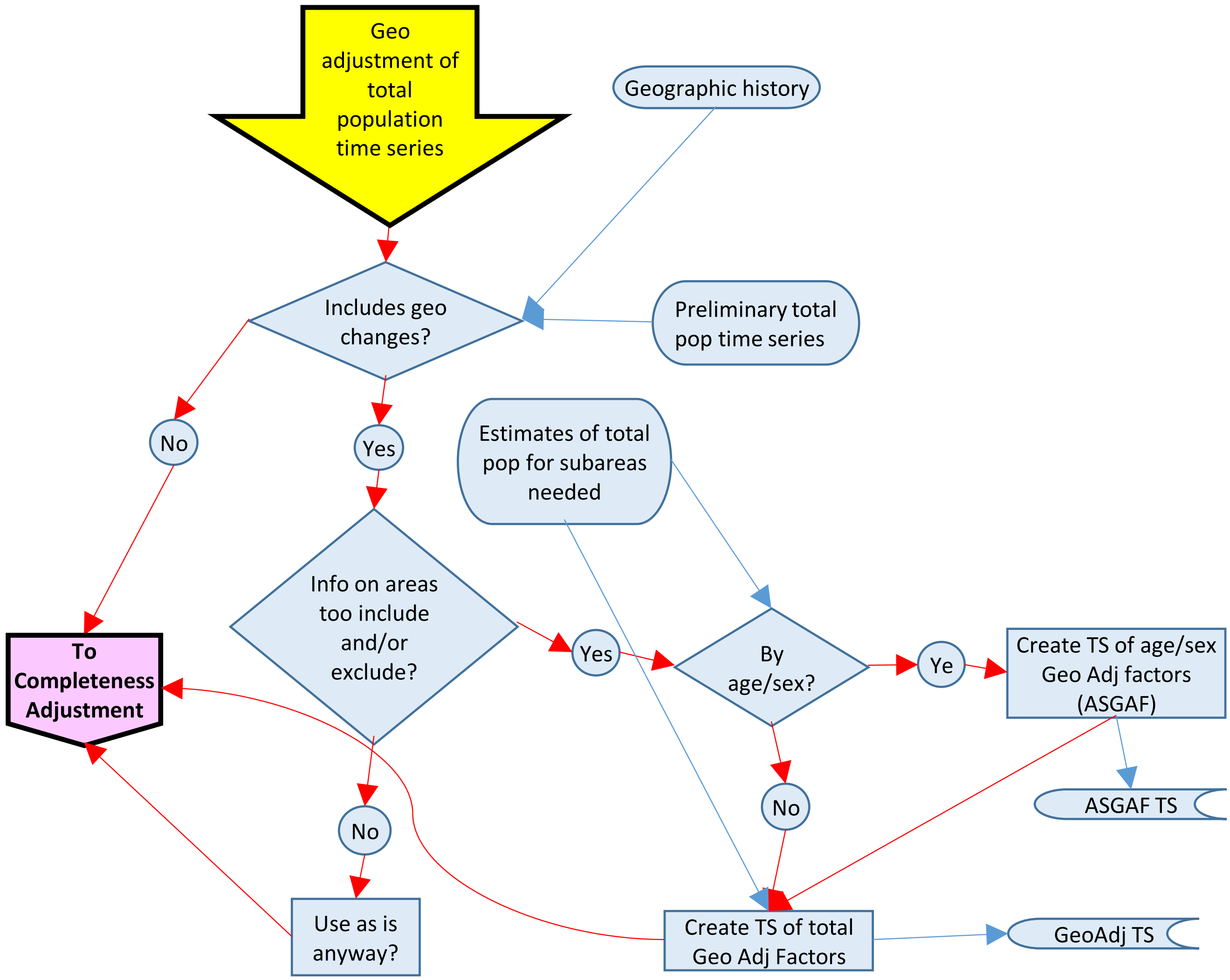


Diagram PopT-3: Time Series of Completeness Adjustment Factors

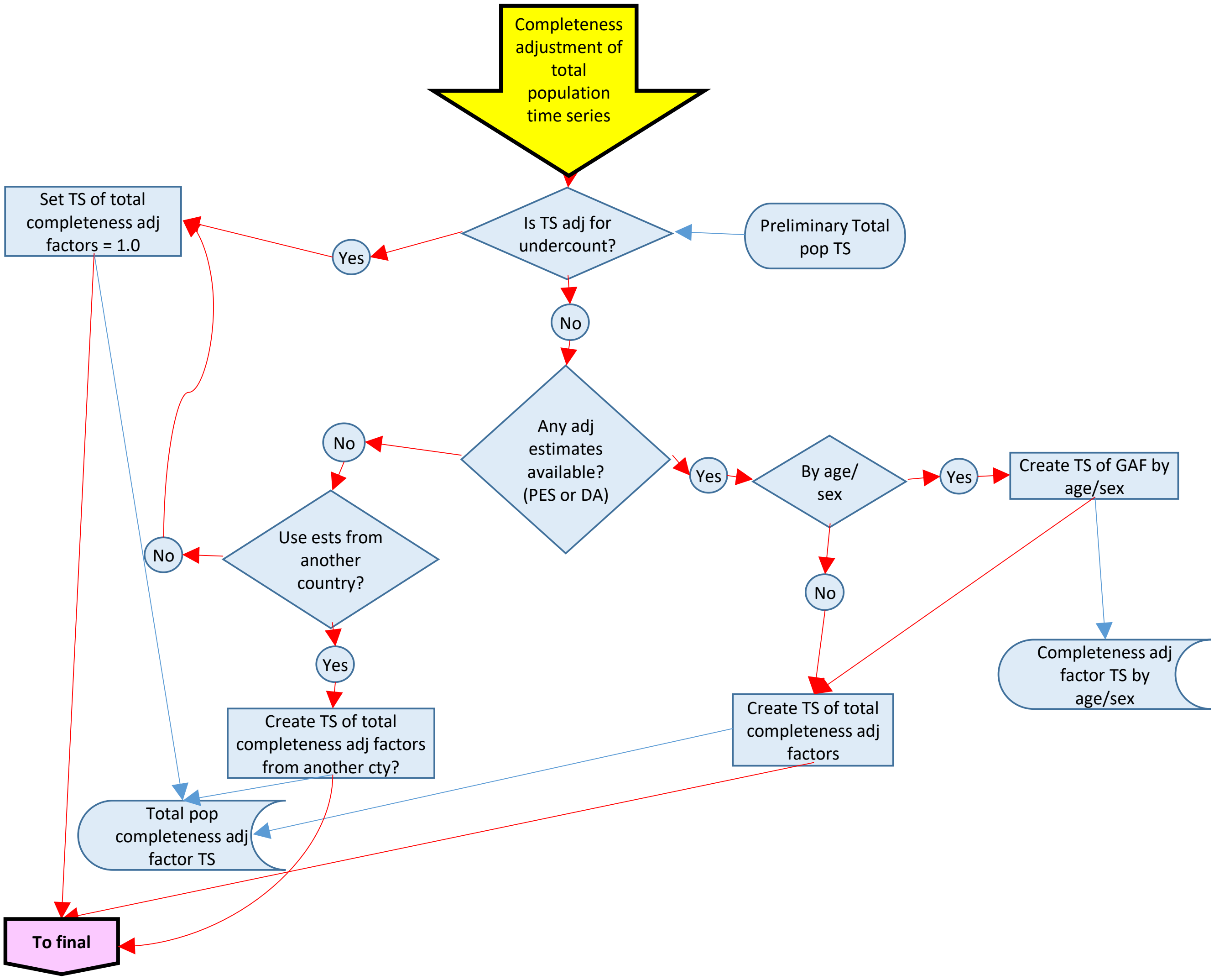


Diagram PopT-4: Final Time Series Adjustment

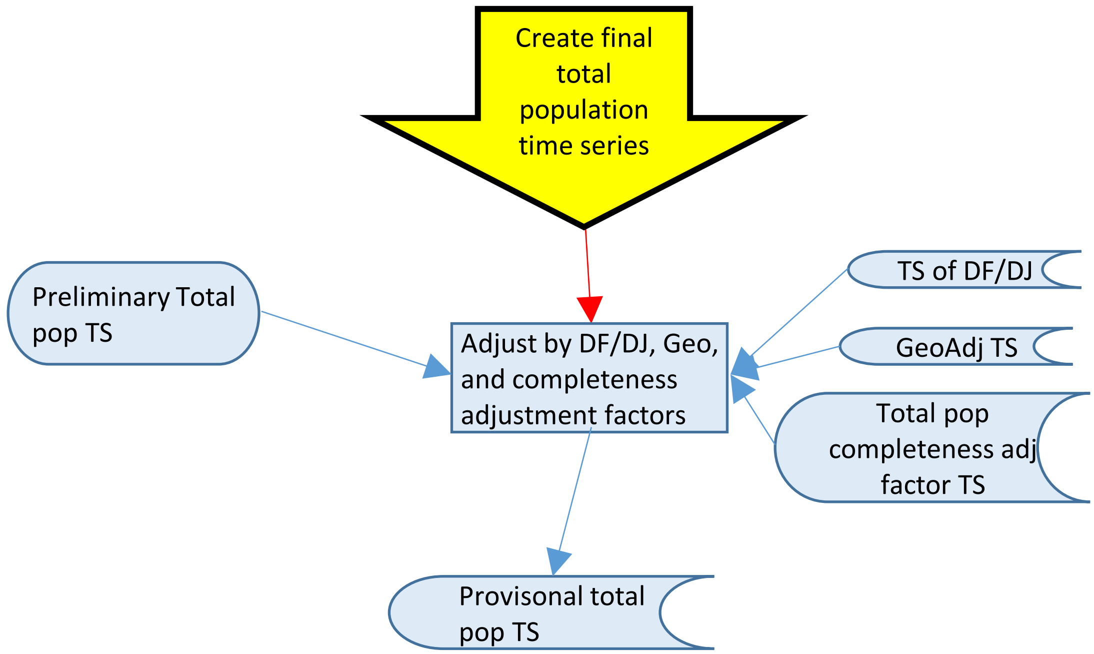


Diagram PopA-1: Analysis and Adjustment of Populations by Single Years of Age

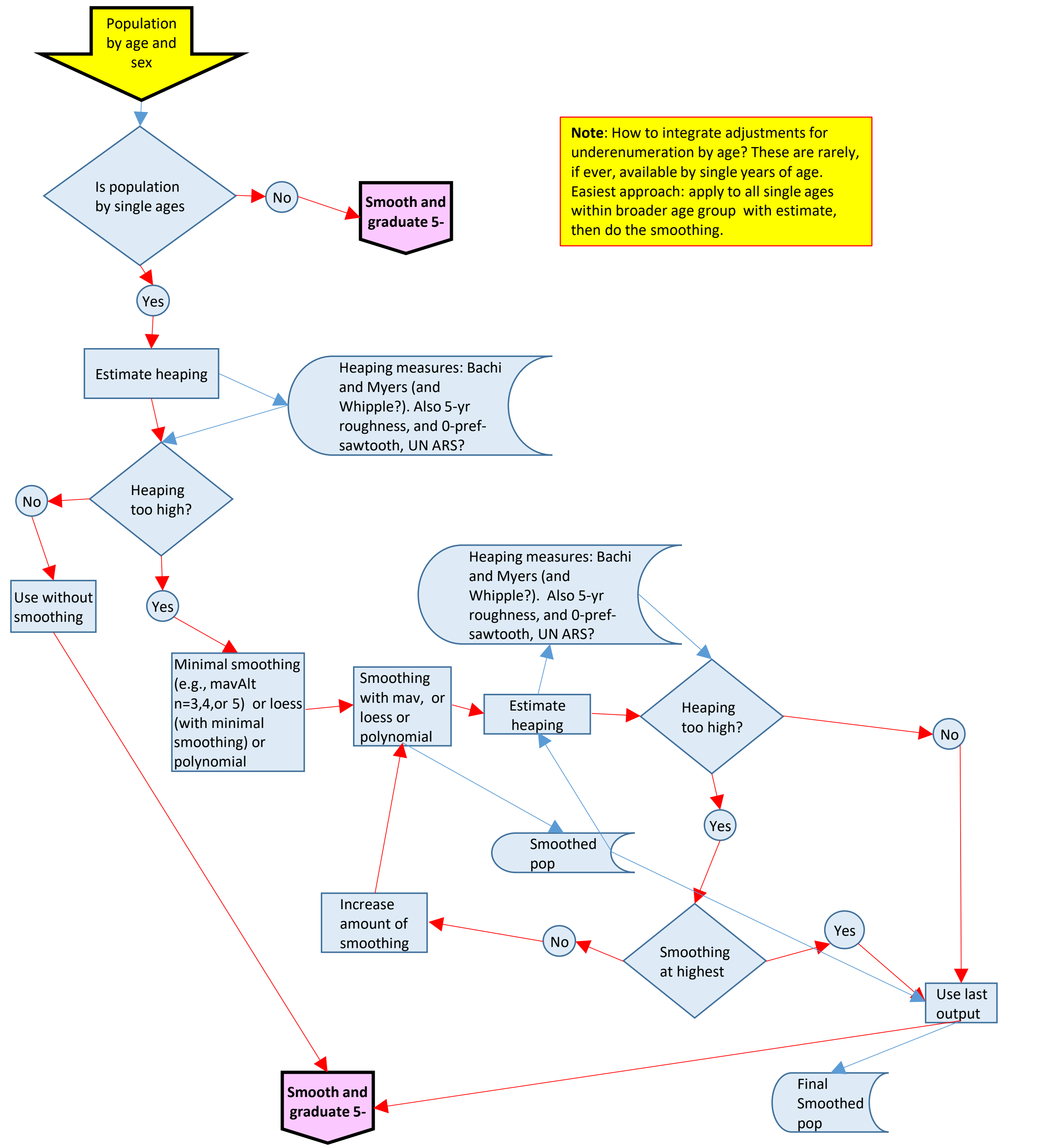


Diagram PopA-2: Analysis and Adjustment of Populations by Sex and 5-year Age Groups

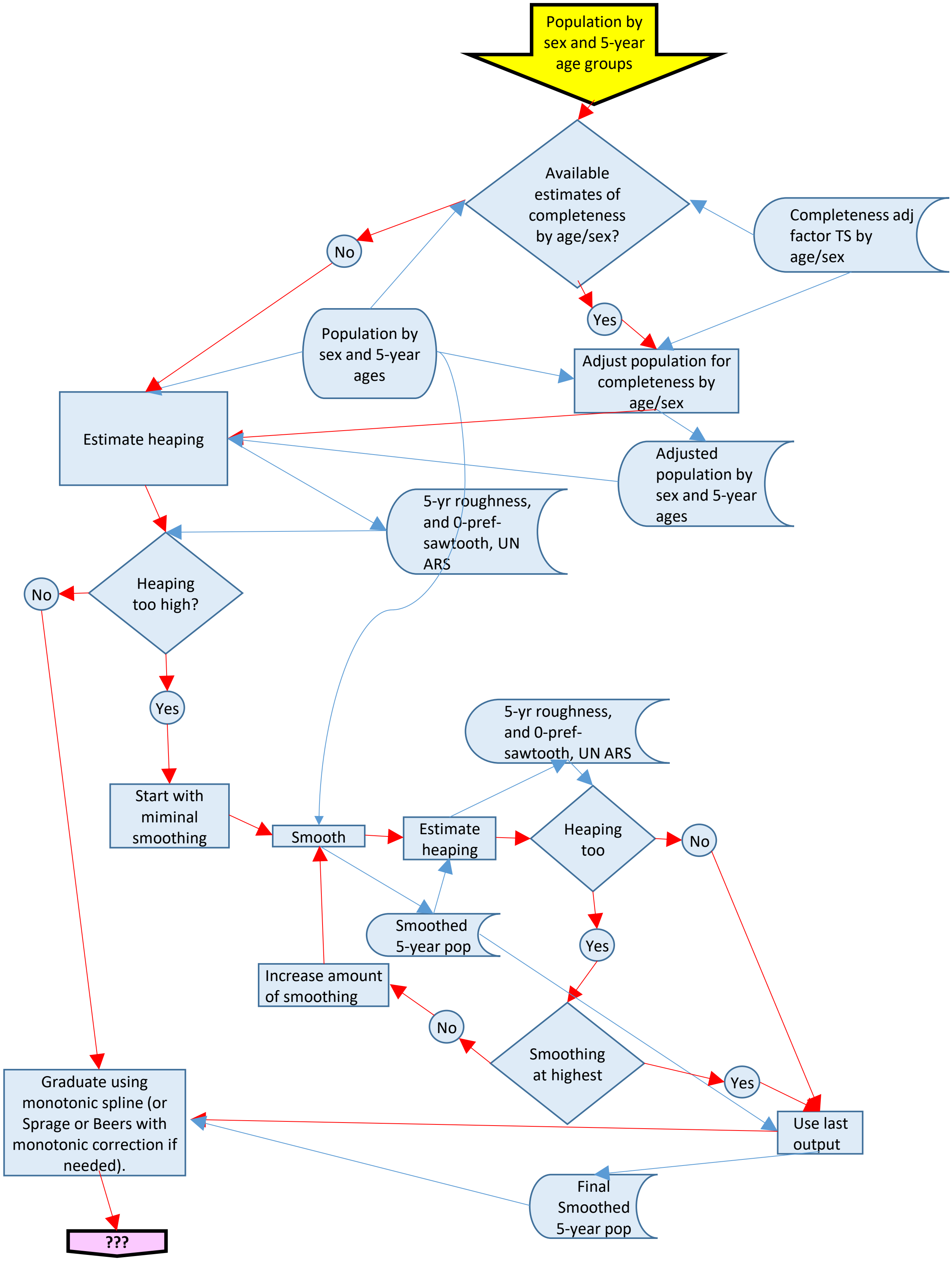


Diagram MortA-1: Adult Mortality Overview

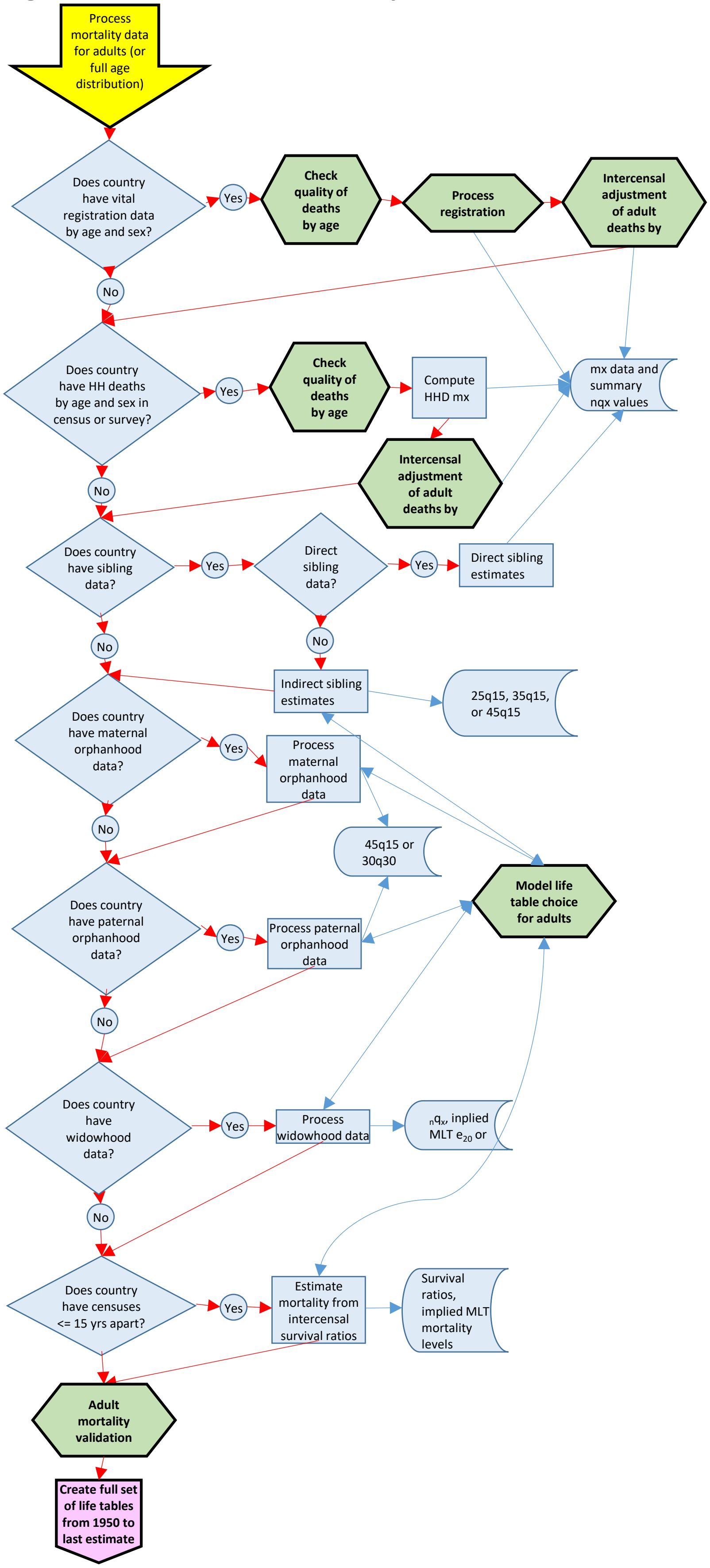


Diagram MortA-2: Adult Mortality Decision Table

Notes	Reg totals	Reg by age	HH Dth	Sibling	Mat orph	Pat orph	Widowhood	Intercensal survival	Action/output
Common problem with small countries	Yes	No	No	No	No	No	No	No	MLT to hit total? Any estimates from cty of e0, other measures?
See sep tree		Yes							mx and life tables
			Yes						mx and life tables (unadj) for census year or intercensal period
			Yes with adj						mx and life tables for intercensal period
				Direct					mx for some adult ages
Needs MLT				Indirect					Various indicators
Needs MLT					Yes				Various indicators
Needs MLT						Yes			Various indicators
Needs MLT							Yes		Various indicators
Needs MLT								Yes	Various indicators
Virtually no adult mortality data for a time period	No	No	No	No	No	No	No	No	1. Use 5q0 if available 2. Use neighboring country 3. Use regional data 4. Extrapolate using data for this country

Diagram MortA-3: Registered Deaths by Age and Sex

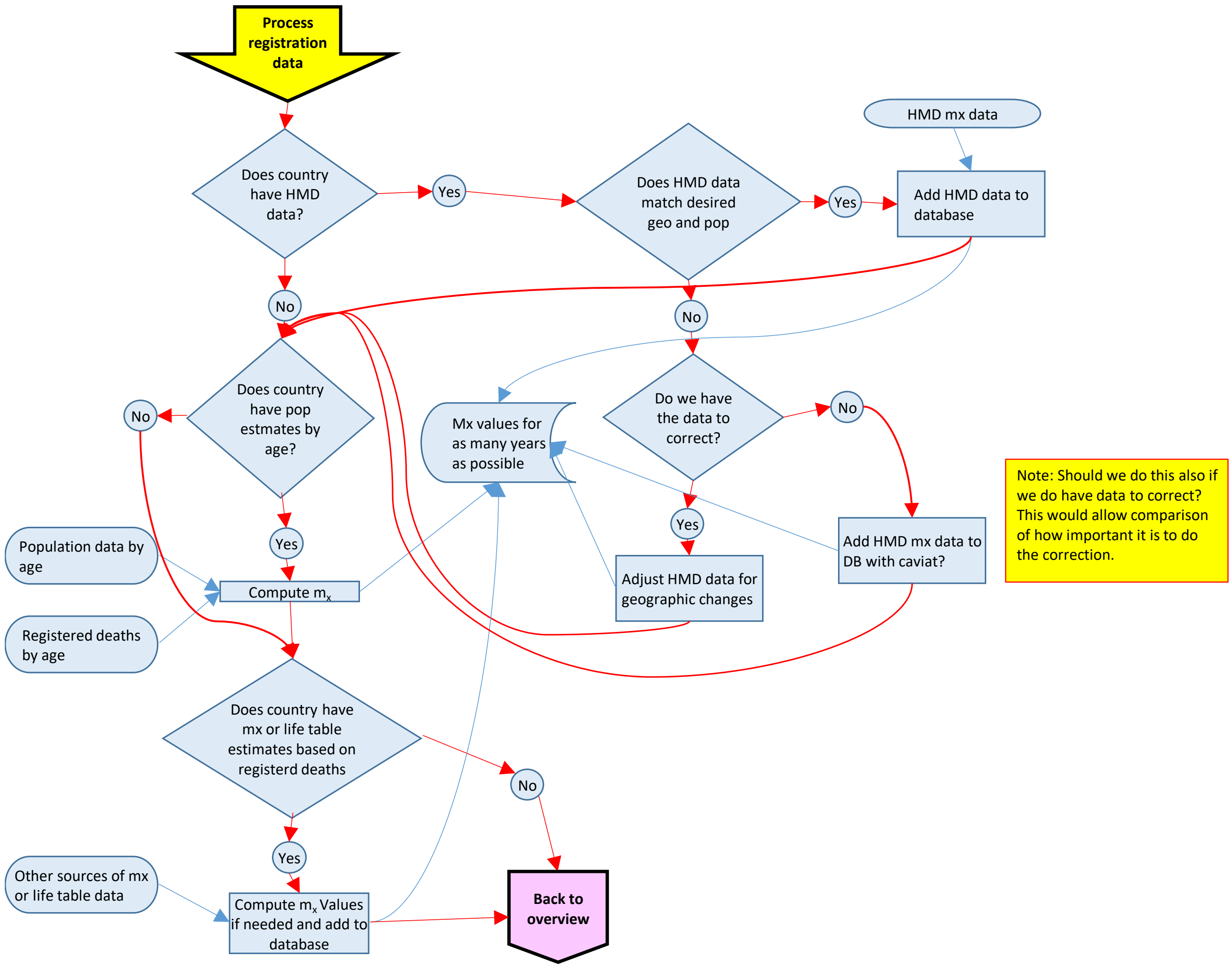
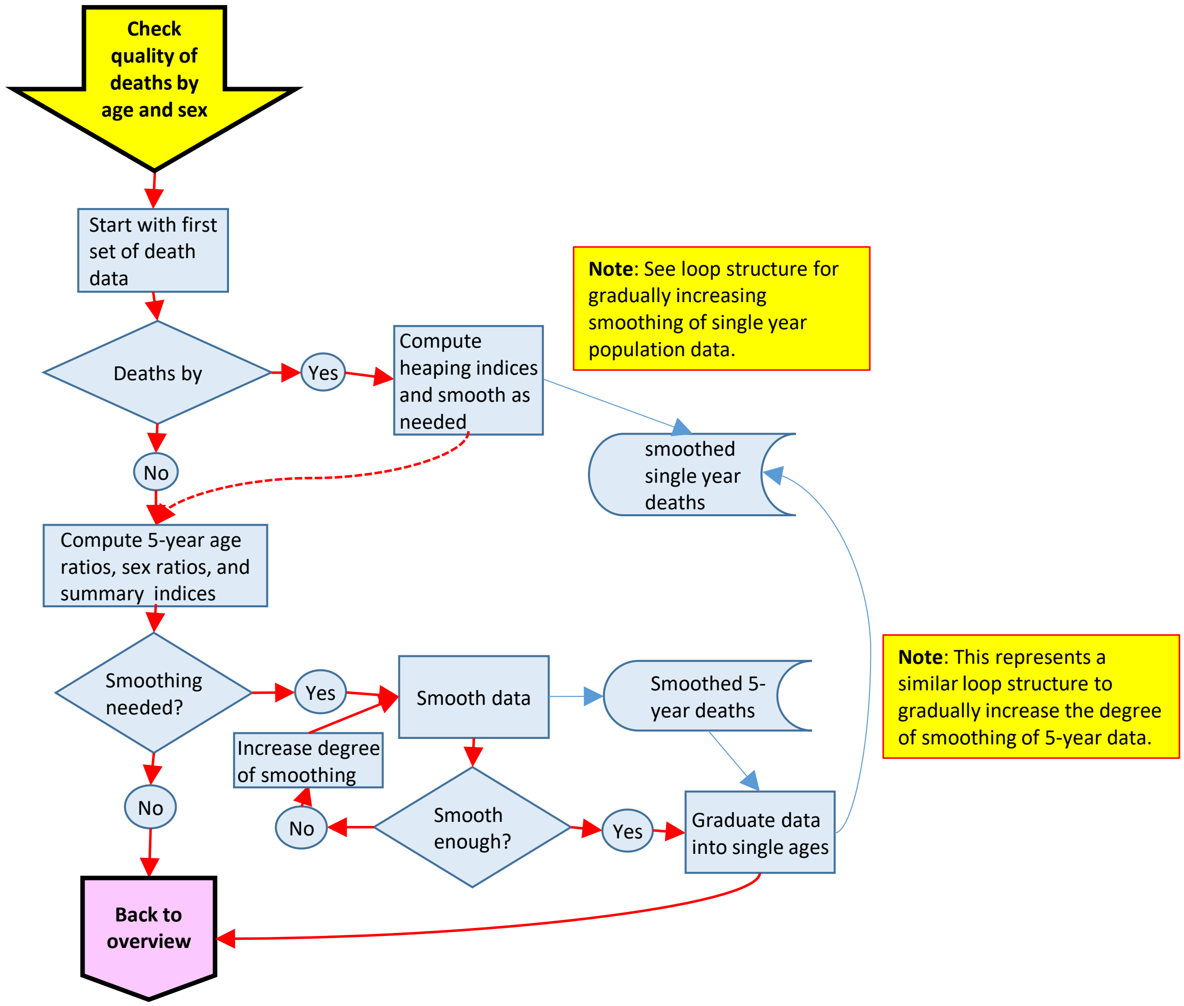


Diagram MortA-4: Adult Death Registration Decision Table

Notes	HMD	HMD geo ok	Fix HMD geo	Pop ests	Mx ests	Action
This is always true	Yes	Yes				Add HMD data to database as far back as possible
	Yes	No	Yes			Add corrected HMD to database
	Yes	No	No			Add uncorrected data to database for comparison?
HMD status irrelevant				Yes		Compute Mx and add to DB
HMD status irrelevant				No		Use previous WPP pop to get Mx and add to DB
HMD status irrelevant				No	Yes	Add other (official?) Mx to DB

Diagram MortA-5: Check Quality of Adult Deaths by Age and Sex



Questions/Notes:

1. What to do with high sex ratio indices of deaths by age?
   a. May indicate differential mortality patterns by sex (like changes in Colombia)
   b. May indicate differential levels of completeness of death registration
   c. Maybe don't bother looking at sex ratios (except to make sure there aren't any sudden changes that may point to data errors or special circumstances). Sex ratios of mx values are usually more informative.
2. Check on cohort patterns?
   a. May indicate abnormal birth cohort size (baby boom or bust)
   b. Possible long-term impacts of events in early childhood?

Diagram MortA-6: Intercensal Adjustment of Deaths by Age and Sex

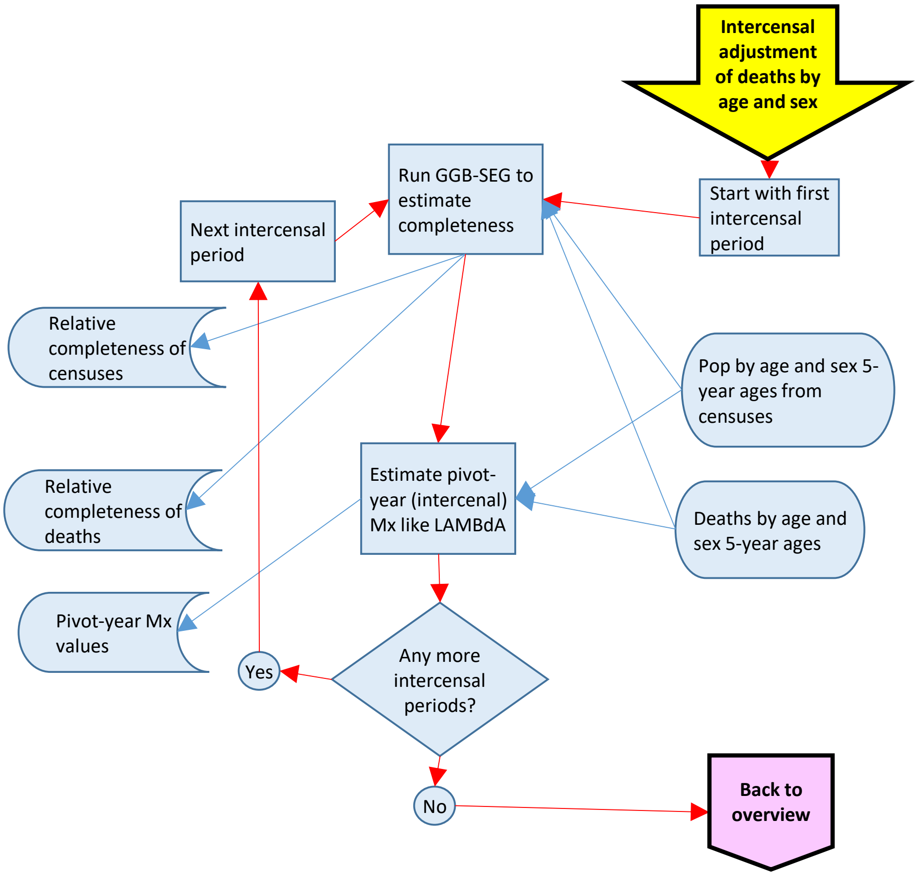


Diagram MortA-7: Adult Mortality Validation and Review

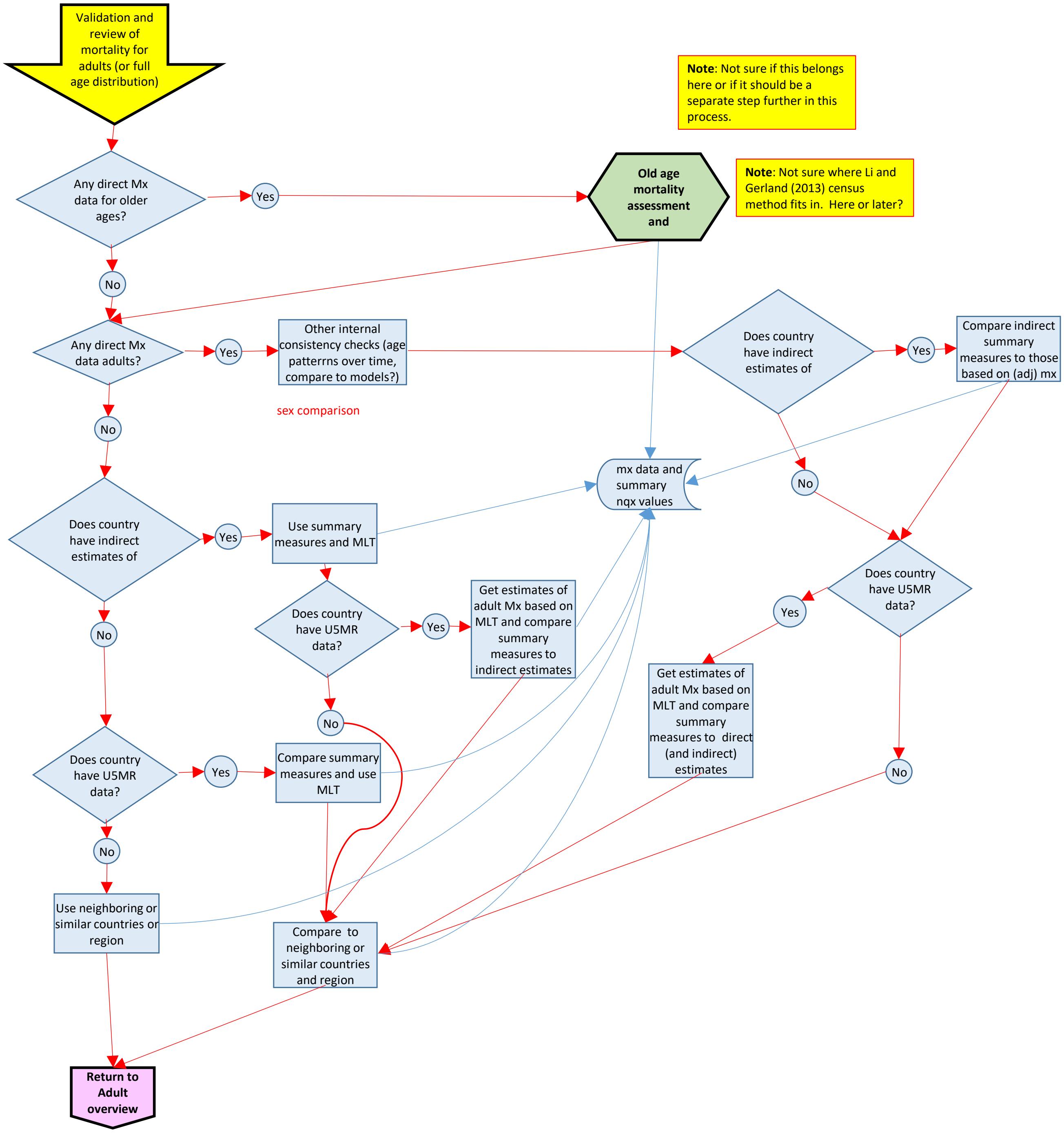


Diagram MortA-8: Old Age Assessment and Adjustment

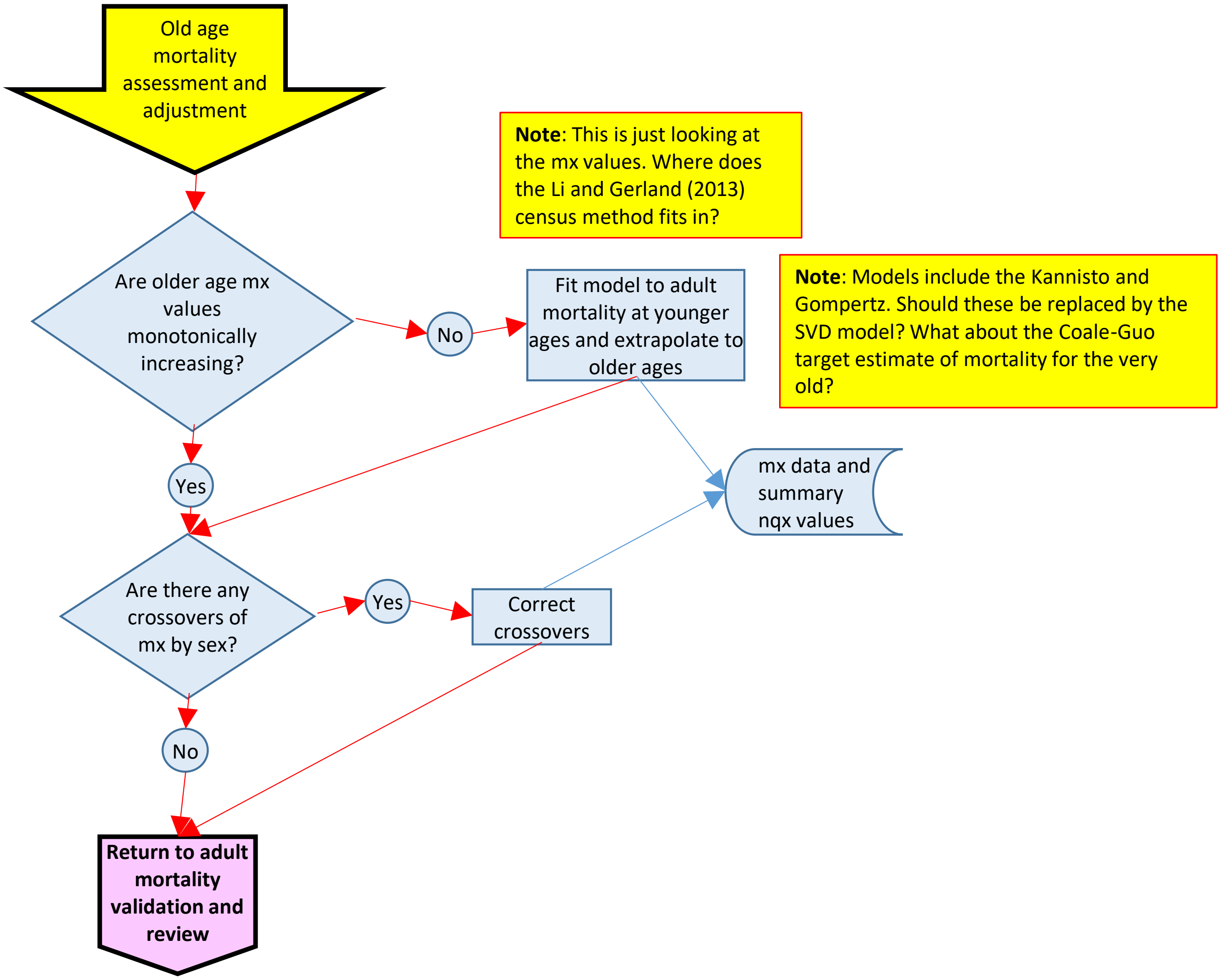


Diagram MortC-1: Child Mortality Overview

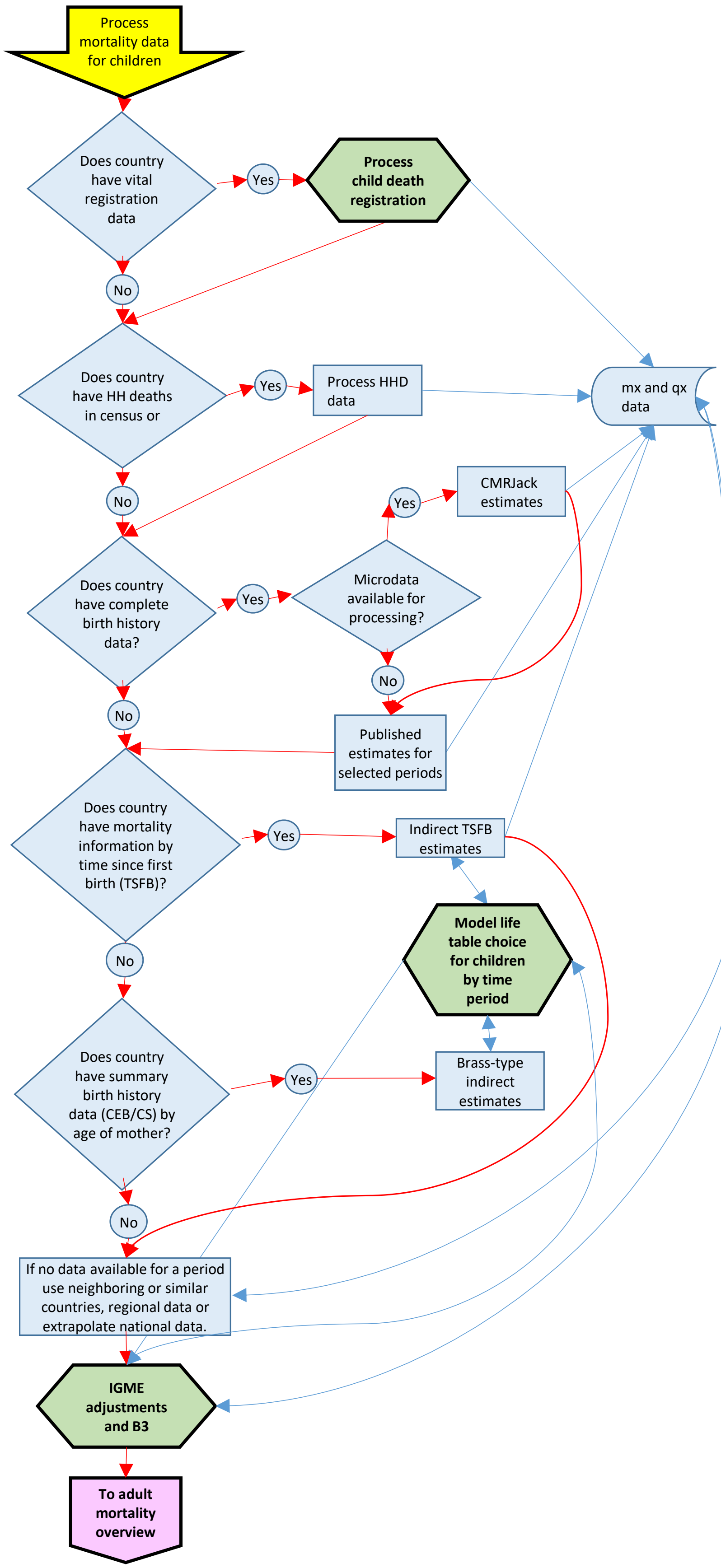


Diagram MortC-2: Child Mortality Registration Data

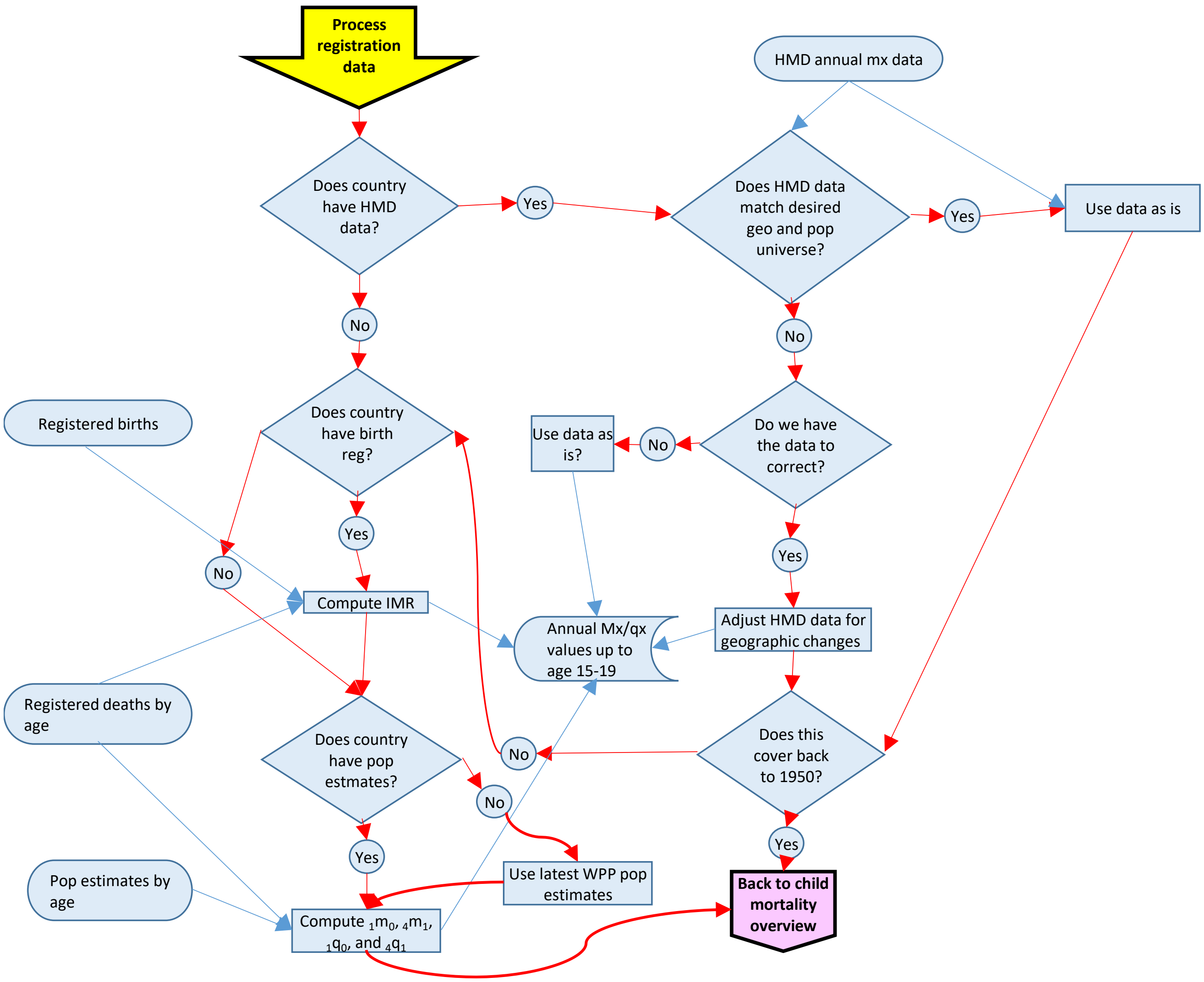


Diagram MortC-3: IGME Adjustments and B3 Model

https://childmortality.org/methods

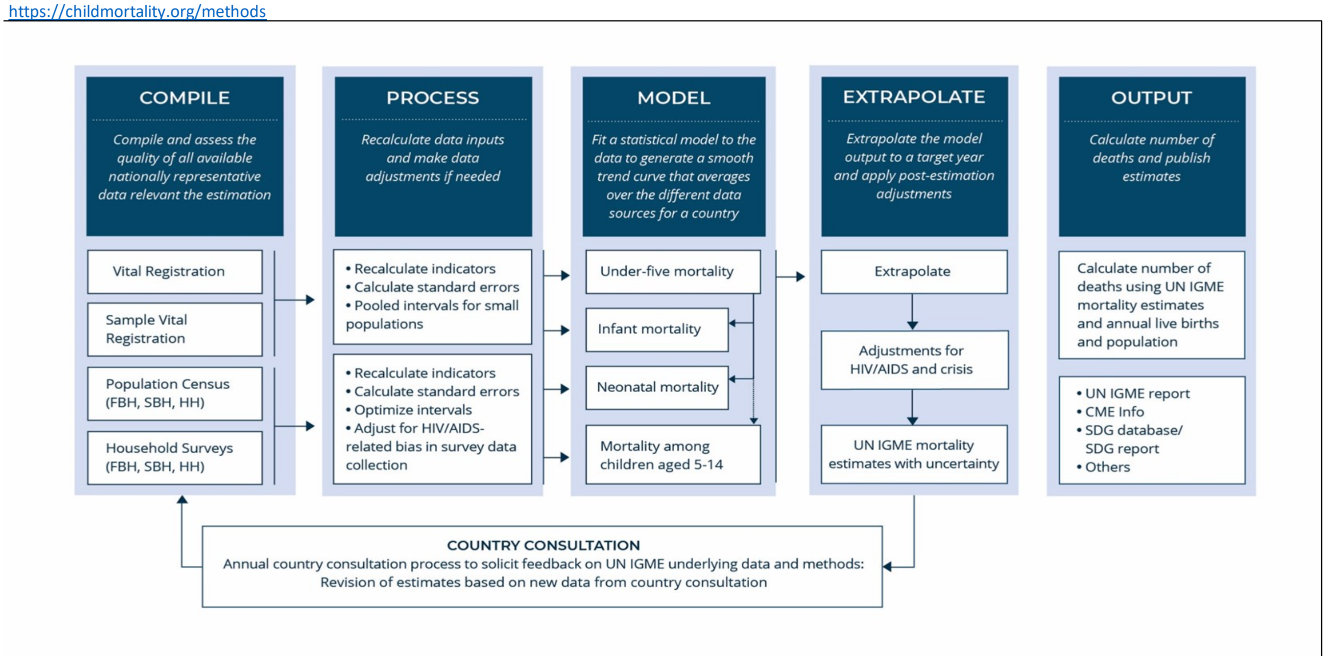


Diagram MortC-4: IGME Notes

Overview of IGME procedures (p. 26) [my emphasis in red]
https://childmortality.org/wp-content/uploads/2019/10/UN-IGME-Child-Mortality-Report-2019.pdf

UN IGME follows the following broad strategy to arrive at annual estimates of child mortality:

1. Compile and assess the quality of all available nationally representative data relevant to the estimation of child mortality, including data from vital registration systems, population censuses, household surveys and sample registration systems

2. Assess data quality, recalculate data inputs and make adjustments as needed by applying standard methods

3. Fit a statistical model to these data to generate a smooth trend curve that averages possibly disparate estimates from the different data sources for a country

4. Extrapolate the model to a target year (in this case, 2018)

p. 28
Whatever the method used to derive the estimates, data quality is critical. UN IGME assesses data quality and does not include data sources with substantial non-sampling errors or omissions as underlying empirical data in its statistical model.

5. Criteria for inclusion/exclusion in the B3 model
    a. Include direct FBH data if available.
    b. Exlude estimates from SBH if estimates from FBH in the same survey are available
    c. Exclude certain data sources (e.g., surveys/censuses) that don't seem consistent based on data quality assessment (i.e., substantial non-sampling errors, omissions or lack of national representation).
    d. Exclude VR data if implied completeness < X%?
    e. Exclude data points from crisis years, fit the B3 model to the remaining data, and then add the crisis-specific mortality rate to the fitted B3 curve

Diagram Fert-1: Compute Fertility Estimates Based on Censuses and Surveys

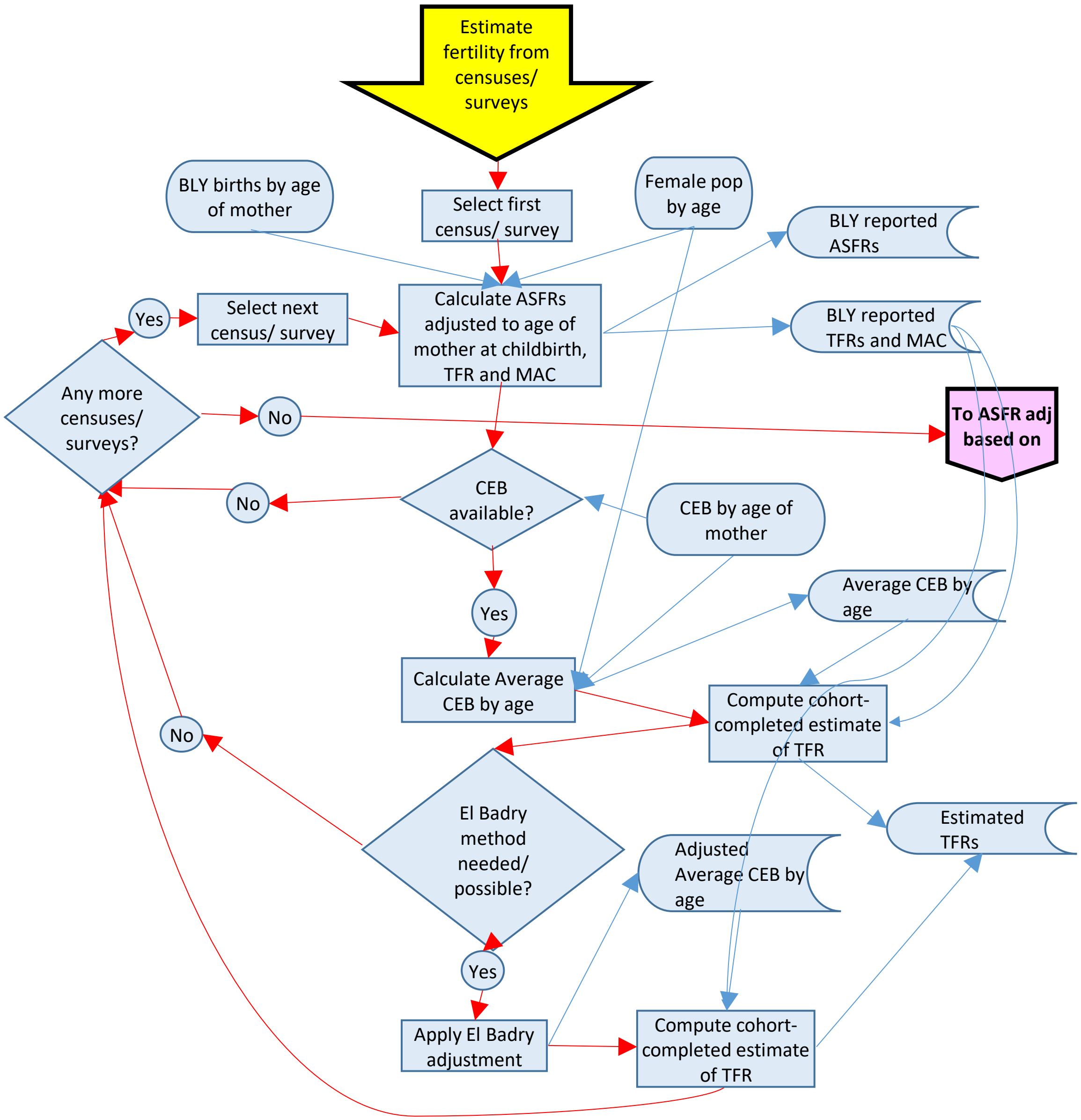


Diagram Fert-2: Evaluate and Adjust Fertility Estimates Based on Censuses and Surveys

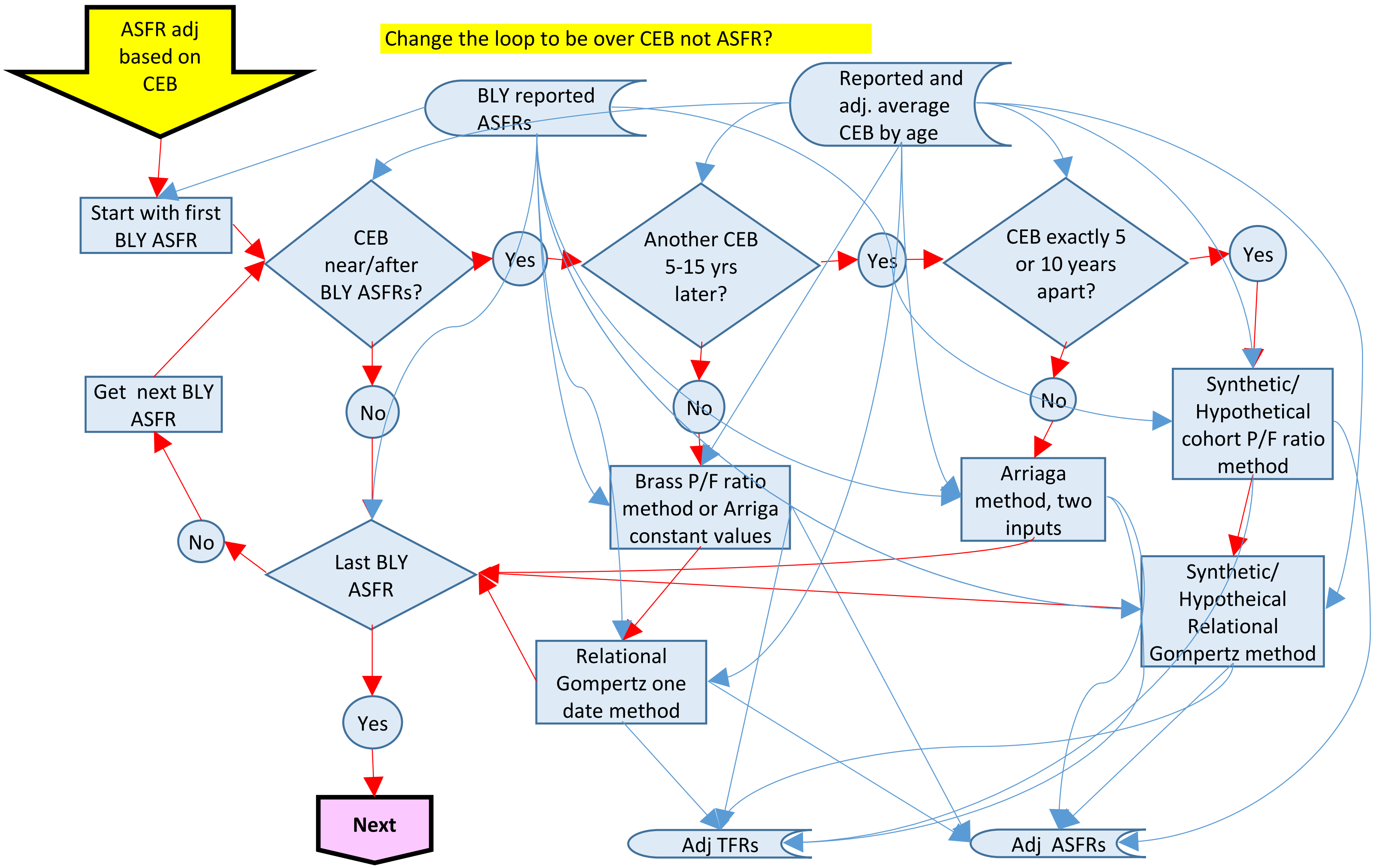


Diagram F-3: Compute Fertility Estimates Based on Registration Data

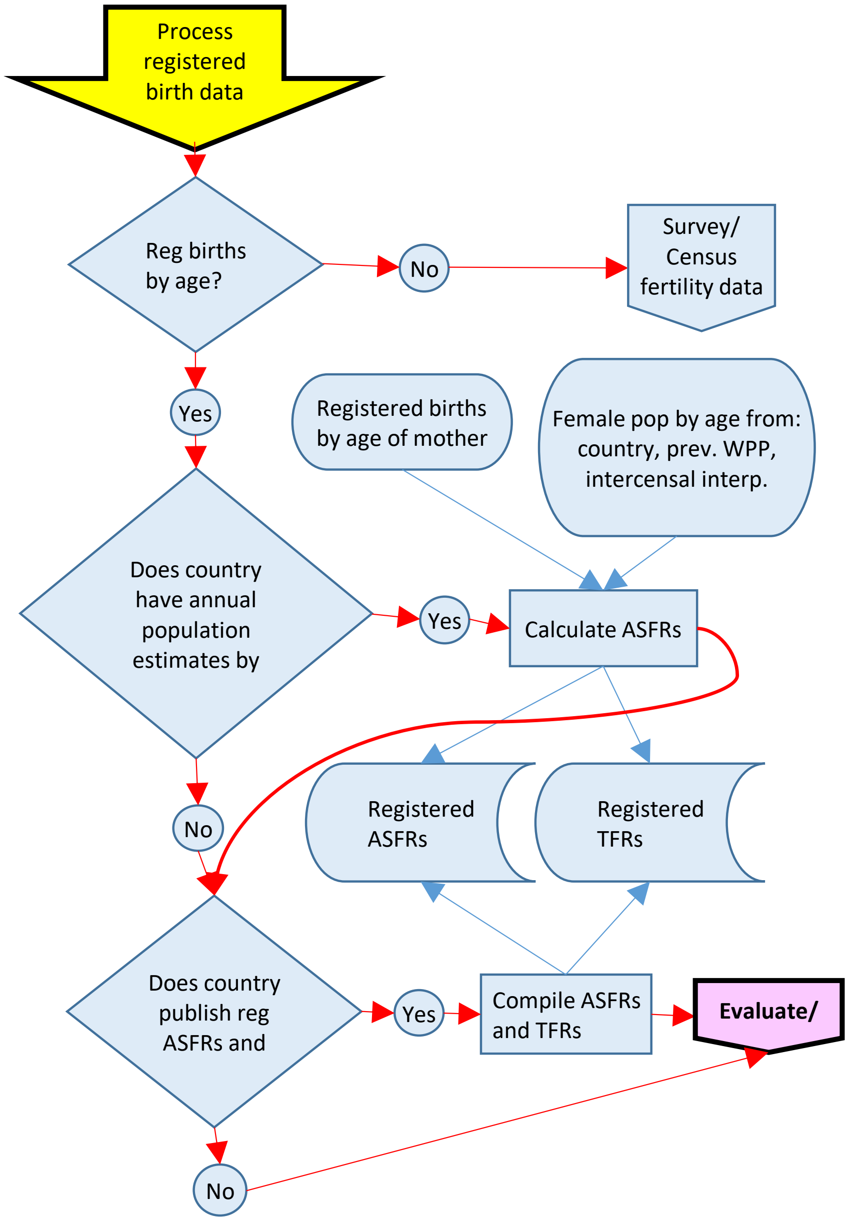


Diagram Fert-4: Evaluate and Adjust Registered Fertility Compared to Data From Censuses and Surveys