

# Estimating uncertainty in geospatial modelling at multiple spatial resolutions: the pattern of delivery via caesarean section in Tanzania

Corrine Warren Ruktanonchai,<sup>1</sup> Jeremiah J Nieves,<sup>1</sup> Nick W Ruktanonchai,<sup>1</sup> Kristine Nilsen,<sup>1</sup> Jessica E Steele,<sup>1</sup> Zoe Matthews,<sup>2</sup> Andrew J Tatem<sup>1</sup>

**To cite:** Ruktanonchai CW, Nieves JJ, Ruktanonchai NW, *et al*. Estimating uncertainty in geospatial modelling at multiple spatial resolutions: the pattern of delivery via caesarean section in Tanzania. *BMJ Global Health* 2020;**4**:e002092. doi:10.1136/bmjgh-2019-002092

**Handling editor** Seye Abimbola

► Additional material is published online only. To view please visit the journal online (<http://dx.doi.org/10.1136/bmjgh-2019-002092>).

Received 17 October 2019  
Revised 2 January 2020  
Accepted 9 January 2020

## ABSTRACT

Visualising maternal and newborn health (MNH) outcomes at fine spatial resolutions is crucial to ensuring the most vulnerable women and children are not left behind in improving health. Disaggregated data on life-saving MNH interventions remain difficult to obtain, however, necessitating the use of Bayesian geostatistical models to map outcomes at small geographical areas. While these methods have improved model parameter estimates and precision among spatially correlated health outcomes and allowed for the quantification of uncertainty, few studies have examined the trade-off between higher spatial resolution modelling and how associated uncertainty propagates. Here, we explored the trade-off between model outcomes and associated uncertainty at increasing spatial resolutions by quantifying the posterior distribution of delivery via caesarean section (c-section) in Tanzania. Overall, in modelling delivery via c-section at multiple spatial resolutions, we demonstrated poverty to be negatively correlated across spatial resolutions, suggesting important disparities in obtaining life-saving obstetric surgery persist across sociodemographic factors. Lastly, we found that while uncertainty increased with higher spatial resolution input, model precision was best approximated at the highest spatial resolution, suggesting an important policy trade-off between identifying concealed spatial heterogeneities in health indicators.

## INTRODUCTION

Achieving the Sustainable Development Goal aims laid out in 2015 necessitates measurement of health outcomes at small geographical areas to ensure ‘no one left behind’.<sup>1</sup> With recent advancements in the collection and distribution of geo-located household surveys, such as those collected via the Demographic and Health Survey (DHS) programme ([www.dhsprogram.com](http://www.dhsprogram.com)), researchers are increasingly using methods such as small area estimation and geostatistical additive models (GAMs) to generate high spatial resolution maps of health and

## Key questions

### What is already known?

- Estimating maternal and newborn health outcomes at small geographical areas is increasingly important in identifying hidden pockets of health inequalities. The use of Bayesian geostatistical models has allowed for the quantification of associated uncertainty with these modelled estimates.

### What are the new findings?

- The trade-off between increasing spatial resolution in model inputs/outcomes and associated uncertainty has not been explored, particularly among maternal and newborn health outcomes.

### What do the new findings imply?

- While uncertainty in model outcomes increases with increasing spatial resolution, model precision was best approximated at the finest spatial resolution for prevalence of delivery via caesarean section in Tanzania. These findings imply an important trade-off between identifying concealed spatial heterogeneities and accuracy of estimates, which should be communicated in policy relevant settings.

development indicators.<sup>2–4</sup> Such subnational, high-resolution estimates have become useful tools for researchers and policy makers alike in uncovering hidden health inequities that would otherwise be masked by aggregate or national-level health indicators, enabling targeted interventions in settings with limited resources.<sup>1 5–8</sup>

Visualising health outcomes and associated uncertainty at high spatial resolutions has distinct policy relevance among maternal and newborn health (MNH) outcomes,<sup>9</sup> as maternal and neonatal mortality both vary geographically and occur relatively rarely. Furthermore, the data associated with maternal mortality are subject to limitations, misclassification and bias,<sup>10 11</sup> particularly



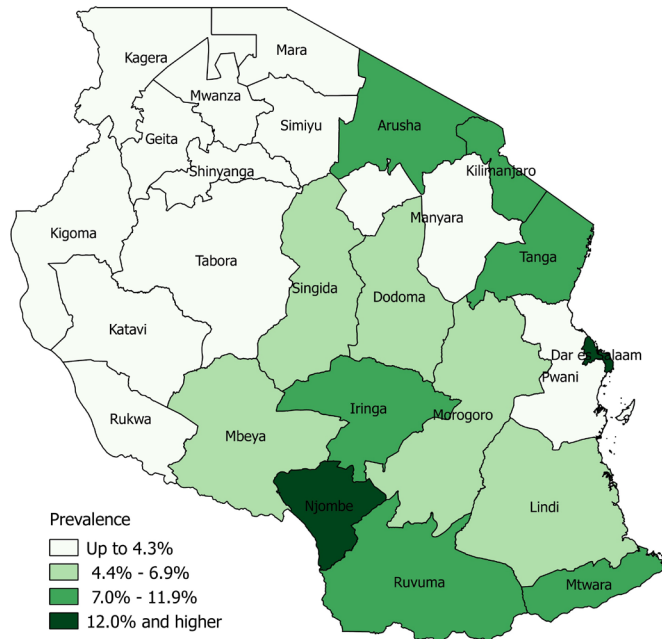
© Author(s) (or their employer(s)) 2020. Re-use permitted under CC BY. Published by BMJ.

<sup>1</sup>School of Geography & Environmental Science, University of Southampton, Southampton, UK

<sup>2</sup>Department of Social Statistics & Demography, University of Southampton, Southampton, UK

### Correspondence to

Corrine Warren Ruktanonchai; [cwr1g15@soton.ac.uk](mailto:cwr1g15@soton.ac.uk)



**Figure 1** Delivery by caesarean section at the administrative 1 level using DHS data, Tanzania, 2015.

within more rural areas of sub-Saharan Africa where many deaths do not occur at hospitals and may go unrecorded.<sup>12</sup> As with maternal mortality, data on life-saving MNH interventions such as antenatal care, skilled birth attendance and delivery via caesarean section (c-section) can be widely obtained at aggregate levels but remain difficult to measure at subnational levels, especially in the most rural and vulnerable areas of the world. While some work has been done modelling key MNH interventions at subnational scales such as maternal health services, exclusive breastfeeding, childhood vaccinations and health systems performances,<sup>5 6 13–15</sup> other vital life-saving interventions that occur less frequently, such as delivery via c-section, have not been modelled previously at high spatial resolutions.

With advancements in computational resources and data availability over recent decades, researchers across disciplines have begun employing Bayesian GAM to map disease and quantify uncertainty in posterior model outcomes, particularly using hierarchical clustered data such as from the DHS.<sup>3 4 16–18</sup> The application of these methods is increasingly pertinent, as access to healthcare services is heterogeneously distributed across landscapes, requiring high-resolution spatial data and modelling techniques to identify the most vulnerable populations. However, these methods and associated spatial data carry limitations and bias, manifesting in uncertainty that should be adequately quantified and communicated to decision makers and non-academic audiences for optimum policy impact.<sup>19</sup> While the use of such GAMs to predict high-resolution health outcomes has improved model parameter estimates and precision among spatially correlated and rare adverse health outcomes<sup>3 20</sup> and allowed for this quantification of uncertainty,<sup>21</sup> no studies

have examined the trade-off between predicting health outcomes at higher spatial resolutions and visualising the spatial distribution of associated uncertainty.<sup>22</sup>

Here, we estimate prevalence of delivery via c-section in Tanzania, using input covariates at varying levels of spatial coarseness within a Bayesian geostatistical model framework. With these models, we investigate how uncertainty varies with spatial resolution, and how this changing uncertainty can be better visualised and communicated. Specifically, we explore the trade-off between model estimates and associated uncertainty at increasing spatial resolutions through exploration of the posterior distribution of modelled delivery via c-section at multiple spatial resolutions.

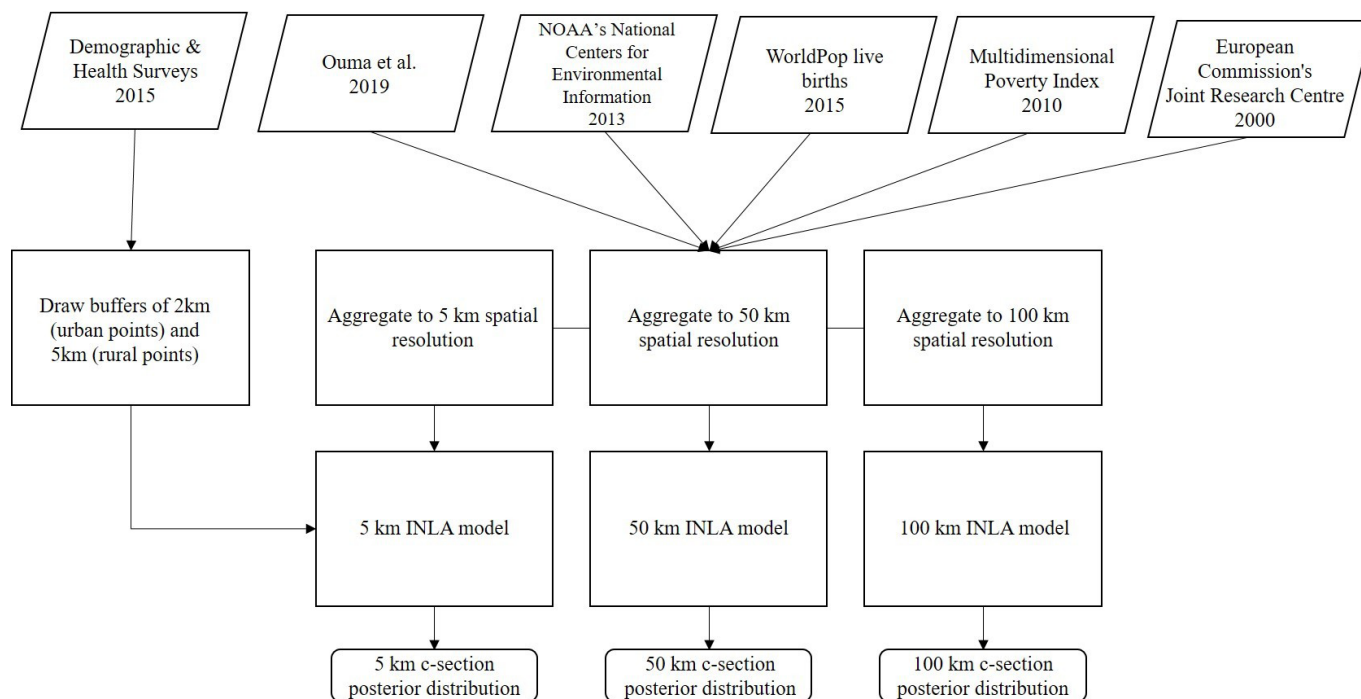
## METHODS

### DHS data

We compiled DHS data from Tanzania for 2015<sup>23</sup> using SAS V.9.4 software<sup>24</sup> and restricted the sample to women with a birth in the preceding 5 years (n=7050 women) with corresponding spatial data, as provided by DHS cluster locations. Briefly, the DHS provides global positioning system (GPS) coordinates for clusters of aggregated household survey data in order to facilitate spatial analyses while also maintaining participant confidentiality. These coordinates are displaced up to 2 km in urban areas and 5 km in rural areas, with up to 1% of points displaced up to 10 km in rural areas.<sup>25</sup> Using these geolocated cluster locations, spatial inference occurs at a higher spatial resolution than the geographic region in which the survey is designed to be representative of. This hierarchical sample design therefore necessitates the use of geostatistical models to make inferences at spatial resolutions finer than the DHS region level.<sup>3</sup> In these analyses, our binary outcome of interest was defined as the number of women who underwent any delivery via c-section for a birth within the preceding 5 years (regardless of whether it was the most recent), as compared with women who had not experienced delivery via c-section for any preceding birth. To maintain survey representativeness, this was calculated at the DHS level 1 resolution, representing the 30 administrative I regions of Tanzania (figure 1).

### Covariate data

In addition to demographic data gathered through the DHS, we also compiled environmental geospatial covariate data that we extracted to DHS cluster locations (figure 2). Because these locations are displaced, we averaged geospatial covariates to 2 km and 5 km buffers for urban and rural locations, respectively. While up to 1% of coordinates in rural areas are displaced within a 10 km radius, the addition of buffers at the 10 km level has been shown to impact very few rural coordinates,<sup>25</sup> while unnecessarily introducing bias in environmental covariates, thereby justifying use of a 5 km buffer in rural areas. First, we gathered data from the European Commission's



**Figure 2** Study analysis flow chart. INLA, Integrated Nested Laplace Approximation; NOAA, National Oceanic and Atmospheric Administration.

Joint Research Centre on accessibility to major cities for the year 2000, representing travel time to the nearest city exceeding a population of 50 000 using land (ie, roads) or water-based travel (ie, rivers and lakes).<sup>26</sup> Next, we included data on annual night light intensity for the year 2013 (an indicator of urbanicity), as generated by the National Oceanic and Atmospheric Administration's National Centers for Environmental Information.<sup>27</sup> We also included live births for the year 2015 at the 1 km resolution, as well as Multidimensional Poverty Index estimates for the year 2010, as obtained via the WorldPop Project ([www.worldpop.org.uk](http://www.worldpop.org.uk)) and outlined by Tatem *et al.*<sup>28,29</sup> Lastly, we included travel time to the nearest public hospital, as outlined by Ouma *et al.*,<sup>30</sup> calculated through a cost–distance algorithm incorporating a travel impedance surface by assigning travel speed to road networks.

These covariates were chosen as previous studies have shown them to be predictive of MNH outcomes and risks,<sup>6,31–34</sup> representing a suite of geospatial covariates with robust predictive power to examine how uncertainty changes as a function of spatial resolution. Notably, we chose to include only geospatial covariates in this model as we could vary the spatial resolution of these variables, thereby addressing our research objectives. Specifically, these datasets were compiled at the 1 km spatial resolution and subsequently averaged at the 5 km, 50 km and 100 km resolutions to facilitate projecting the fitted model onto gridded surfaces at these levels. These surfaces represent a theoretical exploration of the trade-off between increasing gridded spatial resolution and modelled estimates and were chosen to clearly illustrate the difference

in the practical range of estimates, as shown in online supplementary figure S2.

### Model framework

To explore how uncertainty in posterior modelled c-section delivery estimates varied at multiple spatial resolutions, we employed a Bayesian hierarchical model framework with input covariates at varying spatial coarseness. These models have been used extensively with DHS and other household survey data,<sup>1,2,35–38</sup> as they are able to robustly account for the multistage sampling efforts employed during the data collection process, resulting in hierarchically structured data provided through the DHS. Here, our model accounts for the nested structure of DHS data by allowing for variation in the  $n^{\text{th}}$  region among individual respondents, as outlined below. These models were fit independently of each other, resulting in three models with input covariates and modelled outcomes at the 5 km, 50 km and 100 km spatial scales. To predict c-section delivery at a continuous spatial resolution, we implemented these models via stochastic partial differential equation (SPDE) spatial regression approach, implemented using the Integrated Nested Laplace Approximation (INLA) technique within the R-INLA package.<sup>39</sup> This approach was suitable for this analysis as these spatial processes are generally well captured by a Gaussian field with Matérn correlation.<sup>40</sup> These models have similarly been used in previous research combining DHS data and geospatial covariates to predict high-resolution childhood vaccination coverage by disaggregating areal surveillance data.<sup>1,2,34</sup> We employed a similar modelling framework, generally defined as

$$Y_i \sim \text{Binomial}(N_i, p_i), i = 1, \dots, n_A, n_A + 1, n_A + n_p$$

where  $n_A$  represents subnational DHS regions within Tanzania to maintain survey representation;  $Y_i$  represents the number of women delivering via c-section within area,  $A_i$ ;  $p_i$  represents the probability of a woman delivering via c-section over grid points  $n_p$ ; and  $N_i$  represent the number of women surveyed within area,  $A_i$ .

In this framework, the areal units and observation grid points are linked using the following equations

$$\text{logit}(p_i) = \tilde{x}'_i \beta + |A_i|^{-1} \int n(s) ds + \phi_i, i = 1, \dots, n_A$$

$$\text{logit}(p_i) = \tilde{x}'_i \beta + \eta(s_i) + \phi_{A_i}, i = n_A + 1, \dots, n_A + n_p$$

where  $x_i$  and  $\tilde{x}_i$  represent covariates for the  $i^{\text{th}}$  area and grid point, respectively. This framework provides a statistical link between the areal data and high-resolution spatial covariates and random effects, allowing for models at two spatial levels. Further details on the model framework are outlined in Utazi *et al.*<sup>2</sup> See refs 2 34 40 for more detailed information on similar approaches implementing an SPDE approach of Bayesian hierarchical models via R-INLA using DHS data.

### Patient and public involvement

Patients or the public were not involved in the design, conduct, reporting or dissemination of our research. The data used in these analyses were obtained from the DHS programme, which makes global health and demographic data confidentially and freely available to researchers across the world. More information on how the DHS programme conducts the Informed Consent process can be found at <https://dhsprogram.com>.

### RESULTS

Table 1 shows posterior marginal effects for the fixed effects within the 5 km, 50 km and 100 km models as well as model hyperparameters. Fixed effects estimates with upper and lower 95% credible intervals (CIs) that do not cross 1 are considered significant. Overall, we found that modelled c-section prevalence negatively correlated strongly with poverty and slightly with night-time lights across all spatial scales, as shown in online supplementary figure S1. Of note, night-time lights were not significant within the model but presented wide CIs for marginal effects, as seen in table 1. While poverty was not significant at the 5 km scale, it was significant at the 50 km and 100 km scales and showed a consistent pattern at the 5 km scale with other spatial resolutions (online supplementary figure S1). Conversely, these estimates were strongly positively associated with travel time to the nearest hospital across scales, although this was not significant within the model.

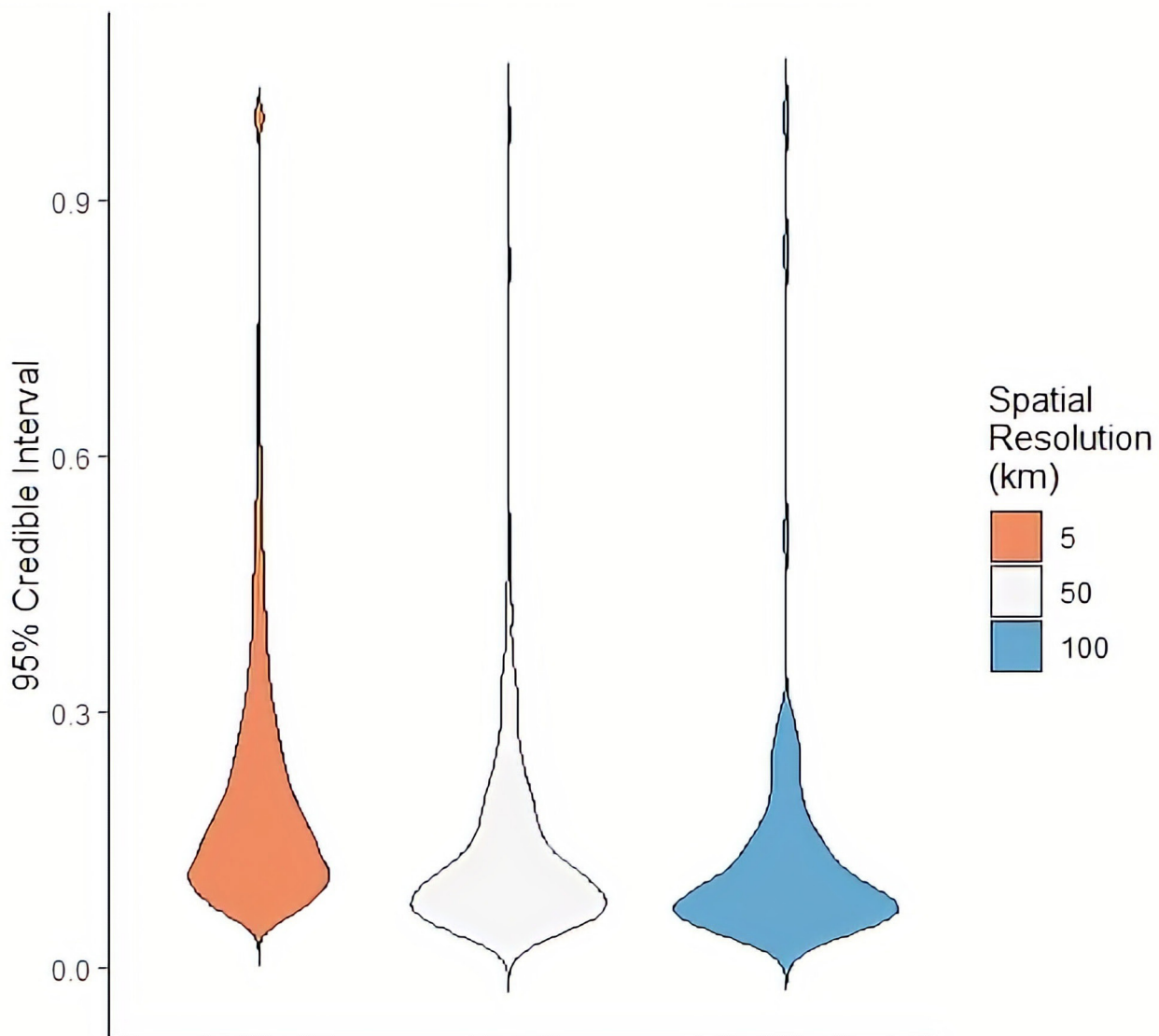
The large precision estimate for the 5 km model as shown in table 1 suggests the spatial process was estimated well with a Gaussian field, while smaller estimates among the 50 km and 100 km models suggest this was not the case. Regardless, the deviance information criterion (DIC) estimates for the latter two models were slightly improved over the 5 km model (table 1). Briefly, these DIC estimates represent a measure of model comparison, trading off between model complexity and model goodness of fit and performing well among Bayesian models in particular.<sup>41</sup> Smaller DIC values represent models with better fit, given model complexity, suggesting these models perform better as compared with the other models.

Figure 3 shows the distributions of the posterior 95% CIs for each model as violin plots. These plots show similar

**Table 1** Marginal effects of the fixed effects and hyperparameters of the posterior c-section models at 5 km, 50 km and 100 km

Parameter	5 km			50 km			100 km		
	Mean	Lower 95% CI	Upper 95% CI	Mean	Lower 95% CI	Upper 95% CI	Mean	Lower 95% CI	Upper 95% CI
Accessibility to cities	1.0001	0.9975	1.0027	0.9999	0.9975	1.0022	0.9999	0.9975	1.0022
Night-time lights	0.9603	0.0522	19.7189	0.666	0.0308	14.7623	0.6574	0.0298	14.9301
Live births	1.0471	0.8645	1.2873	0.9915	0.8313	1.1913	0.9938	0.8313	1.197
Poverty	0.0271	0.0005	2.1231	0.0071	0.0001	0.4548	0.0071	0.0001	0.4682
Travel to nearest hospital	1.0046	0.9946	1.0143	1.0062	0.9965	1.0163	1.0063	0.9964	1.0164
<b>Hyperparameters</b>									
$\theta_1$	Mean	SD	95% CI	Mean	SD	95% CI	Mean	SD	95% CI
	-0.9259	0.6329	(-2.1791 to 0.3104)	0.229	0.942	(-1.636 to 2.071)	0.213	0.945	(-1.646 to 2.070)
$\theta_2$	0.1348	0.4988	(-0.8380 to 1.1226)	0.041	0.724	(-1.376 to 1.472)	0.052	0.718	(-1.352 to 1.474)
Precision	1816.2045	1813.6598	(121.78 to 6626.24)	5.290	5.794	(0.647 to 20.476)	5.814	6.901	(0.648 to 23.539)
$\lambda$	0.5020	0.2699	(0.0497 to 0.9491)	0.354	0.243	(0.031 to 0.876)	0.354	0.243	(0.030 to 0.876)
DIC	155.67			152.4			152.61		
$P_D$	17.65			17.70			17.81		
Marginal likelihood	-118.52			-120.84			-120.89		

DIC, deviance information criterion.

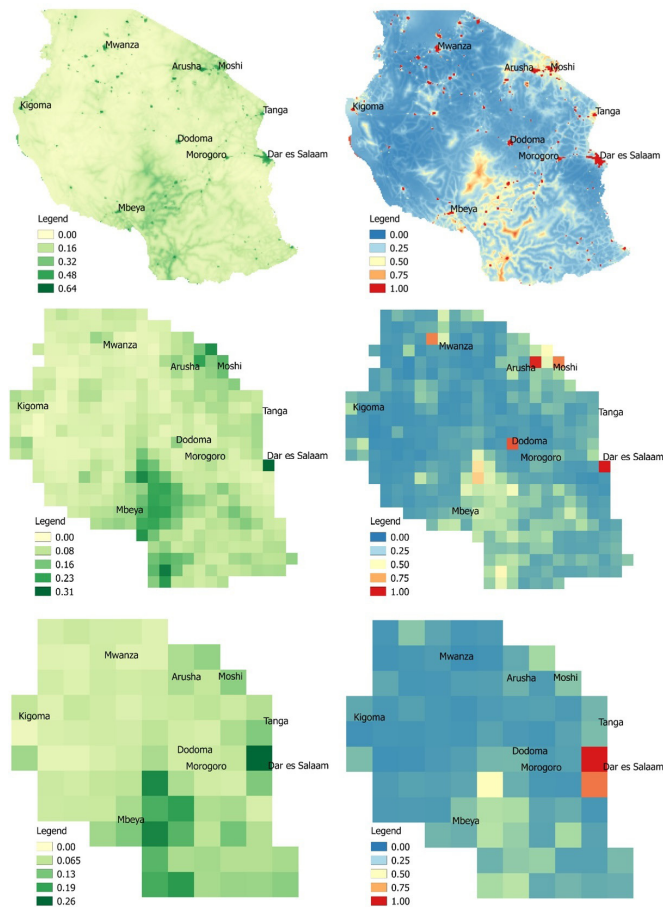


**Figure 3** Violin plot of posterior 95% credible intervals for caesarean section estimates predicted at the 5 km, 50 km and 100 km scale.

summary statistics as boxplots, while also providing information on probability densities, where thinner sections represent a lower probability of a given value occurring. These estimates approximate the trade-off between spatial resolution and uncertainty, representing the density of the width between the posterior upper and lower 95% CI for each grid cell for the 5 km, 50 km and 100 km surfaces. Here, all models had CI widths ranging from near 0 to 1, but CI width became more narrowly distributed and approached zero with decreasing spatial resolution. Mean density at the 100 km and 50 km scales were 0.13 ( $\pm 0.14$ ) and 0.14 ( $\pm 0.13$ ), respectively, while mean density at the 5 km scale was 0.21 ( $\pm 0.16$ ).

Figure 4 visualises prevalence of delivery via c-section and associated uncertainty at the 5 km, 50 km and 100 km spatial resolution. These maps show spatial patterns typical of c-section deliveries, with higher prevalence observed in cities such as Arusha and Dar es Salaam,

and lower prevalence in areas with high inaccessibility to a health facility or among more impoverished women (table 1). Overall, the mean estimated prevalence of obtaining a c-section at delivery at the 5 km resolution was 8.7% ( $\pm 6.2\%$ ), while the mean uncertainty as measured by the posterior distribution was 20.9% ( $\pm 15.7\%$ ). Mean estimated prevalence was slightly lower at the 50 km and 100 km resolutions, measuring 7.9% ( $\pm 4.9\%$ ) and 7.8% ( $\pm 4.8\%$ ), respectively, while mean uncertainty was 14.2% ( $\pm 12.7\%$ ) and 13.1% ( $\pm 14.3\%$ ). Areas of higher c-section utilisation were associated with higher uncertainty and observed around major urban areas, notably Dar Es Salaam, Arusha and Moshi and Dodoma. This trend was observed across spatial resolutions and in accordance with DHS data. Among DHS regions, Dar Es Salaam had the highest prevalence of delivery via c-section at 17%, compared with a national average of 5.9% (figure 1).



**Figure 4** Modelled c-section prevalence estimates (left) and associated 95% credible interval (right) at 5 km (top), 50 km (middle), and 100 km (bottom), Tanzania, 2015.

**DISCUSSION**

Overall, we found increasing model uncertainty associated with increasing spatial resolution, as quantified by increasing 95% CI widths (figure 3). This is unsurprising, given that increasing spatial resolution comes with problems of increasingly sparse events, zero inflation and missing data.<sup>42</sup> Furthermore, both c-section prevalence as well as model uncertainty tended to be higher in urban areas, reflecting greater variance within the data. This uncertainty could be due to increased data availability within large population centres, or potentially due to GPS displacement error within urban areas. While Bayesian hierarchical modelling techniques have developed to help account for these limitations through rigorous quantification of uncertainty,<sup>3</sup> our findings suggest an important and often overlooked trade-off persists between modelling these high-resolution health indicators and corresponding policy relevance if these estimates tend to be highly uncertain. This is evidenced in figure 4, where the highest rates of modelled c-section estimates tend to also have the highest corresponding uncertainty. Despite increasing uncertainty that accompanied increasing spatial resolution, we found that the 5 km model was the most precise, as evidenced by a high precision estimate (table 1). These findings correspond to findings reported in the Tanzania DHS report<sup>23</sup>

where rates of c-section utilisation are higher in urban areas. Other studies similarly suggest that while the prevalence of c-section is increasing globally,<sup>43 44</sup> women in more rural areas who cannot access a health facility quickly have a lower chance of undergoing the procedure in emergency circumstances.<sup>45 46</sup>

Within our models, we further found that poverty was negatively correlated with undergoing delivery via c-section across spatial resolutions and was significant at lower spatial resolutions (50 km and 100 km). These findings are again in line with reported DHS findings suggesting women in the highest wealth quintile were eight times more likely to undergo a c-section as compared with those in the lowest quintile.<sup>23</sup> Because health insurance coverage is generally low in Tanzania and relies heavily on payment at point of service,<sup>47</sup> our findings may suggest that more impoverished women are either unable to afford c-section surgeries when needed or are generally accessing healthcare less frequently across the continuum of pregnancy and childbirth.

Researchers are increasingly quantifying health and development indicators at the district level and high spatial resolutions, with aims of achieving Sustainable Development Goals to ensure ‘no one left behind’.<sup>8 48</sup> While identifying these previously hidden pockets of vulnerable and marginalised populations is vital to improving the health and well-being of all, the geostatistical methods used to accomplish these goals have inherent uncertainty and bias, which should be communicated effectively to policy makers and other non-academic audiences. While studies have recognised the importance in quantifying this uncertainty,<sup>3 22 42</sup> no studies have explored how increasing spatial resolution impacts model estimates and uncertainty. This study is therefore the first to map high resolution estimates of c-section prevalence in Tanzania and examine the trade-off between increasing spatial resolution and associated model uncertainty in these estimates. These results of this study imply an important trade-off between identifying concealed spatial heterogeneities and accuracy of estimates, which should be optimally communicated in policy relevant settings.

**Limitations**

This methodological study was exploratory in nature, examining the trade-off between increasing spatial resolution and model uncertainty and is therefore subject to a variety of limitations. First, we used a suite of standard covariates to explore the impact of spatial resolution alone on model uncertainty, and therefore rigorous covariate selection and model validation efforts were not undertaken during these analyses. As such, the results of these models may not be generalisable to Tanzania, nor to other study countries, and should be used for illustrative purposes only. Furthermore, the spatial resolutions chosen for these analyses represent a theoretical exploration of the impact of increasing resolution on modelled estimates and are unlikely to represent estimates at resolutions at which policy decisions are made. Future work may explore the impact of increasing spatial

resolution on modelled estimates at policy relevant administrative boundaries, as opposed to the rasters employed in these analyses. Second, the suite of covariates and DHS data used are subject to their own biases and limitations—for example, the DHS collects data on births within the previous 5 years, so the estimates presented here may not reflect the current situation within Tanzania. Furthermore, DHS data are not routinely collected registration data and therefore do not capture information on c-sections performed on women who have subsequently died. This potentially represents biased information, as only women obtaining and surviving the procedure are interviewed. Additionally, travel to the nearest public hospital does not account for individuals who may bypass the nearest facility in favour of a facility with higher quality of care. Lastly, the geospatial covariates used have associated error and misclassification bias, particularly night-time lights that may suffer from light refraction errors, for example. Future work should aim to include more recent data on actual health facility used, where possible, and explore the impact of misclassification bias inherent to these environmental covariates.

## CONCLUSIONS

Researchers are increasingly applying Bayesian hierarchical modelling techniques to visualise high-resolution spatial patterns of health indicators. These techniques offer powerful and rigorous methods to quantify and visualise model uncertainty, but few studies have explored how to communicate this uncertainty in policy-relevant settings. Here, we explored how model uncertainty changes with increasing spatial resolution and found that while uncertainty increases with higher spatial resolutions, model precision was best approximated at the highest spatial resolution, suggesting an important policy trade-off between identifying concealed spatial heterogeneities in health indicators. In modelling delivery via c-section at multiple spatial resolutions, we demonstrate poverty to be negatively correlated across spatial resolutions, suggesting important disparities in obtaining life-saving obstetric surgery persist across sociodemographic factors. This work is the first study to explore modelled c-section estimates and uncertainty at varying spatial resolutions and has potential policy implications in terms of visualising spatial patterns of obstetric surgery, as well as focusing MNH data collection efforts within Tanzania.

**Acknowledgements** This work forms part of the output of WorldPop ([www.worldpop.org](http://www.worldpop.org)) at the University of Southampton.

**Contributors** CWR conceived of the study framework and methods for this study, performed all data management and analysis, generated all study visualisations and results interpretations, and wrote and prepared the manuscript for journal submission. JJN and NR contributed insight and feedback on the conceptual study design, as well as statistical feedback on analysis and model results. KN and JES provided input on how the study results fit into the broader fields of maternal health and geography, respectively. Lastly, ZM and AJT supervised CWR during production of this manuscript and oversaw that methods and analysis were executed within a sound and scientifically rigorous framework. All authors reviewed the manuscript and provided feedback before final submission to the journal.

**Funding** The authors would like to acknowledge the support provided by the Economic and Social Research Council's Doctoral Training Programme, which funds CWR.

**Map disclaimer** The depiction of boundaries on the map(s) in this article do not imply the expression of any opinion whatsoever on the part of BMJ (or any member of its group) concerning the legal status of any country, territory, jurisdiction or area or of its authorities. The map(s) are provided without any warranty of any kind, either express or implied.

**Competing interests** None declared.

**Patient consent for publication** Not required.

**Provenance and peer review** Not commissioned; externally peer reviewed.

**Data availability statement** Data are available in a public, open access repository. All data used in these analyses were freely obtained from the DHS program ([dhsprogram.com](http://dhsprogram.com)). The results of these analyses will be freely available via [worldpop.org](http://worldpop.org).

**Open access** This is an open access article distributed in accordance with the Creative Commons Attribution 4.0 Unported (CC BY 4.0) license, which permits others to copy, redistribute, remix, transform and build upon this work for any purpose, provided the original work is properly cited, a link to the licence is given, and indication of whether changes were made. See: <https://creativecommons.org/licenses/by/4.0/>.

## REFERENCES

- 1 Utazi CE, Thorley J, Alegana VA, *et al*. High resolution age-structured mapping of childhood vaccination coverage in low and middle income countries. *Vaccine* 2018;36:1583–91.
- 2 Utazi CE, Thorley J, Alegana VA, *et al*. A spatial regression model for the disaggregation of areal unit based data to high-resolution grids with application to vaccination coverage mapping. *Stat Methods Med Res* 2019;28:3226–3241.
- 3 Gething P, Tatem A, Bird T, *et al*. *Creating spatial interpolation surfaces with DHS data*. Rockville, Maryland, USA: ICF International, 2015. <http://dhsprogram.com/pubs/pdf/SAR11/SAR11.pdf>
- 4 Blangiardo M, Cameletti M. *Spatial and Spatio-temporal Bayesian Models with R - INLA*. John Wiley & Sons, 2015. <https://www.wiley.com/en-gb/Spatial+and+Spatio+temporal+Bayesian+Models+with+R+INLA-p-9781118326558>
- 5 Ruktanonchai CW, Ruktanonchai NW, Nove A, *et al*. Equality in maternal and newborn health: modelling geographic disparities in utilisation of care in five East African countries. *PLoS One* 2016;11:e0162006.
- 6 Ruktanonchai CW, Nilsen K, Alegana VA, *et al*. Temporal trends in spatial inequalities of maternal and newborn health services among four East African countries, 1999–2015. *BMC Public Health* 2018;18:1339.
- 7 Pezzulo C, Bird T, Utazi EC, *et al*. *Geospatial modeling of child mortality across 27 countries in sub-Saharan Africa*. Rockville, MD, USA: ICF International, 2016. <https://eprints.soton.ac.uk/400944/>
- 8 WHO. *State of inequality: reproductive, maternal, newborn and child health*. Geneva, Switzerland: World Health Organization, 2015. [http://apps.who.int/iris/bitstream/10665/164590/1/9789241564908\\_eng.pdf?ua=1&ua=1](http://apps.who.int/iris/bitstream/10665/164590/1/9789241564908_eng.pdf?ua=1&ua=1)
- 9 Utazi CE, Sahu SK, Atkinson PM, *et al*. Geographic coverage of demographic surveillance systems for characterising the drivers of childhood mortality in sub-Saharan Africa. *BMJ Glob Health* 2018;3:e000611.
- 10 Ebener S, Guerra-Arias M, Campbell J, *et al*. The geography of maternal and newborn health: the state of the art. *Int J Health Geogr* 2015;14:19.
- 11 Ahmed S, Li Q, Scraftford C, *et al*. *An assessment of DHS maternal mortality data and estimates. DHS methodological report*. Rockville, Maryland, USA: ICF International, 2014. <https://dhsprogram.com/pubs/pdf/MR13/MR13.pdf>
- 12 Say L, Chou D, Gemmill A, *et al*. Global causes of maternal death: a who systematic analysis. *Lancet Glob Health* 2014;2:e323–33.
- 13 Roberts DA, Ng M, Ikilezi G, *et al*. Benchmarking health system performance across regions in Uganda: a systematic analysis of levels and trends in key maternal and child health interventions, 1990–2011. *BMC Med* 2015;13:285.
- 14 Bhattacharjee NV, Schaeffer LE, Marczak LB, *et al*. Mapping exclusive breastfeeding in Africa between 2000 and 2017. *Nat Med* 2019;25:1205–12.
- 15 Mosser JF, Gagne-Maynard W, Rao PC, *et al*. Mapping diphtheria-pertussis-tetanus vaccine coverage in Africa, 2000–2016: a spatial and temporal modelling study. *The Lancet* 2019;393:1843–55.

- 16 Patil AP, Gething PW, Piel FB, *et al.* Bayesian geostatistics in health cartography: the perspective of malaria. *Trends Parasitol* 2011;27:246–53.
- 17 Soares Magalhães RJ, Clements ACA. Mapping the risk of anaemia in preschool-age children: the contribution of malnutrition, malaria, and helminth infections in West Africa. *PLoS Med* 2011;8:e1000438.
- 18 Schur N, Hürlimann E, Garba A, *et al.* Geostatistical model-based estimates of schistosomiasis prevalence among individuals aged ≤20 years in West Africa. *PLoS Negl Trop Dis* 2011;5:e1194.
- 19 Tatem AJ, Adamo S, Bharti N, *et al.* Mapping populations at risk: improving spatial demographic data for infectious disease modeling and metric derivation. *Popul Health Metr* 2012;10:8.
- 20 Kirby RS, Delmelle E, Eberth JM. Advances in spatial epidemiology and geographic information systems. *Ann Epidemiol* 2017;27:1–9.
- 21 Clements ACA, Lwambo NJS, Blair L, *et al.* Bayesian spatial analysis and disease mapping: tools to enhance planning and implementation of a schistosomiasis control programme in Tanzania. *Trop Med Int Health* 2006;11:490–503.
- 22 Goovaerts P. Geostatistical analysis of disease data: visualization and propagation of spatial uncertainty in cancer mortality risk using Poisson kriging and p-field simulation. *Int J Health Geogr* 2006;5:7.
- 23 Ministry of Health CD, Ministry of Health - MoH/Zanzibar, National Bureau of Statistics - NBS/Tanzania, Office of Chief Government Statistician - OCGS/Zanzibar, ICF. Tanzania demographic and health survey and malaria indicator survey 2015–2016, 2016. Available: <http://dhsprogram.com/pubs/pdf/FR321/FR321.pdf>
- 24 SAS Institute Inc. SAS version 9.4. Cary, NC, USA: SAS Institute Inc. 2013.
- 25 Burgert CR, Colston J, Roy T, *et al.* *Geographic displacement procedure and GEOREFERENCED data release policy for the demographic and health surveys*. Calverton, Maryland, USA: ICF International, 2013. <https://dhsprogram.com/pubs/pdf/SAR7/SAR7.pdf>
- 26 Nelson A. Estimated travel time to the nearest city of 50,000 or more people in year 2000, 2008. Available: <http://forobs.jrc.ec.europa.eu/products/gam/>
- 27 US NOAA National Geophysical Data Center, US Air Force Weather Agency. Version 4 DMSP-OLS Nighttime Lights Time Series (1992–2013; Average Visible, Stable Lights, & Cloud Free Coverages), 2014. Available: <http://ngdc.noaa.gov/eog/dmsp/downloadV4composites.html> [Accessed 22 Aug 2018].
- 28 Tatem AJ, Campbell J, Guerra-Arias M, *et al.* Mapping for maternal and newborn health: the distributions of women of childbearing age, pregnancies and births. *Int J Health Geogr* 2014;13:2.
- 29 Tatem A, Gething P, Bhatt S, *et al.* Pilot high resolution poverty maps, 2013. Available: <http://www.worldpop.org.uk/data/summary/?doi=10.5258/SOTON/WP00290>
- 30 Ouma PO, Maina J, Thurairana PN, *et al.* Access to emergency hospital care provided by the public sector in sub-Saharan Africa in 2015: a geocoded inventory and spatial analysis. *Lancet Glob Health* 2018;6:e342–50.
- 31 Gabrysch S, Campbell OMR. Still too far to walk: literature review of the determinants of delivery service use. *BMC Pregnancy Childbirth* 2009;9:34.
- 32 Gething PW, Johnson FA, Frempong-Ainguah F, *et al.* Geographical access to care at birth in Ghana: a barrier to safe motherhood. *BMJ Public Health* 2012;12:991.
- 33 Mpembeni RN, Killewo JZ, Leshabari MT, *et al.* Use pattern of maternal health services and determinants of skilled care during delivery in southern Tanzania: implications for achievement of MDG-5 targets. *BMC Pregnancy Childbirth* 2007;7:29.
- 34 Bosco C, Alegana V, Bird T, *et al.* Exploring the high-resolution mapping of gender-disaggregated development indicators. *J R Soc Interface* 2017;14:20160825.
- 35 Bearak J, Popinchalk A, Alkema L, *et al.* Global, regional, and subregional trends in unintended pregnancy and its outcomes from 1990 to 2014: estimates from a Bayesian hierarchical model. *Lancet Glob Health* 2018;6:e380–9.
- 36 Kazembe LN, Namangale JJ. A Bayesian multinomial model to analyse spatial patterns of childhood co-morbidity in Malawi. *Eur J Epidemiol* 2007;22:545–56.
- 37 Neal S, Ruktanonchai CW, Chandra-Mouli V, *et al.* Using geospatial modelling to estimate the prevalence of adolescent first births in Nepal. *BMJ Glob Health* 2019;4:e000763.
- 38 Hug L, Alexander M, You D, *et al.* National, regional, and global levels and trends in neonatal mortality between 1990 and 2017, with scenario-based projections to 2030: a systematic analysis. *The Lancet Global Health* 2019;7:e710–20.
- 39 Rue H, Martino S, Chopin N. Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations. *J R Stat Soc Ser B Stat Methodol* 2009;71:319–92.
- 40 Krainski ET, Castro-camilo D. *Advanced spatial modeling with stochastic partial differential equations using R and INLA*. CRC Press, Taylor & Francis Group, 2018.
- 41 Spiegelhalter DJ, Best NG, Carlin BP, *et al.* Bayesian measures of model complexity and fit. *J Royal Statistical Soc B* 2002;64:583–639.
- 42 Richardson S, Thomson A, Best N, *et al.* Interpreting posterior relative risk estimates in disease-mapping studies. *Environ Health Perspect* 2004;112:1016–25.
- 43 Vogel JP, Betrán AP, Vindeoghel N, *et al.* Use of the Robson classification to assess caesarean section trends in 21 countries: a secondary analysis of two who multicountry surveys. *Lancet Glob Health* 2015;3:e260–70.
- 44 Betrán AP, Merialdi M, Lauer JA, *et al.* Rates of caesarean section: analysis of global, regional and national estimates. *Paediatr Perinat Epidemiol* 2007;21:98–113.
- 45 Mishra US, Ramanathan M. Delivery-related complications and determinants of caesarean section rates in India. *Health Policy Plan* 2002;17:90–8.
- 46 Neuman M, Alcock G, Azad K, *et al.* Prevalence and determinants of caesarean section in private and public health facilities in underserved South Asian communities: cross-sectional analysis of data from Bangladesh, India and Nepal. *BMJ Open* 2014;4:e005982.
- 47 Mtei G, Makawia S, Ally M, *et al.* Who pays and who benefits from health care? an assessment of equity in health care financing and benefit distribution in Tanzania. *Health Policy Plan* 2012;27 Suppl 1:i23–34.
- 48 United Nations. The sustainable development goals report 2018. New York, NY, USA United Nations; 2018. <https://unstats.un.org/sdgs/files/report/2018/TheSustainableDevelopmentGoalsReport2018-EN.pdf>