Methodological and practical issues in child mortality estimation

Richard J Silverwood & Simon Cousens

January 2009

Department of Epidemiology and Population Health, London School of Hygiene and Tropical Medicine, Keppel Street, London, WC1E 7HT

E-mail: Richard.Silverwood@lshtm.ac.uk

Introduction

The Millennium Development Goal 4 (MDG4), to reduce under-5 mortality by two-thirds between 1990 and 2015 [1], has caused child mortality to be the focus of much attention in recent years. However, in order to assess progress towards MDG4 it is essential to be able to generate estimates of current levels and trends in under-5 mortality.

Recently, Murray et al [2] proposed a new approach to the estimation of under-5 mortality, since used by others [3], which aims to be reproducible and transparent. This approach uses data derived from different sources (e.g. survey data, vital registration data) at different time points, and applies loess regression to obtain smoothed estimates of past levels of child mortality and to predict future levels based on past and recent trends.

The method proposed by Murray et al is described in more detail elsewhere [2], but is briefly reviewed here. The general approach is to use loess regression [4, 5], a type of local regression, to model under-5 mortality within a given country. The basic loess model is

$$\log(y_i) = \beta_0 + \beta_1 x_i + \beta_2 z_i + \varepsilon_i, \tag{1}$$

where y_i is under-5 mortality, x_i is year, z_i is an indicator variable taking value 1 if the observed value comes from a vital registration system and value 0 otherwise, and $\varepsilon_i \sim N(0, \sigma_\varepsilon)$ is an error term.

The loess function is fitted using weighted least squares regression, with the weights corresponding to each observed under-5 mortality value calculated using a separate weighting function. This weighting function is tuned by a single parameter, α. By fitting the loess model using a variety of values of α, it is possible to vary the sensitivity of predictions to recent data trends.

For each time point of interest the loess model is fitted for each of a pre-determined set of α values, and the multivariate normal distribution defined by the estimated regression coefficients and their variance-covariance matrix is used to simulate 1000 random draws. For each of these random draws under-5 mortality is estimated assuming non-vital registration data (i.e. setting $z = 0$ in (1)). The 1000 estimates per α value are then pooled across the set of α values, and the final estimated under-5 mortality calculated as the median value, with an uncertainty interval corresponding to the 2.5^{th} and 97.5^{th} centiles.

This paper is split into several essentially independent sections in which different issues relating to the loess-based approach to under-5 mortality estimation are discussed. In the first section we investigate a perceived problem with Murray et al's implementation of the approach, which leads to their published uncertainty intervals being unduly narrow. In many countries, both direct and indirect estimates derived from the same survey are currently used when estimating under-5 mortality, which is a questionable approach given the clear lack of independence between the estimates. Therefore, in the second section we investigate the effect of using only the direct estimates in the these situations. A further issue relating to estimates derived from survey data is that the loess-based approach as formulated by Murray et al assumes independence of data points for different time periods from the same survey, which is unlikely to be true in most situations. One way to relax the assumption of independence is through the use of multilevel modelling. In the third section we investigate the effect of using a random intercepts loess-based model. In the final section we briefly discuss some further outstanding issues with the loess-based approach to estimating under-5 mortality.

The data used by Murray et al are available from the Institute for Health Metrics and Evaluation (IHME) website [6], and we use these data in our analyses. Murray et al initially examined data from 189 countries, though they only provide under-5 mortality estimates for 172 of these due to the exclusion of countries with very small populations. We restrict our attention to these 172 countries. Within each country we exclude the same data points as were excluded

by Murray et al, though for many of these it is unclear what the precise basis for exclusion was.

Implementation by of the loess-based approach Murray et al

On publication of the under-5 mortality estimates produced using the loess-based approach of Murray et al [2], Wardlaw [7] identified several examples of countries where the estimated uncertainty bounds appeared surprisingly narrow, namely Congo, Belarus, Côte d'Ivoire, and Ukraine.

Method

For these four countries, we reanalyse the mortality data used by Murray et al and re-estimate under-5 mortality using the method described in their paper. We then compare these re-estimated values to the estimates and projections of Murray et al (obtained from the IHME website [8]).

Results

In Fig. 1–Fig. 4, the under-5 mortality estimates of Murray et al and our re-estimated values, with corresponding uncertainty intervals, are plotted for these four countries. The values plotted in these figures are also presented for selected years in Table 1. In all four countries Murray et al's point estimates and our re-estimates of under-5 mortality are seen to be virtually identical at each time point (note that due to the random draw simulation element of the approach, the two sets of estimates would not necessarily be expected to be identical). For the Congo and Côte d'Ivoire, however, the uncertainty intervals corresponding to our re-estimated values are far wider than those reported by Murray et al, particularly for years within the range of observed data.

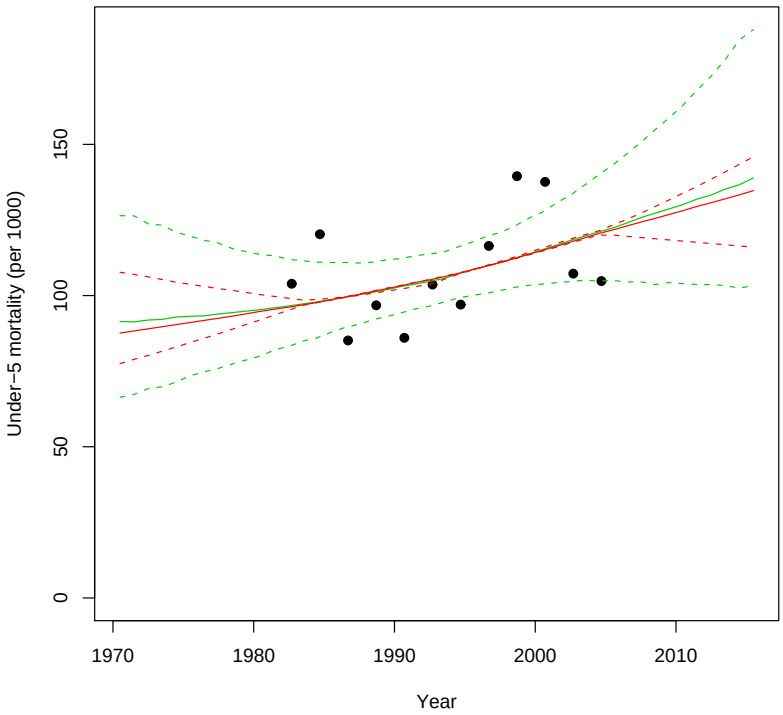

Fig. 1: Comparison of estimates for the Congo. Black points represent observed non-vital registration data; solid red line represents estimates of Murray et al; solid green line represents our re-estimated values; dashed lines represent uncertainty intervals.

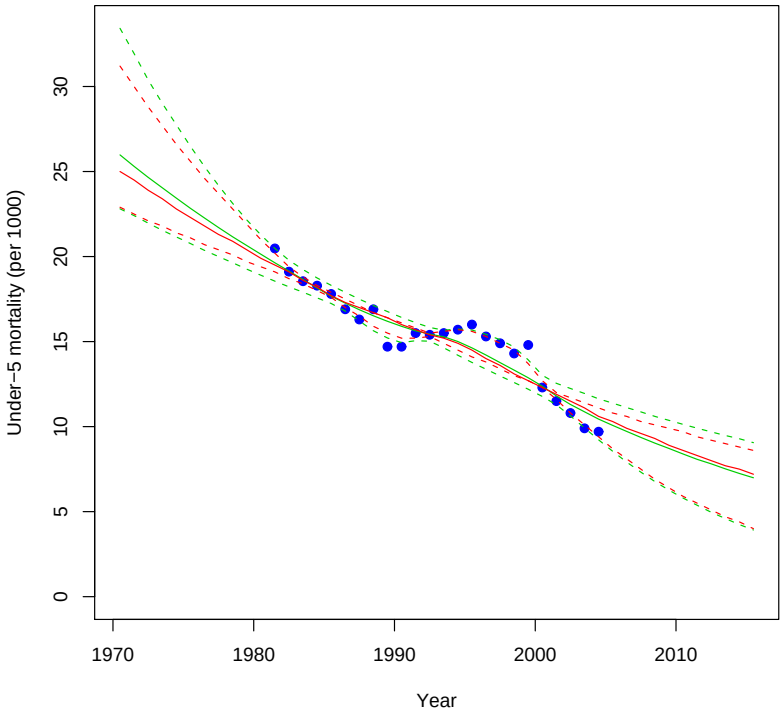

Fig. 2: Comparison of estimates for Belarus. Blue points represent observed vital registration data; solid red line represents estimates of Murray et al; solid green line represents our re-estimated values; dashed lines represent uncertainty intervals.

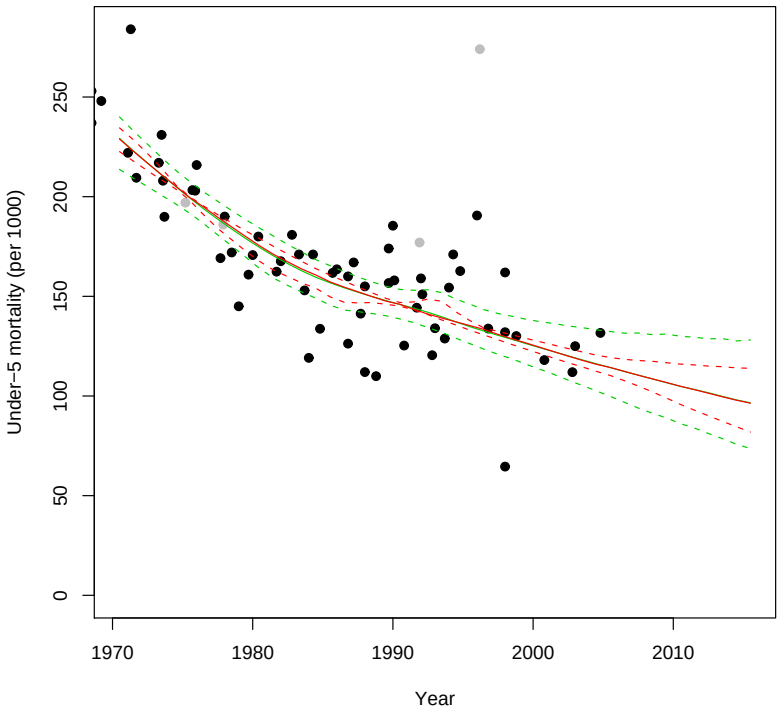

Fig. 3: Comparison of estimates for Côte d'Ivoire. Black points represent observed non-vital registration data; grey points represent excluded data; solid red line represents estimates of Murray et al; solid green line represents our re-estimated values; dashed lines represent uncertainty intervals.

		Under-5 mortality (per 1000)			
		Murray et al		Our re-estimate	
Country	Year	Estimate	UI	Estimate	UI
Congo	1970	88	78–108	91	67–129
	1980	95	92–100	95	80–114
	1990	103	102–103	103	94–112
	2000	115	115–116	115	104–128
	2010	128	118–134	130	104–163
	2015	135	116–146	138	103–187
Belarus	1970	25	23– 31	26	23– 33
	1980	20	19– 21	20	19– 21
	1990	16	15– 16	16	15– 16
	2000	12	12– 13	12	12– 13
	2010	9	6– 10	8	6– 10
	2015	7	4– 9	7	4– 9
Côte d'Ivoire	1970	229	223–234	229	214–240
	1980	175	168–179	174	164–184
	1990	146	145–147	146	139–153
	2000	124	121–127	124	113–137
	2010	105	96–116	105	86–130
	2016	96	82–114	96	73–128
Ukraine	1970	26	24– 30	26	24– 30
	1980	22	21– 23	22	21– 23
	1990	18	18– 18	18	17– 19
	2000	15	15– 15	15	15– 16
	2010	12	10– 12	12	10– 13
	2016	10	8– 12	10	8– 12

Table 1: Comparison of Murray et al's reported estimates [2] with our re-estimated values. UI is uncertainty interval.

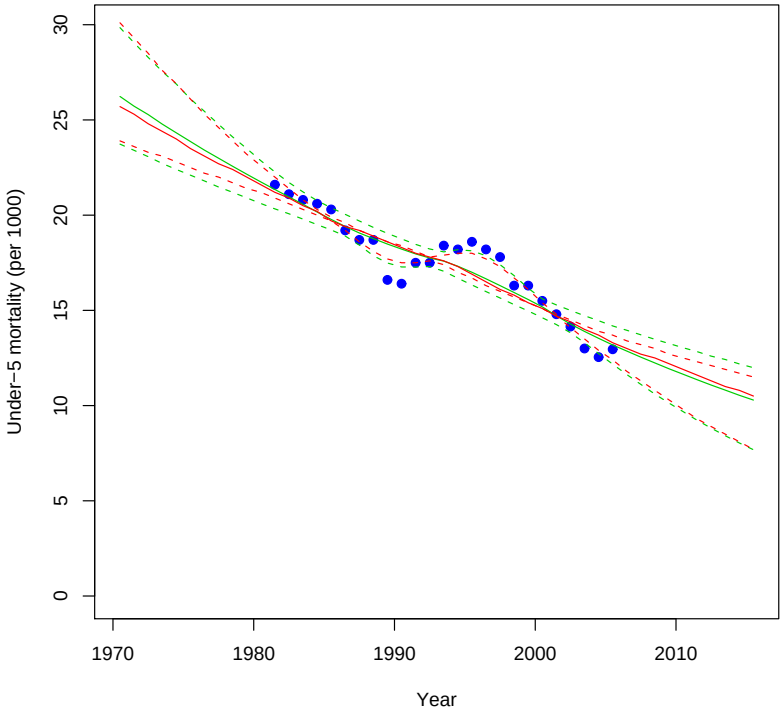

Fig. 4: Comparison of estimates for Ukraine. Blue points represent observed vital registration data; solid red line represents estimates of Murray et al; solid green line represents our re-estimated values; dashed lines represent uncertainty intervals.

Discussion

In the loess-based approach described by Murray et al [2], the uncertainty surrounding a given estimate is calculated by combining within-α value parameter uncertainty (reflecting sampling and other variation in the data points) with between-α value model uncertainty. From examination of the computer code used by Murray et al to calculate the reported under-5 mortality estimates, which they kindly shared with the Inter-Agency Coordination Group on Child Mortality Estimation, it appears that there has been insufficient allowance made for within-α value parameter uncertainty for years within the range of observed data.

This observation is borne out by the differences between the estimates seen in Fig. 1–Fig. 4 and Table 1. In Belarus and Ukraine the observed data follow a clear trend with little deviation from that trend due to sampling variation. This means that for each α value the fitted loess regression models fit the observed data well, and hence there is little within-α value parameter uncertainty. Thus, when computing overall uncertainty, similar intervals will be obtained regardless of whether or not the within-α value parameter uncertainty is fully taken into account. Additionally in Belarus and Ukraine, the trends in the observed data include time periods in which the trend is close to linear on the log-transformed under-5 mortality scale. In the middle of these time periods each α value will result in a similar fitted model, meaning there is little between-α value model uncertainty. This results in the relatively narrow uncertainty intervals seen at certain time points both in Murray et al's reported estimates and our re-estimated values. Furthermore, as both Belarus and Ukraine are estimated to have relatively high levels of vital registration coverage (most recent estimates of 97.6% and 96.2% respectively [9]), the generally narrow uncertainty intervals for these two countries do not seem implausible.

In the Congo and Côte d'Ivoire the observed data do not follow such an obvious trend, and there is much greater variation between data points close together in time. Thus for each α value there will be relatively large within-α value parameter uncertainty. Whether or not the within-α value parameter uncertainty is fully taken into account

will therefore make a large difference to the final estimated uncertainty intervals. This situation is exacerbated in the Congo where the scarcity of data means that only α values ≥ 1 are used. This leads to each α value resulting in similar fitted loess regression models (and hence virtually no between-α value model uncertainty), meaning that ignoring within-α value parameter uncertainty results in particularly narrow uncertainty intervals.

To conclude, we believe that Murray et al's implementation of their proposed method did not in practice properly account for sampling uncertainty within the range of observed data. Whilst this will lead to underestimation of the overall uncertainty in all countries, in countries with only a small amount of noisy data the underestimation may be especially marked.

The use of direct and indirect data

In many countries, both direct and indirect estimates derived from the same survey are currently used when estimating under-5 mortality. For example, Armenia has 12 direct estimates of under-5 mortality derived from the 2000 Demographic and Health Survey (each estimate covering a different time period) as well as 5 indirect estimates (which cover some of the same time period covered by the direct estimates). The same information thus being used twice, each time carrying the same weight as data from a source which only provides either direct *or* indirect estimates, is a questionable approach.

Method

We examine the impact on the estimates and uncertainty intervals obtained of using only the direct estimates (making the assumption that the birth histories used to calculate the direct estimates are valid).

We re-examine the data for the 172 countries for which Murray et al provided under-5 mortality estimates [2], identifying the countries in which this problem occurs. For these countries we reanalyse the data after first excluding the relevant indirect estimates and compare the estimated under-5 mortality (and associated uncertainty) to that when including all data. As the number of data points is reduced by the exclusion of the indirect estimates, the range of α values which can be used may also be reduced. We thus recalculate the set of α values to use in the manner detailed by Murray et al [2].

Results

Of the 172 countries analysed by Murray et al [2], 66 have data from surveys which provide both direct and indirect estimates. Table 2 summarises the number of data points for each of these countries. They have between 8 and 168 (median 59, IQR 31) data points in total, of which between 3 and 30 (median 10, IQR 8) are indirect estimates from sources which also provide direct estimates. These indirect data points make up between 5.8% and 62.5% (median 16.5%, IQR 11.3%) of the total data points for these countries. After the removal of these data points, each country has between 3 and 148 (median 46, IQR 33) remaining. The country with 3 remaining data points (Bhutan) has insufficient data for analysis and is excluded, so that 65 countries are reanalysed. For each country we estimate under-5 mortality, first using all available data, then again after having excluded indirect estimates from sources which also provide direct estimates.

In order to examine the effect of excluding indirect estimates from sources which also provide direct estimates across all 65 countries, the two sets of under-5 mortality estimates in 1970, 1990 and 2010 are plotted against one another in Fig. 5, Fig. 6 and Fig. 7, respectively. In each plot the two sets of estimates can be seen to be very similar in almost all countries, with Ethiopia in 2010 showing the only sizeable disparity.

	Number of observations				
	Min.	Q_1	Median	Q_3	Max.
Total	8	46	59	77	168
Indirect estimates from sources which also provide direct estimates	3	5	10	13	30
% of data which are indirect estimates from sources which also provide direct estimates	5.8	11.4	16.5	22.7	62.5
Total after removal of indirect estimates from sources which also provide direct estimates	3	36	46	69	148
% of data remaining after removal of indirect estimates from sources which also provide direct estimates	37.5	77.3	83.4	88.6	94.2

Table 2: Number of observations of under-5 mortality for the 66 countries with data from surveys which provide both direct and indirect estimates. Q_1 is the first quartile, Q_3 is the third quartile.

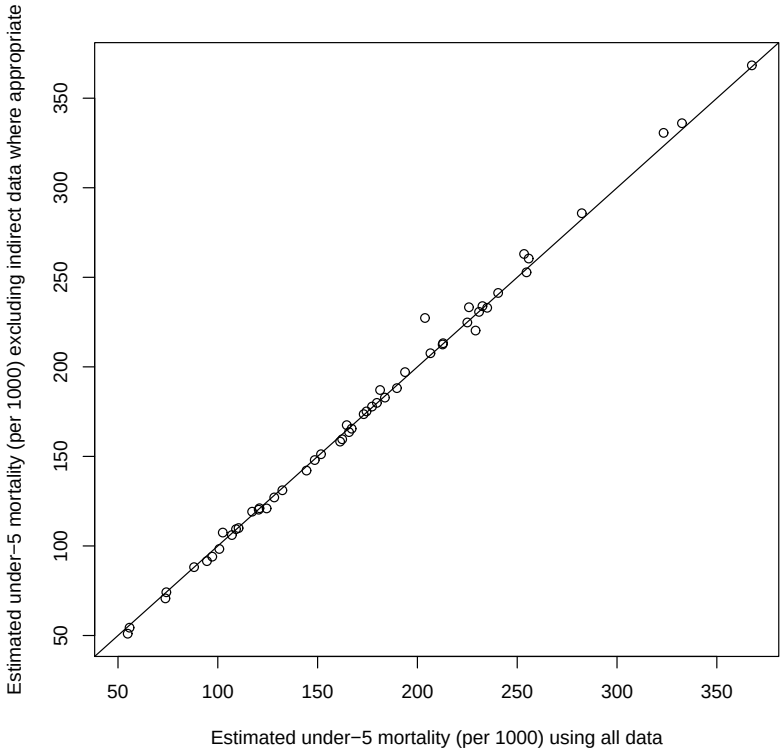

Fig. 5: Comparison of estimated under-5 mortality in 1970 when using all data and when excluding indirect estimates from sources which also provide direct estimates.

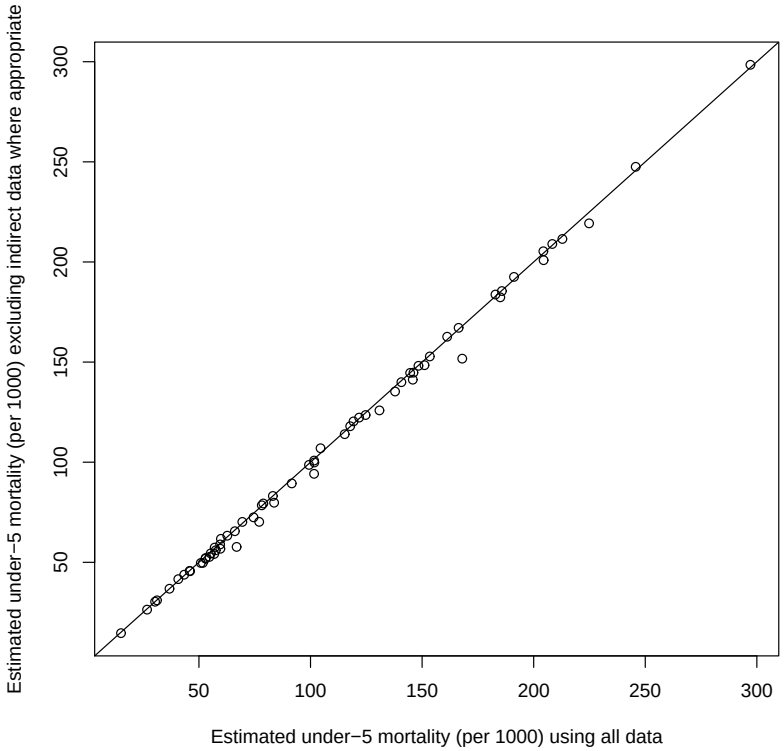

Fig. 6: Comparison of estimated under-5 mortality in 1990 when using all data and when excluding indirect estimates from sources which also provide direct estimates.

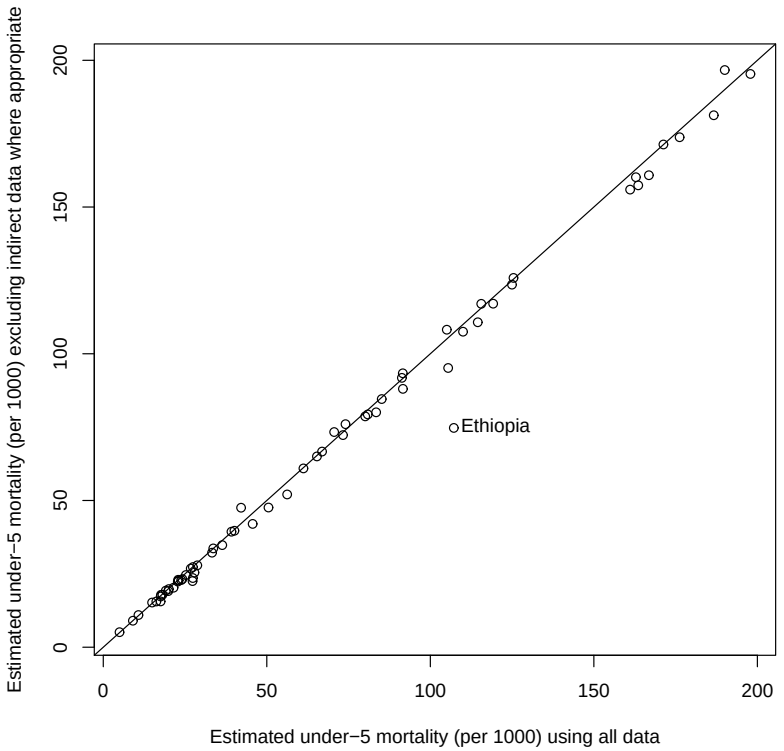

Fig. 7: Comparison of estimated under-5 mortality in 2010 when using all data and when excluding indirect estimates from sources which also provide direct estimates, with outlier identified.

When comparing the uncertainty intervals corresponding to the two sets of estimates it is preferable to consider the relative uncertainty interval widths (uncertainty interval width expressed as a percentage of the estimated under-5 mortality), rather than the absolute uncertainty interval width, due to potential differences in the underlying level of under-5 mortality. Fig. 8, Fig. 9 and Fig. 10 show relative uncertainty interval width when excluding indirect estimates from sources which also provide direct estimates plotted against relative uncertainty interval width when using all data for 1970, 1990 and 2000, respectively. In 1970 and 1990 the exclusion of indirect estimates results in increased relative uncertainty in the majority of countries, in some cases a considerable increase. By 2010, however, the relative uncertainty interval width differs little between the two sets of estimates in the majority of countries. The anomalously high relative uncertainty interval widths seen for Republic of Korea in 1990, whether indirect estimates are excluded or not, are due to the estimated under-5 mortality being low (15.1 per 1000 when using all the data) but there still being some degree of uncertainty (uncertainty interval 10.4–22.2 when using all the data).

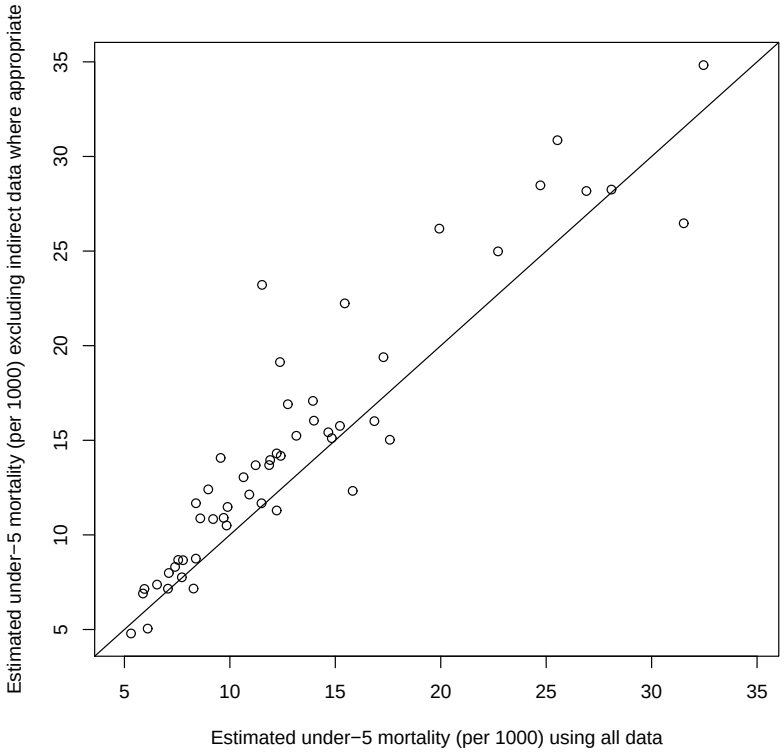

Fig. 8: Comparison of relative under-5 mortality uncertainty interval width (uncertainty interval width as a percentage of the corresponding estimate of under-5 mortality) in 1970 when using all data and when excluding indirect estimates from sources which also provide direct estimates.

Table 3 presents these results in the form of the percentage change in estimated under-5 mortality when going from the full data set to the restricted data set. Equivalent percentage changes in the relative uncertainty interval width are also given. The number of countries contributing at each time point differs as estimates are only calculated for time points at least as recent as the first observed data point for each country. The median effect at each time point from 1970 onwards is to decrease estimated under-5 mortality by a small amount, which increases at time progresses. For the middle 50% of countries estimates prior to 2010 change by a maximum of 2.5%. However, in some countries there are large increases or large decreases in estimated under-5 mortality. The larger effects are seen at time points beyond the range of observed data. As the model at this point continues any trends seen over the period of observed data, any small differences in the fitted models will be amplified in the predicted under-5 mortality values. On average the relative uncertainty interval width increases a little at each time point, though again there is considerable variability.

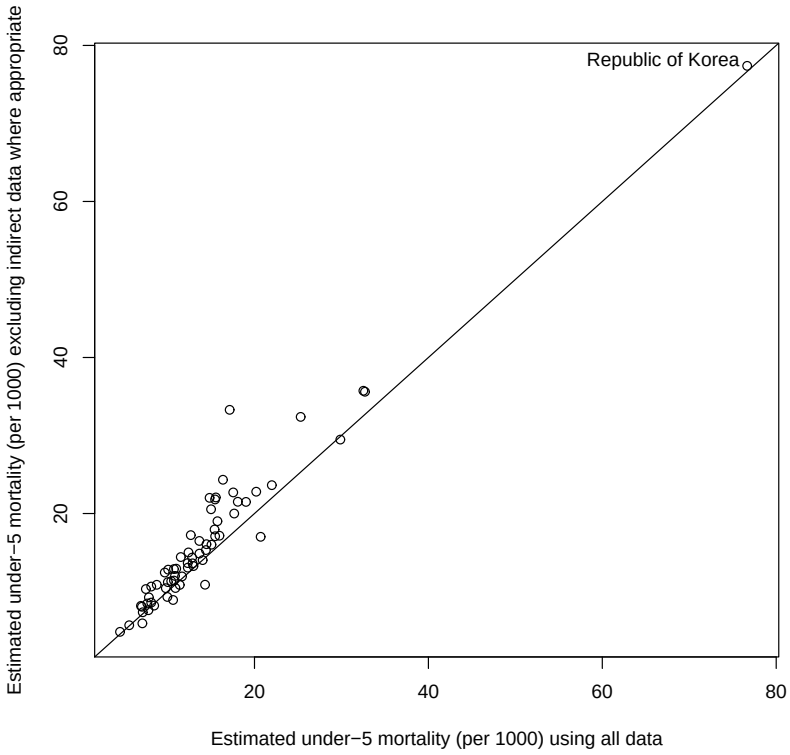

Fig. 9: Comparison of relative under-5 mortality uncertainty interval width (uncertainty interval width as a percentage of the corresponding estimate of under-5 mortality) in 1990 when using all data and when excluding indirect estimates from sources which also provide direct estimates.

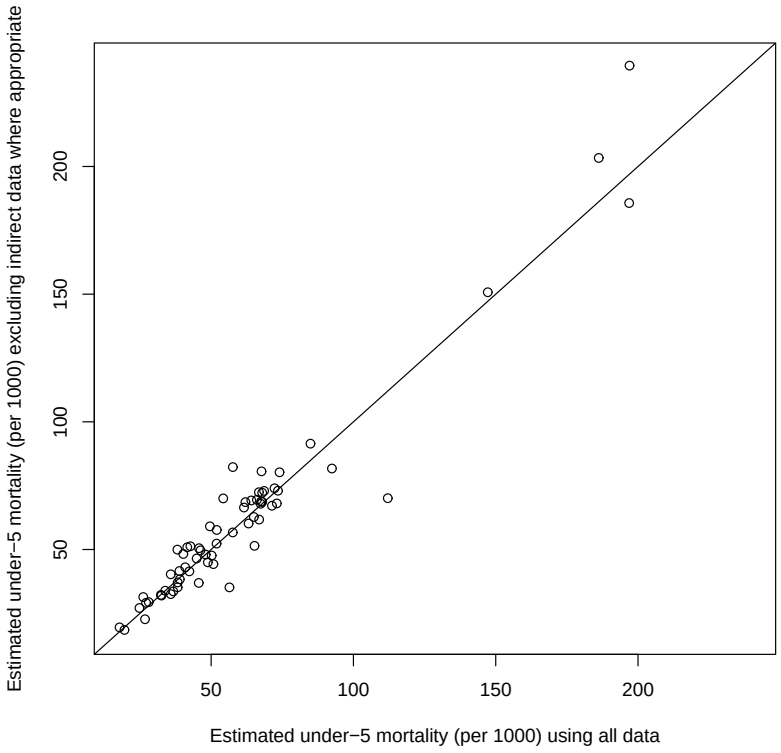

Fig. 10: Comparison of relative under-5 mortality uncertainty interval width (uncertainty interval width as a percentage of the corresponding estimate of under-5 mortality) in 2010 when using all data and when excluding indirect estimates from sources which also provide direct estimates.

For the middle 50% of countries relative uncertainty interval width changes by 20% or less.

		% change in estimated U5M			% change in relative U5M UI width		
Year	No. countries	Q_1	Median	Q_3	Q_1	Median	Q_3
1960	31	−0.8	+0.1	+0.6	−2	+5	+8
1970	52	−1.0	−0.1	+0.5	+3	+13	+20
1980	63	−1.7	−0.4	+0.3	+2	+8	+18
1990	65	−2.2	−0.8	+0.4	+6	+11	+21
2000	65	−2.5	−1.0	+0.0	−4	+3	+9
2010	65	−3.6	−1.5	+0.2	−5	+2	+9
2015	65	−4.1	−1.7	+0.2	−5	+2	+11

Table 3: Percentage (%) change in estimated under-5 mortality (U5M) and estimated U5M uncertainty interval (UI) width when going from using all data to excluding indirect estimates from sources which also provide direct estimates. Q_1 is the first quartile, Q_3 is the third quartile.

These results are best illustrated through examples where the estimated under-5 mortality changes considerably through adoption of this approach. However, it should be emphasised that in most countries the estimated under-5 mortality changes relatively little. Fig. 11 compares the fitted models for Ethiopia and Fig. 12 for Nigeria. In both plots the solid black points represent observed data from surveys providing only direct or indirect estimates, or direct estimates from surveys providing both direct and indirect estimates, and the hollow black points represent observed indirect estimates from surveys providing both direct and indirect estimates. The solid black line represents the fitted model when using all data (i.e. both solid and hollow points), the solid red line represents the fitted model when indirect data from surveys providing both direct and indirect data are excluded (i.e. solid points only), and the dashed lines represent the uncertainty intervals.

In Ethiopia (Fig. 11) the indirect estimates from surveys providing both direct and indirect estimates occur between 1975 and 1996 and are often greater than the other observations in that period, particularly from 1985 onwards. As a result, when these data points are excluded the estimated under-5 mortality is increasingly reduced as time progresses, leading to a 34% decrease (from 96 to 63 per 1000) in 2015. An additional consequence of the exclusion of indirect estimates from surveys providing both direct and indirect estimates is that the variability in observed under-5 mortality is reduced over much of the period considered, again particularly from 1985 onwards. This results in the uncertainty interval width associated with each estimate of under-5 mortality being greatly reduced across much of the plot, for example from a relative width of 144% to 85% (a 41% decrease) in 2015.

In Nigeria (Fig. 12) the indirect estimates from surveys providing both direct and indirect estimates occur between 1966 and 1986 and are seen to be lower than the other observations in the period prior to 1976. Exclusion of these data points thus leads to an increase in estimated under-5 mortality in the 1970s and 1980s, for example an 11% increase (from 204 to 227 per 1000) in 1970, but changes of 1% or less from 1990 onwards. In Ethiopia the exclusion of more extreme data points reduces the variability observed and hence the uncertainty interval width. In Nigeria the exclusion of more extreme data points similarly reduces the variability observed but an additional consequence is that the short-term trend in estimated under-5 mortality over this period (steep decline) differs markedly from the long-term trend (relatively stable). As a result the overall uncertainty interval width is actually increased in this period, for example a 30% increase (from a relative width of 20% to 26%) in 1970, though less so in more recent times.

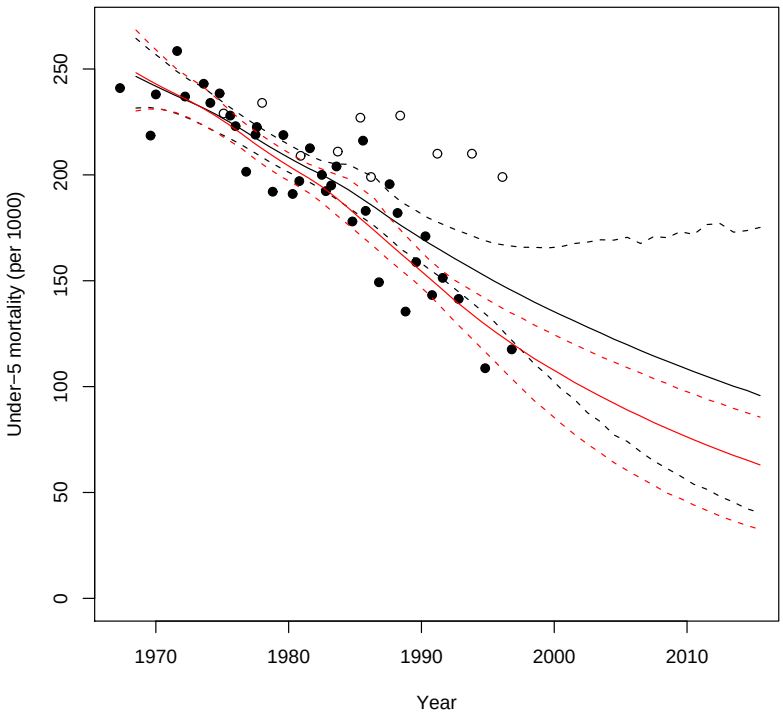

Fig. 11: Comparison of fitted models for Ethiopia. Solid black points represent observed data from surveys providing only direct or indirect estimates, or direct estimates from surveys providing both direct and indirect estimates; hollow points represent observed indirect estimates from surveys providing both direct and indirect estimates; solid black line represents fitted model when using all data (i.e. both solid and hollow points); solid red line represents fitted model when indirect estimates from surveys providing both direct and indirect estimates are excluded (i.e. solid points only); dashed lines represent uncertainty intervals.

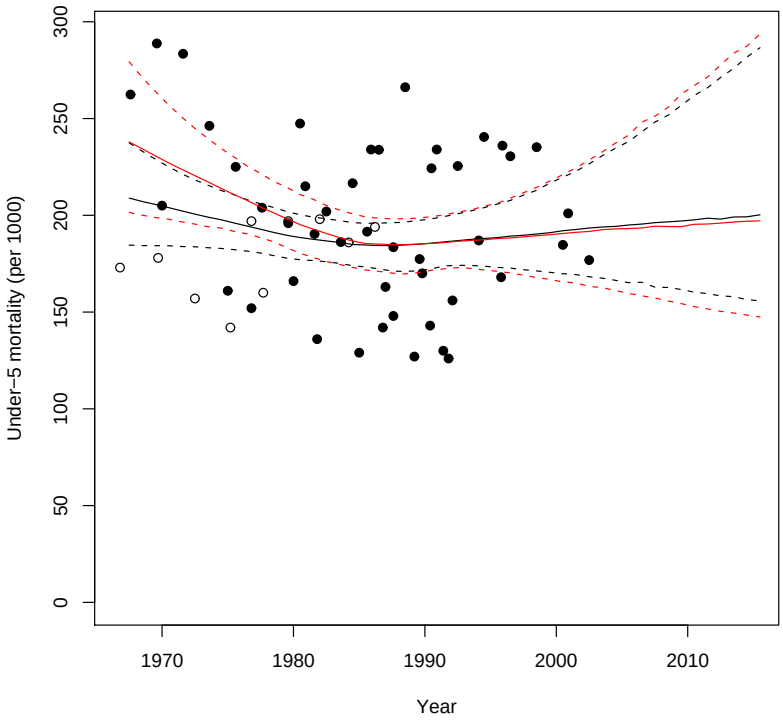

Fig. 12: Comparison of fitted models for Nigeria. Solid black points represent observed data from surveys providing only direct or indirect estimates, or direct estimates from surveys providing both direct and indirect estimates; hollow black points represent observed indirect estimates from surveys providing both direct and indirect estimates; solid black line represents fitted model when using all data (i.e. both solid and hollow points); solid red line represents fitted model when indirect estimates from surveys providing both direct and indirect estimates are excluded (i.e. solid points only); dashed lines represent uncertainty intervals.

Discussion

Using both direct and indirect estimates derived from the same survey in the estimation under-5 mortality, as is presently the case, is a questionable approach. We have reanalysed the available data for 65 countries where this is an issue using only the direct estimates in these situations.

In most countries this alternative approach makes relatively little difference to either the estimated under-5 mortality or the associated uncertainty. Somewhat surprisingly, the percentage change in estimates does not seem to appear to be strongly correlated with the proportion of data points which are excluded (results not shown).

One country (Bhutan) with both direct and indirect estimates derived from the same survey has insufficient data for analysis once the relevant indirect data are excluded. Where this is the case we recommend that, unless further data are available, estimates will need to incorporate expert opinion and should acknowledge that fact in a clear fashion.

It should also be noted that the exclusion of data points may affect the range of α values which can be used in the estimation process, particularly in countries where a large proportion of data points are excluded and only a small number remain. In these instances the set of α values to be used may be found by following the procedure described by Murray et al [2] in the usual manner. Restricting the present analysis to the 24 countries where the set of α values remained unchanged provides values almost identical to those in the unrestricted analysis in Table 3 (results not shown). This suggests that the observed results are not just an artifact of the changing sets of α values.

The need for a multilevel approach?

An important technical point, not discussed in the original publication [2], is that when survey data have been utilised the data from a single survey have often been used to produce multiple data points. For example, a survey conducted in 2005 may have been used to construct estimates of under-5 mortality at several different time points from the early 1980s up to 2005. The approach suggested by Murray et al, referred to here as the 'conventional loess-based approach', assumes independence of all data points. However, when several data points come from the same source, this assumption of independence is unlikely to hold. For example, under-5 mortality in 1990 estimated from a given Demographic and Health Survey (DHS) may be more similar to under-5 mortality in 2000 estimated from the same DHS than to under-5 mortality in 2000 estimated from, say, a Multiple Indicator Cluster Survey (MICS). In this way, the data can be considered as belonging to 'clusters' defined by their source.

One way to relax the assumption of independence is through the use of multilevel modelling [10]. Multilevel models allow data to be viewed as a series of levels nested within one another to form a hierarchy. Explicitly defining the structure in this way as part of the modelling process enables the clustering effects to be accounted for.

We propose a random intercepts multilevel loess-based model for the estimation of under-5 mortality and apply it to the data used by Murray et al for countries with multiple data sources. We compare the resulting estimates and uncertainty intervals to those obtained using the conventional approach.

Method

Under-5 mortality data can be considered as a two-level hierarchy, with individual data points nested within data sources. Note that 'data source' in this context is used to differentiate between data sources within which the observed or derived estimates of under-5 mortality may be expected to have some level of dependency. For survey data, data source thus refers not only to the survey type but also to the point at which data were collected, so that, for example,

a DHS conducted in 1995 is considered as a different data source from a DHS conducted in 2000. However, all vital registration (VR) data within a country are considered as coming from a single data source.

Hierarchical data of this nature can be described through the use of multilevel models, the simplest of which is the random intercepts model. The basic loess model (1) can be extended to become a random intercepts model,

$$\begin{aligned} \log(y_{ij}) &= \beta_{0j} + \beta_1 x_{ij} + \beta_2 z_{ij} + \varepsilon_{ij} \\ \beta_{0j} &= \beta_0 + u_j, \end{aligned} \tag{2}$$

where y_{ij}, x_{ij} and z_{ij} are now the ith values within the jth data source, and $u_j \sim N(0, \sigma_u)$ and $\varepsilon_{ij} \sim N(0, \sigma_\varepsilon)$ are both independent and identically distributed, with all u_j independent of all ε_{ij}. This formulation allows the intercept of the fitted model to differ between data sources. Now $\beta_0 + \beta_1 x + \beta_2 z$ gives the 'average' relationship between $\log(y)$ and x and z.

The random intercepts model (2) can again be fitted using a weighted approach, although now via maximum likelihood or restricted maximum likelihood rather than weighted least squares. By using the same weighting function as for the conventional loess model, (2) becomes the random intercepts loess model. Use of this model is referred to here as the 'multilevel loess-based approach'.

The data used by Murray et al are reanalysed using both the conventional and multilevel loess-based approaches. From the 172 countries for which Murray et al provide under-5 mortality estimates, we exclude countries with only a single data source (many countries), countries with two data sources but only a single data point for one of them (Romania, Samoa, OPT) and countries with three or fewer data points in total (Bhutan) as a simple multilevel approach within country cannot be utilised in these instances. After exclusions, 112 countries remain in our analysis.

For each country, the two approaches are used to estimate past trends in under-5 mortality from the earliest to the latest observed data point, and to forecast under-5 mortality from this point until 2015. As in the previous section, where both direct and indirect data are derived from the same survey only the direct data are used. For the conventional approach, the set of α values used for each country are found using the approach detailed by Murray et al [2]. For the multilevel approach, there is an additional condition that α values must lead to the inclusion of more than one data source at all times.

We also compare the performance of two approaches in terms of prediction. In order to do this we identify the most recent non-VR datapoint and use this as the target for prediction. We then remove all data points in the two years prior to the target data point, as well as any other data points coming from the same data source as the target data point. We estimate under-5 mortality and the corresponding uncertainty interval for the year in which the target data point was observed using both the conventional and multilevel loess-based approaches. The estimated uncertainty interval is then examined to see if it contains the target data point, and the occurrence of this is compared between the approaches. As the prediction analysis involves use of a restricted dataset, some countries which generally have sufficient data to allow a multilevel model to be fitted no longer do so, and only 95 countries are included.

Results

Both the conventional and multilevel loess-based approaches for estimating under-5 mortality are applied to the 112 countries. Differences between the resulting estimates across all the countries are examined, but first some individual country results are presented as examples to illustrate how the two approaches can in some circumstances lead to very different results.

Fig. 13 and Fig. 14 show the fitted conventional and multilevel estimates of under-5 mortality for Liberia and Lesotho, respectively. In Liberia (four data sources: Population Growth Survey 1969–70, Population Growth Survey 1970–71, Census 1974 and DHS 1986), the conventional loess-based approach to estimation, which assumes independence of the data (i.e. ignoring the connections between the data points in Fig. 13), results in a relatively low rate of decrease in under-5 mortality. However, it is clear from examining the connected data points that the observed rate of decrease in under-5 mortality within each data source is greater than this. This feature is captured in the more rapidly declining estimates found using the multilevel approach. In Lesotho (ten data sources) the opposite situation arises, whereby the assumption of independence results in the estimation of a fairly rapid decline in under-5 mortality, but the within data source trends are far more stable — indeed, in some more recent data an increasing trend is observed. Again, the estimated under-5 mortality corresponding to the multilevel approach reflects these within-data source trends, with estimates slowly decreasing until the mid-1980s before becoming essentially constant.

One feature common to both Fig. 13 and Fig. 14 is that the uncertainty intervals corresponding to the multilevel approach are much wider than those corresponding to the conventional approach in the region of the plot where data are observed. Indeed, particularly for Lesotho, the uncertainty intervals corresponding to the conventional approach often appear implausibly narrow. Whilst the uncertainty intervals for the conventional estimates rapidly become wider once the point in time being considered is beyond the last data point, those for the multilevel approach retain a similar width, so that that the widths of the uncertainty intervals under the two approaches are more comparable for predicted (as opposed to estimated) values.

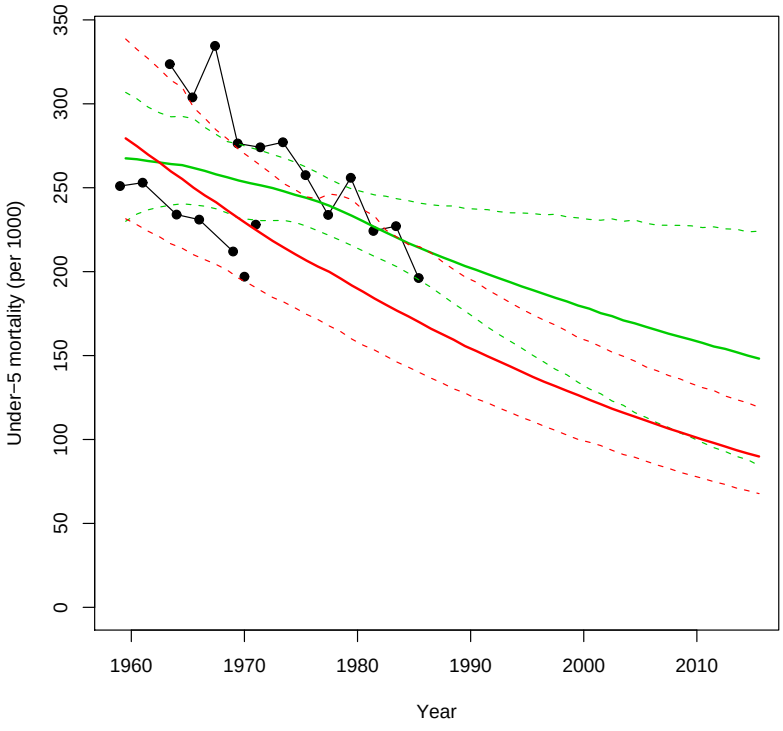

Fig. 13: Comparison of conventional and multilevel loess-based approaches for Liberia. Black points and lines represent observed non-vital registration data; solid green line represents fitted conventional loess model; solid red line represents fitted multilevel loess model; dashed lines represent uncertainty intervals.

There are also many countries where the adoption of a multilevel approach makes little difference to the estimates of under-5 mortality. One such country is Costa Rica, shown in Fig. 15. Here, as the within data source trends in under-5 mortality differ little from the overall trend assuming independence, the conventional and multilevel estimates

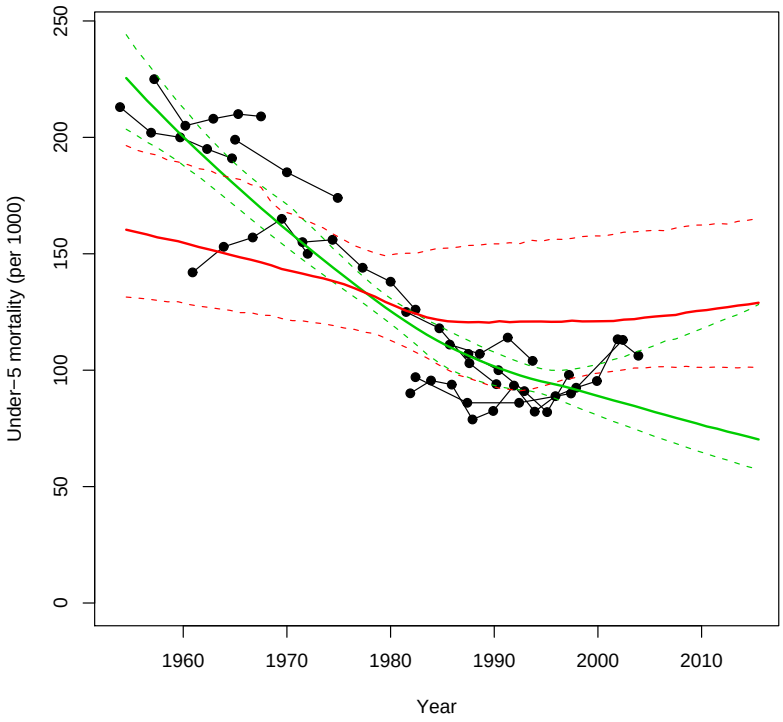

Fig. 14: Comparison of conventional and multilevel loess-based approaches for Lesotho. Black points and lines represent observed non-vital registration data; solid green line represents fitted conventional loess model; solid red line represents fitted multilevel loess model; dashed lines represent uncertainty intervals.

are virtually identical. As with the other examples, however, the uncertainty intervals corresponding to the multilevel estimates appear wider over much of the plot.

The percentage changes in both the estimates themselves and the relative uncertainty interval widths when comparing the multilevel to the conventional approach for Liberia, Lesotho and Costa Rica are displayed in Table 4. In Liberia, adoption of the multilevel approach leads to reduced estimates of under-5 mortality in every year from 1970 onwards — a 31% reduction in 2000 and a predicted 39% decrease for 2015. Whilst the multilevel uncertainty intervals are much wider in Liberia between 1960 and 1990 (almost twice the relative width in 1980), by 2010 they are considerably narrower. Use of the multilevel approach in Lesotho results in increased estimates of under-5 mortality from 1980 onwards — 38% in 2000 and 83% in 2015. Here, the uncertainty intervals under the multilevel approach are 3–4 times the relative width prior to 2000, but narrower from 2010 onwards. Estimated under-5 mortality in Costa Rica can be seen to be essentially the same under the two approaches though, as was noted from the plot, uncertainty is increased somewhat for the years prior to 1990.

In order to examine the differences between the two approaches across all 112 countries, plots of the multilevel estimates against the conventional estimates in 1970, 1990 and 2010 are provided in Fig. 16, Fig. 17 and Fig. 18, respectively. In both 1970 and 1990, time points within the period for which empirical data are available for most countries, there are only a handful of countries in which the estimates under the two approaches differ noticeably, but by 2010, beyond the range of available data, there are several countries where the differences are clearly substantial. However, there is little evidence that one approach produces systematically higher estimates than the other.

Fig. 19, Fig. 20 and Fig. 21 show multilevel relative uncertainty interval width plotted against conventional relative uncertainty width for 1970, 1990 and 2000, respectively. In 1970 and 1990 virtually every country has greater relative uncertainty surrounding the estimate using the multilevel approach, and in some countries the relative uncertainty

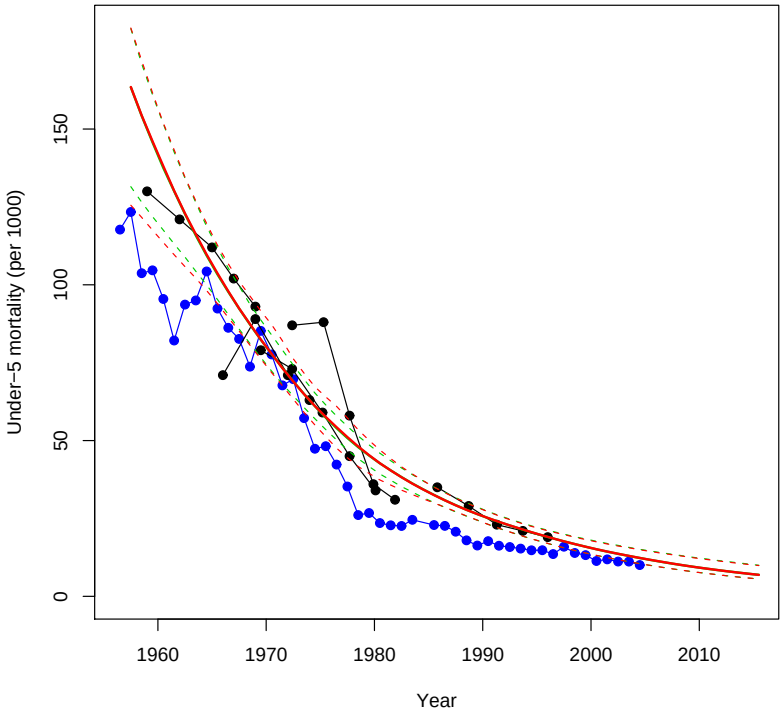

Fig. 15: Comparison of conventional and multilevel loess-based approaches for Costa Rica. Black points and lines represent observed non-vital registration data; solid green line represents fitted conventional loess model; solid red line represents fitted multilevel loess model; dashed lines represent uncertainty intervals.

	Liberia		Lesotho		Costa Rica	
Year	Estimate	UI width	Estimate	UI width	Estimate	UI width
1960	+2.8	+48	−22.2	+227	+0.2	+14
1970	−10.1	+96	−10.0	+172	+0.0	+32
1980	−18.4	+183	+2.8	+230	+0.0	+43
1990	−24.1	+38	+19.7	+272	+0.0	−2
2000	−30.5	−15	+37.5	+83	+0.2	−4
2010	−36.5	−35	+66.0	−33	+0.2	−3
2015	−39.4	−38	+82.7	−49	+0.1	−3

Table 4: Percentage (%) change in estimated under-5 mortality (U5M) and relative under-5 mortality uncertainty interval (UI) width when going from conventional to multilevel loess-based approach in a selection of example countries. UI is uncertainty interval.

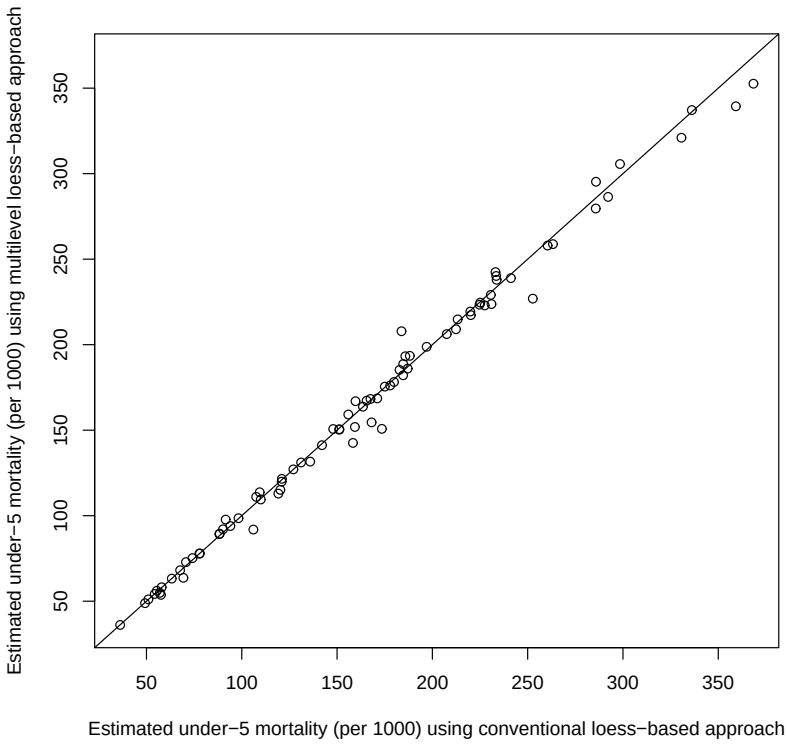

Fig. 16: Comparison of estimated under-5 mortality in 1970 when using the conventional and multilevel loess-based approaches.

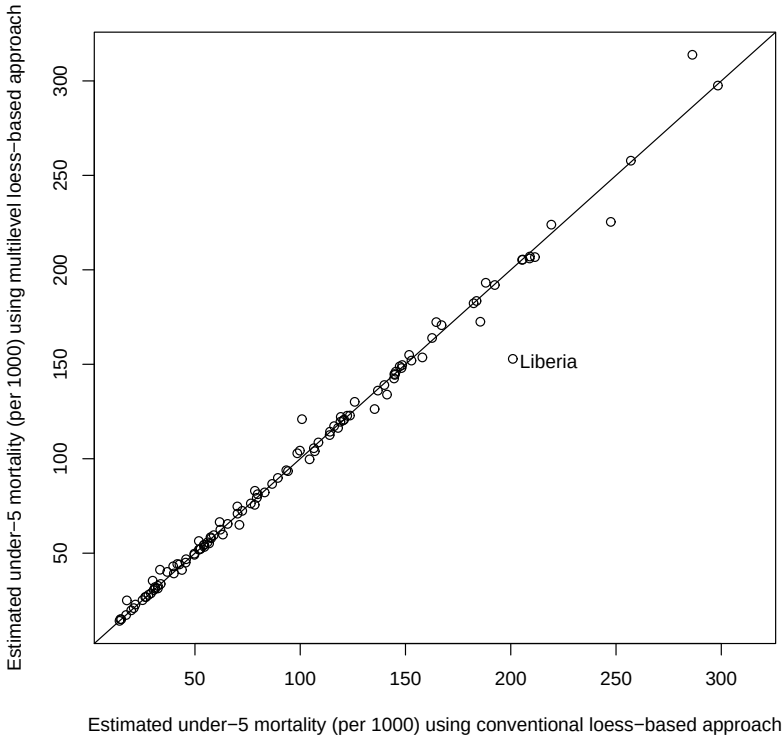

Fig. 17: Comparison of estimated under-5 mortality in 1990 when using the conventional and multilevel loess-based approaches, with outlier identified.

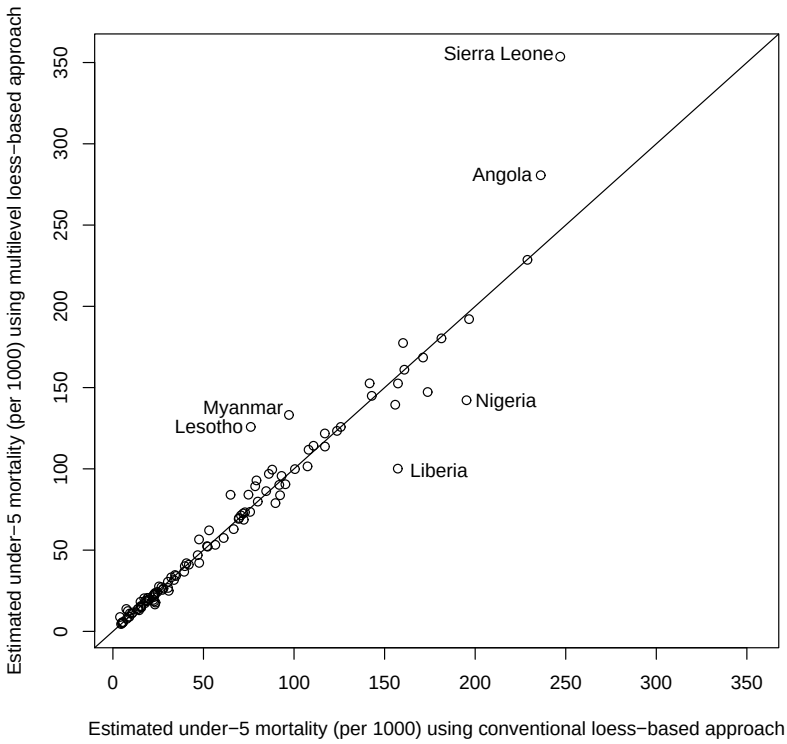

Fig. 18: Comparison of estimated under-5 mortality in 2010 when using the conventional and multilevel loess-based approaches, with outliers identified.

interval width is several times greater than that for the corresponding conventional estimate. By 2010, however, although there remain differences between relative uncertainty interval width under the two approaches in most countries, there are similar numbers of countries where multilevel relative uncertainty is the greater and where conventional relative uncertainty is the greater. Additionally, there are no countries where either relative uncertainty interval width is many times greater than the other. Again, the very high relative uncertainty interval width seen in some countries is due to low under-5 mortality estimates rather than particularly high absolute uncertainty.

Table 5 summarises the percentage difference between the estimates using the multilevel approach and the estimates using the conventional approach across all 112 countries. The number of countries contributing at each time point again differs as estimates are only calculated for time points at least as recent as the first observed data point for each country. The median percentage difference in the estimates does not differ appreciably from 0 in any year. The first and third quartiles of percentage change in estimated under-5 mortality shows that in the middle 50% of countries estimates differ by less than 3% between 1970 and 2000, though variability increases somewhat in years towards or beyond the limits of the observed data. The percentage change in relative uncertainty interval width, however, is marked, with a median increase of at least a 60% at each year between 1960 and 1990, although this is much reduced in 2000 and by 2010 there is essentially, on average, no difference. The variability in the percentage change in relative uncertainty interval width also decreases as time progresses.

Table 6 shows the results of the prediction analysis. It should be noted that the uncertainty intervals are for the 'true' value of under-5 mortality whereas the observed values are estimates of the true value with associated standard error. Thus although the uncertainty intervals are approximate 95% intervals, the observed values would be expected to lie within their bounds somewhat less than 95% of the time. Using the conventional loess-based approach the target data point is included within the estimated uncertainty interval 60 out of 95 times (63%), whereas using the multilevel loess-based approach it is within the uncertainty interval 64 times out of 95 (67%). These results suggest that the

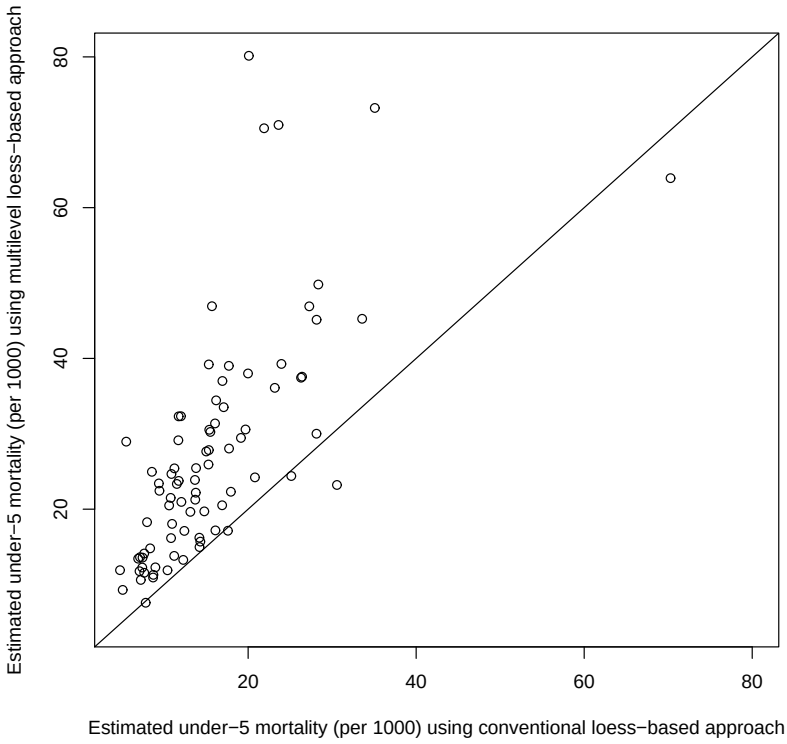

Fig. 19: Comparison of relative under-5 mortality uncertainty interval width (uncertainty interval width as a percentage of the corresponding estimate of under-5 mortality) in 1970 when using the conventional and multilevel loess-based approaches.

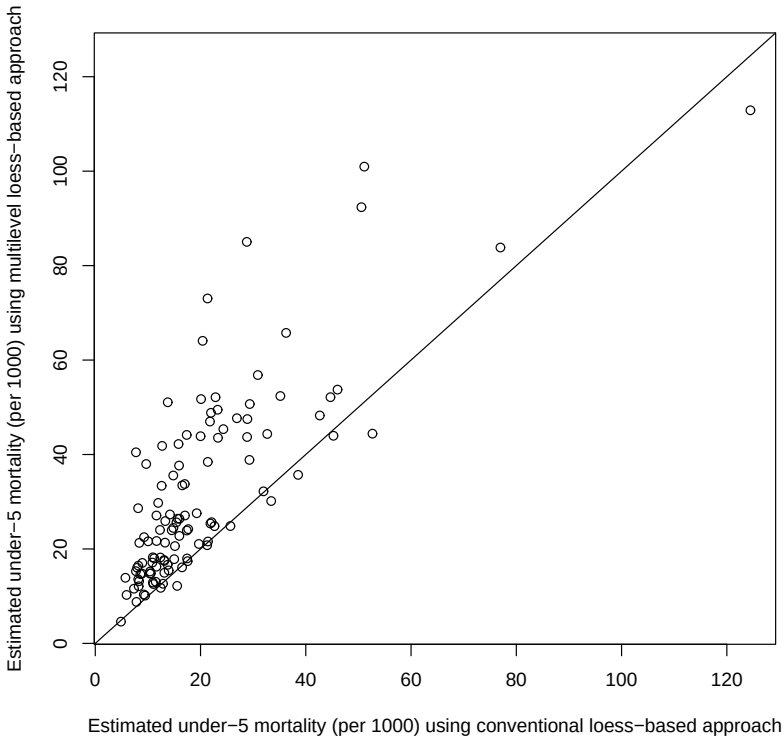

Fig. 20: Comparison of relative under-5 mortality uncertainty interval width (uncertainty interval width as a percentage of the corresponding estimate of under-5 mortality) in 1990 when using the conventional and multilevel loess-based approaches.

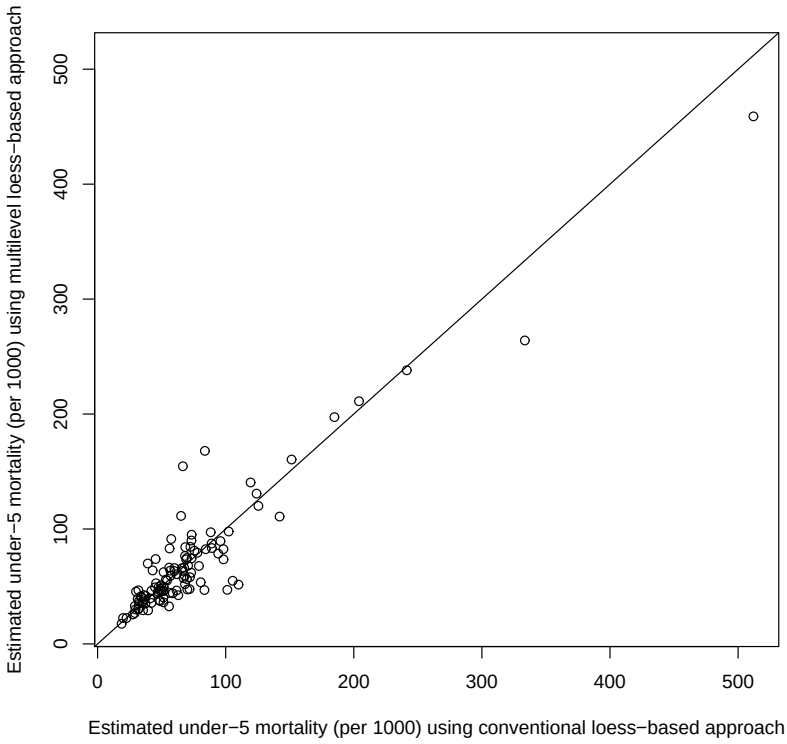

Fig. 21: Comparison of relative under-5 mortality uncertainty interval width (uncertainty interval width as a percentage of the corresponding estimate of under-5 mortality) in 2010 when using the conventional and multilevel loess-based approaches.

		% change in estimated U5M			% change in relative U5M UI width		
Year	No. countries	Q_1	Median	Q_3	Q_1	Median	Q_3
1960	51	−3.6	−0.9	+0.8	+20	+60	+102
1970	82	−1.5	−0.2	+1.1	+37	+73	+103
1980	102	−0.3	+0.0	+1.7	+34	+68	+118
1990	111	−0.8	+0.0	+1.4	+16	+61	+98
2000	111	−1.7	+0.1	+2.6	+1	+24	+44
2010	111	−3.0	−0.1	+3.8	−15	−2	+10
2015	111	−3.9	−0.3	+5.0	−17	−5	+5

Table 5: Percentage (%) change in estimated under-5 mortality (U5M) and estimated under-5 mortality uncertainty interval (UI) width when going from conventional to multilevel loess-based approach. Q_1 is the first quartile, Q_3 is the third quartile.

multilevel approach may have slightly better predictive capabilities though the difference is not large.

		Conventional approach		
		Included	Not included	Total
Multilevel approach	Included	52	12	64
	Not included	8	23	31
	Total	60	35	95

Table 6: Number of occasions on which the target data point is included or not included within the estimated uncertainty interval in the prediction analysis.

When deciding upon the set of α values to use for the multilevel approach we introduced an additional condition that α values must lead to the inclusion of more than one data source at all times, which may mean that in a given country the conventional and multilevel approaches are applied using different sets of α values. However, we found that this is in fact only the case in a single country out of the 112 examined, so there is no problem with the comparability of the results.

Discussion

The loess-based approach to the estimation of under-5 mortality posited by Murray et al [2] assumes the independence of all data points. However, as several data points often come from the same source, this assumption of independence is unlikely to hold. One possible approach which allows this clustering of data to be accounted for is multilevel modelling.

Of the 172 countries for which Murray et al provided under-5 mortality estimates, we have reanalysed the 112 with multiple data sources using both the 'conventional' loess-based approach and a multilevel equivalent. Estimated under-5 mortality under the two approaches is generally found to be similar, though in some cases the differences are substantial. The estimated uncertainty surrounding the estimates, however, almost always increases markedly under the multilevel approach. Although some of the examples cited above are chosen to exemplify the potential disparity between estimates resulting from the two approaches and are thus far from typical, it is clear that this choice of methodology can have a large impact on the results obtained.

The two approaches will result in widely differing estimates in countries where there is homogeneity in the observed within-data source trends in under-5 mortality, but where these within-data source trends differ from the trend estimated assuming independence. The multilevel approach is likely to lead to wider uncertainty intervals within the time period for which data are observed, but this is particularly true in countries where there is heterogeneity in the within-data source trends.

Whilst the use of a multilevel approach is a more correct way of dealing with dependencies within the data, it is not clear that the estimates produced are necessarily superior. If the estimated under-5 mortality is similar under the two approaches this probably reflects comparable validity of the point estimates derived by the two approaches, and the (generally) wider uncertainty intervals provided by the multilevel approach would seem to be preferable to the often implausibly narrow confidence intervals provided by the conventional approach. However, if estimated under-5 mortality differs markedly between the two approaches due to the within-data source trends differing from the trend assuming independence then a multilevel approach is only appropriate if the within-data source trends are unbiased. One way in which large differences could arise between the within-data source trends and the trend assuming independence is through 'date shifting' which would bias the within-data source trends. In such circumstances the trend

assuming independence may represent a truer reflection of reality. It is thus our recommendation that both conventional and multilevel estimates should be calculated, with countries for which the sets of estimates differ markedly being identified for further consideration as to why this is the case.

A random slopes and intercepts loess model may be considered preferable to the random intercepts loess model utilised here as it would allow the slope as well as the intercept to vary between data sources. However, application of a random slopes and intercepts model for the estimation of under-5 mortality would necessitate the exclusion of many countries where there is insufficient data to fit this type of model. A random intercepts only model is therefore preferred. In countries with sufficient data to fit a random slopes and intercepts loess model, the resulting estimates are generally very similar to those when using a random intercepts only loess model (results not shown).

In countries with a very small number of data sources even the use of a random intercepts only model may be questioned. However, as we are only interested in the fixed effects, the number of level-2 units is of less concern than may be the case in other applications. Even with these concerns, a multilevel approach is a more correct way of handling the data structure, and as more data sources are added over time countries where this is an issue will become fewer.

As a multilevel approach cannot be utilised in countries with only a single source of data (mainly those countries which rely solely on VR data), the conventional approach may have to be retained in these cases. However, it should be noted that in such circumstances the uncertainty intervals will not account for between-data source variation, and hence the uncertainty intervals for such countries will tend to be narrower, perhaps inappropriately so, than for countries with multiple data sources. A further extension would be to develop a method by which uncertainty interval width in these cases can be inflated in an appropriate manner.

Further issues

There are further outstanding issues with the loess-based approach to estimating under-5 mortality. Murray et al, in their original analysis [2], exclude any α values which result in an annual rate of decline in under-5 mortality more than 3 standard deviations from the mean rate across their empirical dataset. Consequently, only α values corresponding to annual rates of decline less than 8.94% or annual rates of increase less than 1.80% are included. Whilst it may be sensible to place a constraint on the admissible α values in order to avoid spuriously high rates of decline or increase corresponding to low α values, to exclude α values which result in even a 2% annual rate of increase in under-5 mortality may be considered overly-restrictive, especially in countries which have experienced conflict or high levels of HIV. The above analyses use the method described by Murray et al to select the set of α values to use, including the imposition of this growth rate constraint.

An additional issue with the loess-based approach is that when predicting under-5 mortality at time points beyond the most recent observed data point the absolute uncertainty interval width decreases as time progresses in a small proportion of countries, which may seem somewhat counter-intuitive. However, if relative rather than absolute uncertainty interval widths are considered then no such decrease in uncertainty interval width is observed. As can be seen from Fig. 13, narrowing absolute uncertainty interval width for predicted under-5 mortality may also occur when using the multilevel loess-based approach.

Also of concern is how estimates and their associated uncertainty can be optimally presented. Although a plot of the data for a country overlaid with the estimates and upper and lower uncertainty interval limits illustrates the final results, it does not enlighten as to the fitted loess trajectories which contribute to them. We propose additionally plotting the estimated trajectories corresponding to the (country-specific) minimum and maximum α values used in

the estimation process. This is illustrated in Fig. 22 for Lesotho. Both trajectories appear to be acceptable fits to the data, though one takes a short-term perspective and one a long-term perspective. However, the overall estimates are clearly closer to the trajectory corresponding to the maximum α value, perhaps suggesting an overweighting towards long-term trends.

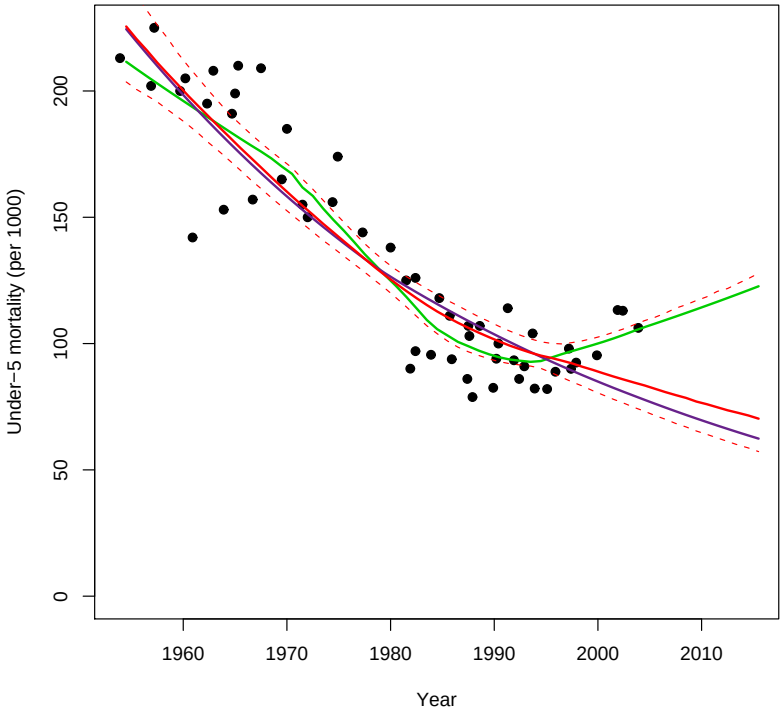

Fig. 22: Estimated under-5 mortality in Lesotho. Black points represent observed non-vital registration data; green and purple lines represent the estimated trajectories corresponding to the minimum (= 0.5) and maximum (= 2.0) α values respectively; solid red line represents the overall estimates; dashed red lines represent overall uncertainty intervals.

More generally, the extent to which the overall estimates compromise between the short- and long-term effects is determined by the set of α values which are used. Different alpha-value sets could be chosen which would give more (or less) weight to relatively short-term trends. The minimum α value which can be used is defined by data availability (and also by the growth rate constraint described above), but the maximum α value and the method for selecting which values to use between the minimum and the maximum is essentially arbitrary. It has been suggested that perhaps there is an overweighting towards long-term trends, as is seen in Fig. 22. This could be remedied by reducing the maximum α value or increasing the density of lower α values relative to the density of higher α values.

The current implementation of the loess-based approach implicitly assumes that sampling variability is the same for each data point, but this is unlikely to be true. A further extension to the approach could be to incorporate measures of sampling variability into the modelling process so that data points with low sampling variability (i.e. those estimated from very large samples) are given more weight relative to those with high sampling variability (i.e. those estimated from small samples).

All our analyses and results assume that the data points (with the exception of VR data) are unbiased estimates of the true under-5 mortality. However, in some situations this assumption may not hold. One country where the validity of this assumption may be doubted is Congo (see Fig. 1) where the data points (all estimated from the same 2005 DHS) indicate an increasing trend in under-5 mortality which is at odds with the trends seen in the surrounding countries over this period. Bias is clearly an important issue which may affect our ability to provide reliable overall

estimates, but the identification and treatment of biased data points are problems which remain largely unresolved. In their original analysis Murray et al [2] do exclude some data points (which are consequently also excluded from our analyses) on the basis of acknowledged bias, as well as some 'extreme outliers that clearly differ from the rest of the data points', which may also indicate bias.

As estimated levels of child mortality have been the focus of much recent attention, it is imperative that they are estimated using appropriate statistical methodology. We believe that Murray et al's work represents an important step forward, particularly in its attempt to provide uncertainty estimates which take account of both uncertainty due to sampling variability and uncertainty due to (one aspect of) model specification.

However, it is essential that the proposed method is implemented correctly and that further development of the method is undertaken in the areas discussed in this paper.

Acknowledgements

This work was funded by grants from Unicef and the World Health Organisation.

References

[1] United Nations. *United Nations Millennium Declaration, Resolution 52/2.* United Nations, New York, 2000.

[2] C. J. Murray, T. Laakso, K. Shibuya, K. Hill, and A. D. Lopez. Can we achieve Millennium Development Goal 4? New analysis of country trends and forecasts of under-5 mortality to 2015. *Lancet*, 370(9592):1040–54, 2007.

[3] H. Masanja, D. de Savigny, P. Smithson, J. Schellenberg, T. John, C. Mbuya, G. Upunda, T. Boerma, C. Victora, T. Smith, and H. Mshinda. Child survival gains in Tanzania: analysis of data from demographic and health surveys. *Lancet*, 371(9620):1276–83, 2008.

[4] W. S. Cleveland. Robust locally weighted regression and smoothing scatterplots. *J Am Stat Assoc*, 74(368):829–36, 1979.

[5] W. S. Cleveland and S. J. Devlin. Locally weighted regression: an approach to regression analysis by local fitting. *J Am Stat Assoc*, 83:596–610, 1988.

[6] Institute for Health Metrics and Evaluation. Worldwide Child Mortality (Mortality Data). `http://www.healthmetricsandevaluation.org/print/mortality.xls`, accessed November 2007.

[7] T. Wardlaw. New estimates of global under-5 mortality. *Lancet*, 370(9596):1413, 2007.

[8] Institute for Health Metrics and Evaluation. Worldwide Child Mortality (Estimates and Projections). `http://www.healthmetricsandevaluation.org/print/estimates.xls`, accessed November 2007.

[9] World Health Organisation. Estimated coverage of mortality data for latest year. `http://www.who.int/whosis/database/mort/table4.cfm`, accessed July 2008.

[10] H. Goldstein. *Multilevel Statistical Models.* Hodder Arnold, London, 2003.