Original Research Article

'Spread' of Artemisinin Resistance in Myanmar Revisited

Abstract

Geostatistical models have been widely used to represent disease prevalence in spatial disease mapping. More recently, these models have been employed to estimate geographic 'spread' of antimalarial drug resistance in South East Asia from genetic mutations identified in the K13 gene of the malaria *plasmodium falciparum* parasite. Here, I revisit the question of 'spread' of artemisinin resistance (AR) as represented by resistant *plasmodium falciparum* in Myanmar, when re-evaluated from K13 mutant alleles data published by Tun et al (2015). The new analysis gives a broader perspective by incorporating information published by the World Health Organization (WHO) in 2015 and subsequently in 2018 of the K13 mutant alleles confirmed to confer or to be associated with artemisinin resistance. This provides insights, which hitherto have not been described and reveals the disparity between the estimation of 'spread' of AR by Tun et al (2015) and that of AR prevalence based on metrics, which (are) currently supportable) are supported by published WHO's data.

Key words: K13, genetic markers, overestimation, prevalence, malaria, antimalarial drug resistance.

1. Introduction

Malaria is caused by a microscopic *Plasmodium* parasite conveyed into humans by some of the mosquito species in the *Anopheles* genus. Antimalarial drugs have long been developed to combat the disease, but their effectiveness, typically has been shown to decrease with time as antimalarial drug resistance evolves in certain (a) regional populations. More recently, there has been an emphasis on geostatistical maps to describe the regional status of antimalarial drug resistance in epidemiological mapping studies (e.g. Grist et al., 2016) and (Ménard et al., 2016; Tun et al., 2015; Zaw et al., 2018). Such mappings attempt to depict the prevalence of antimalarial drug resistance in a geographical region using geostatistical models parameterized with data taken from infected persons, but (these data) are typically limited in sample size and spatial distribution. Their accuracy is dependent also upon the choice of metrics used to denote antimalarial drug-resistance as well as what constitutes statistical significance for the gathered data. For instance, relatively recent antimalarial drugs such as artemisinin, the uncertainty surrounding all these factors remains largely unresolved. This leaves assessments of antimalarial drug resistance and interpretation of its 'spread' open to overestimation and sensationalization (sm) as a scaremongering tactic (Chookajorn, 2018; Meshnick, 2012; Tun et al., 2015), which could arguably lead to a conflict of interest in disease-mitigation strategies.

The efficacy of an antimalarial drug is typically evaluated by recording the time taken for an arbitrary fixed percentage (usually 50%) of malaria parasites to be cleared from an infected person after administering the drug. If the parasite clearance time, when measured over a large number of infected individuals in a given geographical region, is statistically slower than a specific, yet arbitrary average, then antimalarial

(drug) resistance to the drug in use is deemed to be present in the region. Although the assumption of delayed parasite clearance times equating with the presence of artemisinin resistant *Plasmodium falciparum* parasites and hence artemisinin resistance (AR) is generally accepted, this has been questioned by some leading authors (Ferreira et al., 2013; Krishna & Kremsner, 2013). In line with Tun et al. (2015) however, the assumption is accepted as axiomatic in this paper.

In the case of artemisinin, concerns of artemisinin-resistant *Plasmodium falciparum* spreading across parts of South-East Asia, and in particular Myanmar, have been loudly broadcast (Tun et al., 2015). The thesis is assisted by the discovery of a genetic association between specific mutations of the Kelch 13 (K13) gene in the *Plasmodium* parasites with longer malarial parasite clearance times. But, it is also important to recognise that parasite clearance delay times vary both with different K13 codon mutations and sampling location. For example, the artemisinin resistance (AR) 'effect' of codon F446I has been measured as mild in northern Myanmar where this mutation predominates (Tun et al 2015, White and Woodrow, 2017). In short, the precise causal reasons for the genetic associations and variations in their effect against artemisinin resistance remain unknown. Nevertheless, identification of specific mutant alleles, which confer resistance to artemisinin has been the subject of much ongoing research since the discovery was published (Ariey et al., 2014.)

In 2015, the World Health Organisation (WHO) published a list of K13 mutant allele codon numbers, which at the time were either *confirmed to confer* AR or to be *candidates associated* with AR (WHO 2015), and this was later updated to reflect the current knowledge in (2018 in) a list published (in) by WHO (2018). Such definitive genetic marker information provides a statistical handle, enabling AR to be more rapidly estimated through inference from blood samples taken from infected persons. It is the association of AR resistant *Plasmodium falciparum,* with these K13 gene mutations, when observed in spatial locations where infected persons are present that provides the marker tool for mapping estimated AR prevalence on a geographical basis.

In Tun et al. (2015) however, the set of *all* mutant alleles with a codon number greater than 440 (henceforth referred to as the g440 metric) collected in their sampling, was assumed to confer AR. Surprisingly, no supporting scientific data were provided for this assumption and it appears to have been an arbitrary choice based only on codons located in the propeller region of the K13 protein. The g440 metric, was then employed in geostatistical models to generate maps of the estimated 'spread' of artemisinin resistant plasmodium from K13 gene mutation from blood sampled data obtained at sites (sampled) across Myanmar. Again, neither the robustness nor the importance of the underlying assumptions made (that all the mutant alleles greater than codon 440 confer AR) was addressed in producing the (se) maps in question. From a scientific standpoint, all these factors are crucial, given that not all these codons necessarily confer AR and that many, such as A578S, are known (to) not to confer AR (Ménard et al., 2016; Woodrow & White, 2017) . Fairhurst (2016) for example, provides a list of mutant alleles known to (then) confer

AR, which included only 5 of the mutant alleles employed in the g440 metric of Tun et al (2015).

This raises the important, but unanswered, question of how the choice of g440 metric may have influenced the inference and geostatistical mapping of estimated 'spread' of AR in Myanmar by Tun et al (2015). The purpose of this paper, is therefore to determine how the choice of metric for K13 mutant allele codons conferring AR from the K13 genetic marker data as published by Tun et al (2015) in Myanmar:

i. influences the estimation of AR prevalence,
ii. compares with estimates based on K13 mutant alleles codons that have been recognised as conferring AR by the WHO.

2. Materials and Methods

2.1 *Sample sizes*

A first avenue into the strength of the statistical inference, possible from the Tun et al (2015) data, is provided by examining the sample sizes at each of the sampled sites. Figure 1 shows a plot of the sample size distribution when viewed across all samples taken from all the study sites.

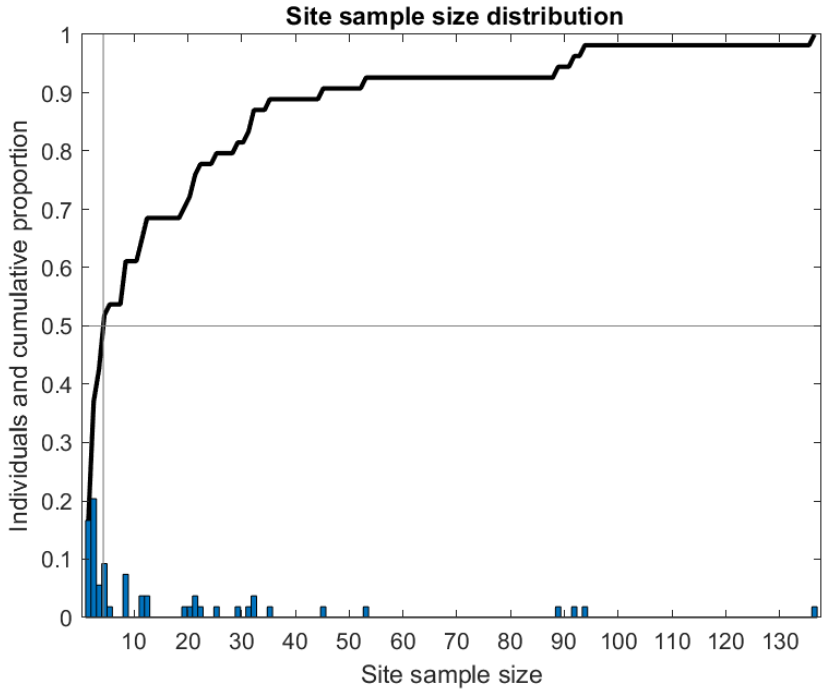

Figure 1. Site sample sizes of the Tun et al (2015) data. The blue bars show relative frequencies of site sample sizes as a proportion of the total samples collected (n=940) with the cumulative proportion superimposed (solid black line). The 50% median percentile (vertical thin line) demonstrates that the majority (28) of the 54 sites sampled had fewer than 5 individuals.

This plot immediately reveals the sparsity of larger samples obtained at any specific location and that the vast majority (28, greater than 50%) of the samples were

obtained from site samples of less than 5 individuals. The spatial extent of the corresponding information supporting any thesis of 'spread' of AR is thus severely limited. In Tun et al (2015), this limitation is concealed by appealing to plots of local prevalence based on agglomerated data shown only for the local administrative regions (in their Figure 3).

2.2 Choice of AR metric

In Table 1, four different AR metrics named Metric 1 to Metric 4 are defined in terms of the K13 mutant alleles of the codon numbers implicated in AR based on the official lists of 2015 and 2018 published by the WHO (WHO 2015, 2018).

WHO 2015		**Myanmar**		**WHO 2018**		**Myanmar**	
confirmed	**candidate**	**Metric 1**	**Metric 2**	**confirmed**	**candidate**	**Metric 3**	**Metric 4**
n=4	*n=9*	*n=1*	*n=8*	*n=9*	*n=11*	*n=5*	*n=14*
493	**441**	**580**	**441**	**446**	**441**	**446**	**441**
539	**446**		**446**	**458**	**449**	**458**	**446**
543	**449**		**449**	**476**	**469**	**476**	**449**
580	*553*		**458**	*493*	**481**	**561**	**458**
	458		**561**	*539*	**527**	**580**	**469**
	561		**574**	*543*	**537**		**476**
	568		**580**	*553*	**538**		**481**
	574		**675**	**561**	*568*		**527**
	675			**580**	**574**		**537**
					673		**538**
					675		**561**
							574
							580
							675
P441L , **F446I** , **G449A** , **N458Y** , **C469F** , **M476I**, **A481V** , *Y493H* , **P527H** , **N537I** , **G538V** , *R539T*, *I543T* , *P553L*, **R561H** , *V568G* **, P574L** , **C580Y** , *F673I* , **A675V**							

Table 1. **LABELLING OF TABLES SHOULD BE ABOVE THEM**

K13 mutant allele codons either *confirmed to confer AR* or identified as *candidates to confer AR* by the WHO in the official lists at 2015 (WHO 2015) and 2018 (WHO 2018). Those shown in bold type are included in the g440 metric of mutant allele codons found and used to estimate AR 'spread' in Myanmar by Tun et al (2015). All codons are listed in full notation (with prefix and suffix letters) in the bottom panel. Those codons employed in Metrics 1 to 4, as defined in this paper, to estimate AR prevalence in Myanmar are listed in the adjacent 4 separate columns.

These codons fall into two categories within each study year, respectively defined as either *confirmed to confer* AR, or *candidates to confer* AR at the time of the WHO publications in 2015 and 2018. These findings, therefore, provide 4 logical options for defining an AR metric in terms of codon subgroups:

Metric 1. *Confirmed codons* to confer AR in 2015.

Metric 2. *Confirmed and candidate* codons to confer AR in 2015.
Metric 3. *Confirmed codons* to confer AR in 2018.
Metric 4. *Confirmed and candidate* codons to confer AR in 2018.

These metrics provide a straightforward scientific basis for assessing the estimation of AR prevalence by incorporating official information on K13 mutant alleles codons as published by the WHO in 2015 and in 2018. Metric 1 and Metric 2 respectively correspond to the best- and worst-case scenarios based on the WHO (2015) list. Metric 1 represents the minimum number of codons at 2015 confirmed to confer AR, while Metric 2 represents the potential perceived as maximum number of codons that could confer AR. Metric 3 and Metric 4 respectively correspond to parallel definitions as at 2018 published data based on the WHO (2018) list.

Metric 1 consists of the sole codon C580Y, which was the only codon in the Tun et al (2015) data as confirmed to confer AR (at) in 2015 by the WHO. A comparison of the codons listed for Metric 2 (8 in total), with those listed for Metric 3 (5 in total), importantly shows that three of the Metric 2 candidates in 2015 codons (P441L, P574L, A675V) were still candidates and were not subsequently confirmed to confer AR in 2018 confirmed data. One other WHO 2015 listed candidate, codon *V568G*, and not found in the Myanmar data, also remained a candidate in the list of 2018 published data by WHO. If this process of incomplete progression (from candidate to confirmed data) continues to the next updated WHO list, yet to be published, it will likely imply that Metric 4 data with inclusion of all perceived candidate codons also provides an *overestimation* of the AR prevalence in Myanmar.

Metric 3 therefore, provides a current benchmark as the minimum subset of 5 mutant allele codons found by Tun et al (2015) as currently confirmed to confer AR by the WHO. Metric 4 gives an upper bound on Metric 3 as a maximal subset (14 in total) to confer AR as currently reported on the assumption that the 9 candidates listed in WHO (2018) will subsequently be confirmed to confer AR. However, the estimate of AR 'spread' by Tun et al (2015), when using the g440 metric, namely with all 24 mutant alleles with codon number greater than 440, still exceeds this upper bound by including an additional 10 codons without support and never included in the WHO 2018 list which was published 3 years later. LONG SENTENCE – CONSIDER BREAKING INTO SHORT ONES FOR CLARITY. In simple terms therefore, this presentation seems to expose the extent of the overestimation of AR prevalence more conspicuous as (consciously) made by Tun et al. (2015) (of AR prevalence) in Myanmar's case at the time of that publication.

2.3 Estimation of AR spread

In estimating 'spread' of AR, it is critical to define exactly what is meant by 'spread'. Woodrow and White (2017) allude to this by differentiating between independent 'de novo' geographic emergences of AR (Takala-Harrison et al., 2015) in the sense of a 'soft sweep' of competing mutants arising and 'spread' as a 'hard sweep' in which with time, many of the soft sweep mutants decline to a few in particular. In Tun et al (2015), the usage of the word 'spread' of AR also comes with a dynamic innuendo of

spreading spatially in time (also pointed out by Plowe and Ringwald (2015)), though this obviously cannot be supported as the data are effectively limited to a snapshot and, therefore, unlike a disease time series (e.g. Kaestli et al., 2016), lack a temporal component.

In the current paper, I use the term 'spread' of AR in the spatial sense and distinguish this from 'spread' of AR in the population sense. In the former sense, spread is determined by a spatial measure, most fundamentally as the number of sampled sites where mutant alleles were found, which I refer to as spatial sites coverage. In the latter sense, spread is determined by the frequency of mutant alleles found in the sampled individuals from the study population.

Spatial sites coverage provides a crude summary statistic for assessing the spread of AR in terms of the proportion of sampled sites where cases of AR (by a chosen metric) were found. However, (this) such statistical method takes no account of the number of individuals sampled, other than (that) one or more (were) likely to be found (to be) presenting with an AR mutant allele at each site. Thus, if only a few individuals with mutant alleles were found with AR genes but were located over several closely located sites, AR spread would have received a higher estimate than if many individuals were found at a few but widely dispersed sites.

On the other hand, population spread as the summary statistic for assessing spread of AR, namely as the proportion of sampled individuals where cases of AR (by a chosen metric) were found, (is) could also be a crude estimate for AR spread because it could take no account of any spatial information. Population spread accuracy in determining geographic prevalence is dependent only on whatever spatial diversity is (unwittingly) inherent but not explicitly recorded in the configuration of site sampling locations.

When contrasted with both the above crude metrics of spread, a geostatistical model could enable a more sophisticated estimate of spread of AR to be derived by taking into account the number of cases of AR (whatever the chosen metric), and both the spatial site sampling locations and number of individuals sampled at each study site. This is achieved by spatial interpolation across the geographic domain at points where no sampling occurred, based on data, statistical modelling assumptions and knowledge. In addition, geostatistical maps can easily enable a visualisation of the spatial extent of spread of AR to be readily compared using different metrics of interest for the given data. In this paper, I determine geostatistical maps of AR prevalence in Myanmar through the widely used Bayesian hierarchical model approach (Lawson, 2008). The models incorporate either the commonly used Poisson (e.g. Walker and Gotway, 2004) or binomial (e.g. Morgan et al 2004) underlying distributional assumptions together with a Besag-York-Mollie (BYM) specification and were evaluated using the R-INLA software (Appendix).

3. Results

Figure 2 compares the distributions (for) of AR spread obtained by the two definitions of population proportions and spatial sites coverage in percentage, using

the g440 AR metric of Tun et al (2015). A comparison of the bar plots shown for spread in the spatial sites' coverage sense (Figure 2B) with those (for) of spread in the population sense (Figure 2A) demonstrates that the estimate of AR spread is more accentuated when evaluated by spatial sites coverage.

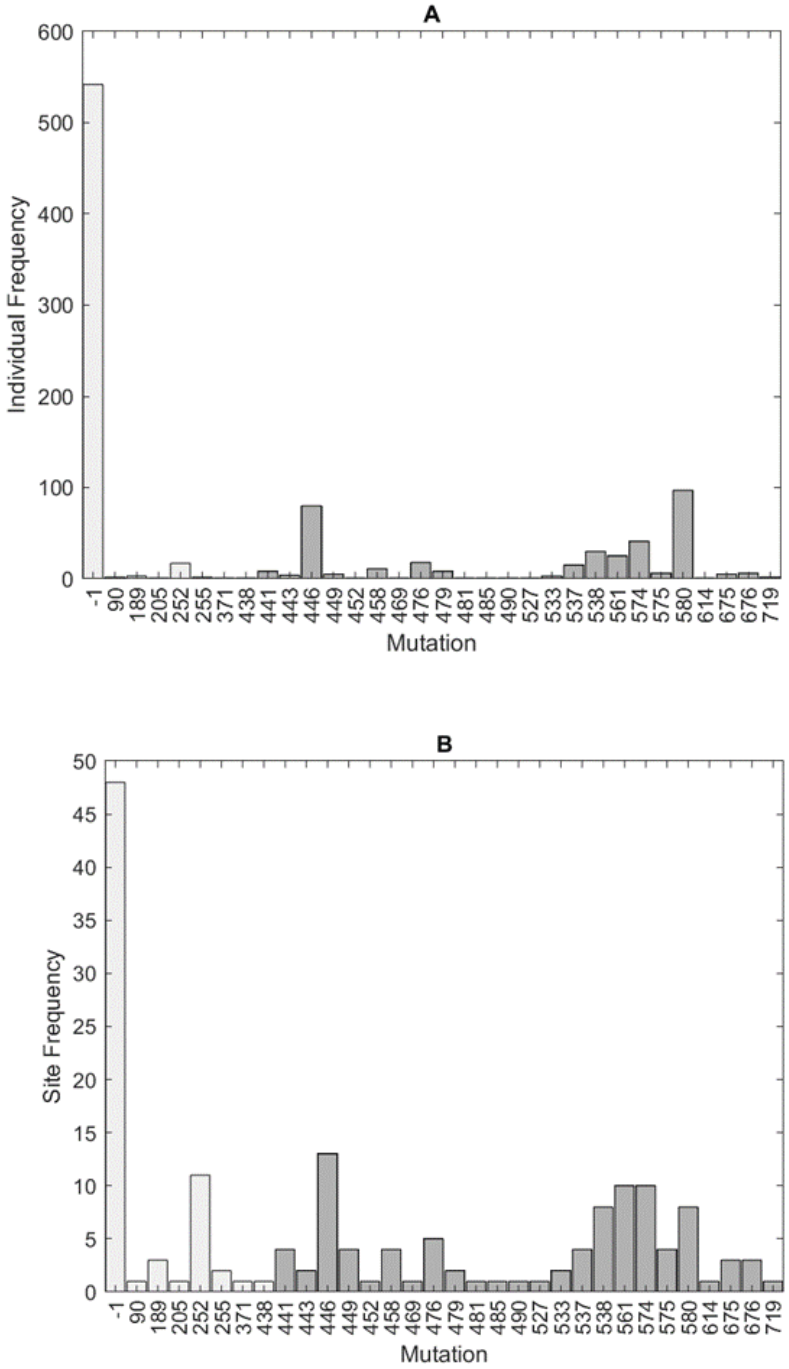

Figure 2. Bar plots showing the codon number frequencies of K13 mutant alleles in the Tun et al (2015) data: (A) by number of individuals (sample total *n*=940) and (B) by number of sites (sample total *n*=54). The shaded bars indicate those codon numbers included in the g440 metric (those codon numbers > 440) which was used in the estimation of AR 'spread' by Tun et al (2015).

Figure 3 shows geostatistical maps of AR prevalence in Myanmar using the g440 AR metric of Tun et al (2015) data, generated by R-INLA software (Lindgren & Rue, 2015) with a BYM specification (Besag et al., 1991) and either a Poisson or binomial underlying distribution (Appendix?). Results for the Poisson model are shown in the top row, with Figure 3A showing the map of estimated AR prevalence and Figure 3B showing the associated, but uncertainty map. The results for the binomial model are shown in the bottom row with Figure 3C showing the prevalence map and Figure 3D depicting the uncertainty map. The maps in Figure 3(A and C) are strongly reminiscent, as would be expected, of those maps generated by the geostatistical Bayesian models used for estimating AR prevalence in Tun et al (2015) data.

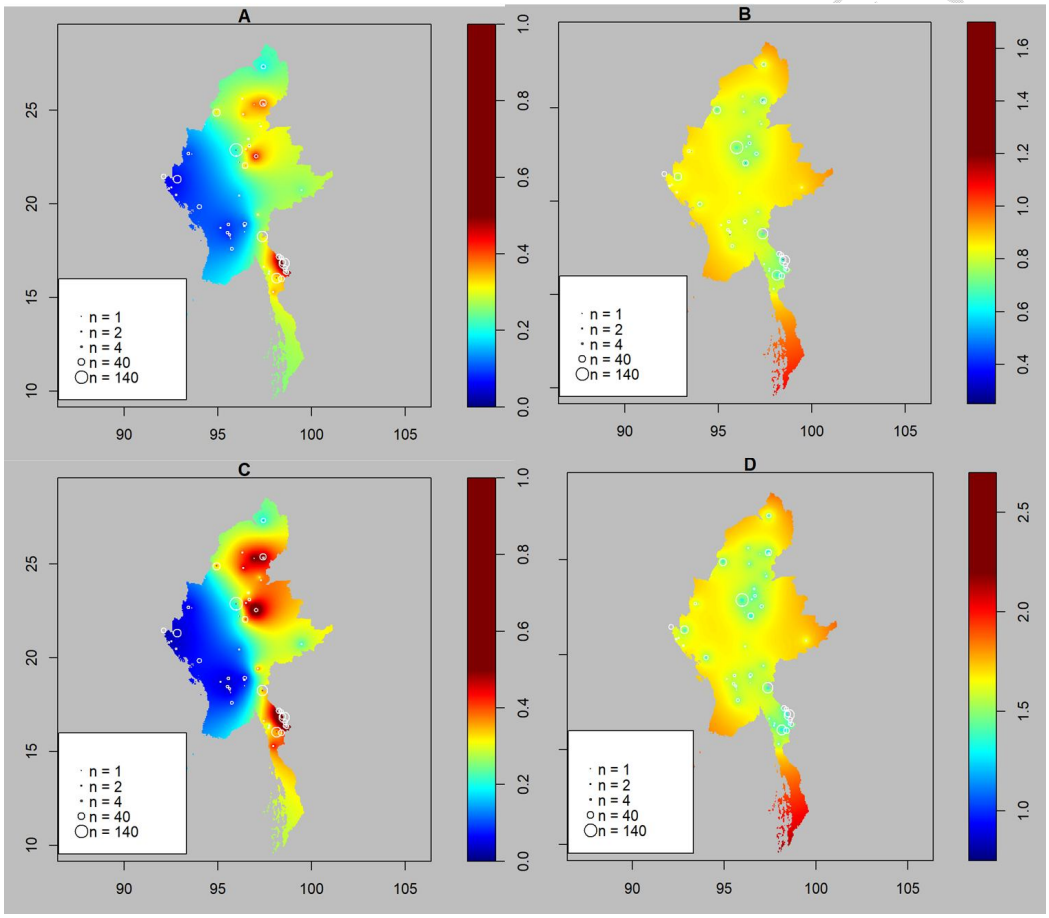

Figure 3. Geostatistical maps of AR 'spread' in Myanmar generated by R-INLA with the BYM model specification using the g440 metric of Tun et al (2015) data for mutant allele codons with an underlying Poisson distribution: (A) prevalence and (B) associated uncertainty and with an underlying binomial distribution: (C) prevalence and (D) associated uncertainty. The colour scale in (A) and (C) ranges from dark blue (0%) to dark red (>=50%) as was scaled in Tun et al (2015). Site sample sizes are superimposed as circles.

A comparison of estimated AR prevalence in the Figure 3 maps, obtained by using the g440 metric, was made with estimated AR prevalence by each of the four WHO's metric data. The respective RINLA maps as here generated, are shown in Figure 4 for Metric 1 and Metric 2 based on the WHO (2015) list of codons and in Figure 5 for Metric 3 and Metric 4, based on the WHO (2018) list.

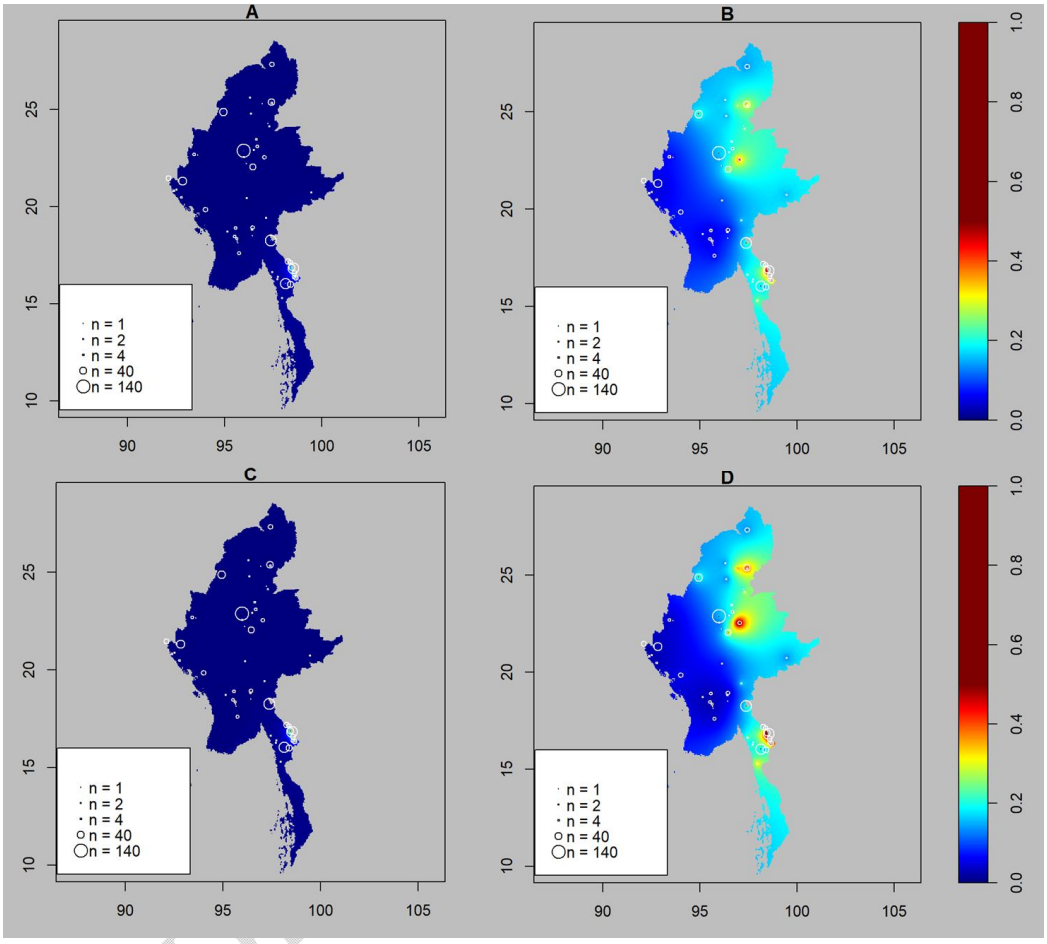

Figure 4. Geostatistical maps of estimated AR prevalence in Myanmar as generated by R-INLA with the BYM model specification using Metric 1 or Metric 2 and defined in Table 1 respectively of either K13 mutant allele *confirmed* or *confirmed and candidate* codons based on the WHO 2015 list (WHO 2015), as presented in the Tun et al (2015) data. Respective maps are shown for an underlying Poisson distribution in (A) with Metric 1, (B) with Metric 2 and for an underlying binomial distribution in (C) with Metric 1 and (D) Metric 2. The colour scale in all the maps ranges from dark blue (0%) to dark red (>=50%) as scaled in Tun et al (2015) data, and site sample sizes are superimposed as circles.

In Figure 4A for the Poisson model and Figure 4C for the binomial model, where the Metric 1 consists of only the sole codon C580Y as confirmed to confer AR by the WHO as at 2015. There is minimal estimated spread of AR, as represented by a tiny zone located in a south eastern location near the Kayin border with Thailand. These results contrast strongly with the maps generated by each model shown respectively

in Figure 4B and Figure 4D, where the confirmed and candidate codons in Metric 2 (9 in total), as listed by the WHO at 2015 are included. The presented maps seem to show a wider regional estimated prevalence of AR in the eastern half of the country. Additional hotspots are shown together with two clearly defined (additional hotspots) ones in the north of the country at Kachin (driven by codon F446I) and immediately below that in a more central location (driven by codon P574L).

Similarly, comparisons are made in Figure 5 for Metric 3 and Metric 4, based on the more recently published WHO's (2018) list for confirmed and candidate mutant codons as at 2018. Here, the codons that were included into the AR metric were the confirmed codons for Metric 3, resulting in the R-INLA model maps as shown in Figures 5A and 5D respectively, for the Poisson and binomial models and the confirmed and candidate codons for Metric 4 as shown in Figures 5B and 5C (for the) respective. (Poisson and binomial models).

Again, the disparity between estimated prevalence of AR by Metric 3 in the geostatistical maps, as shown in Figure 5 (A and C) when compared with those obtained by the g440 metric as shown in Figure 3(A and C) is readily apparent. In the former, there are only two miniscule hotspots that are indicated by red/yellow locations, consistent with those of Metric 2, one in the northern side of the country in Kachin and (one) again in (a) south-eastern location near the Kyin border with Thailand. However, in the latter location, these hotspots are completely subsumed within the much larger red/yellow zone, which extends along the entire eastern border region of the country on the right side of the map. In addition, there is also now a larger hotspot, immediately below Kachin, contained within the central region of this zone.

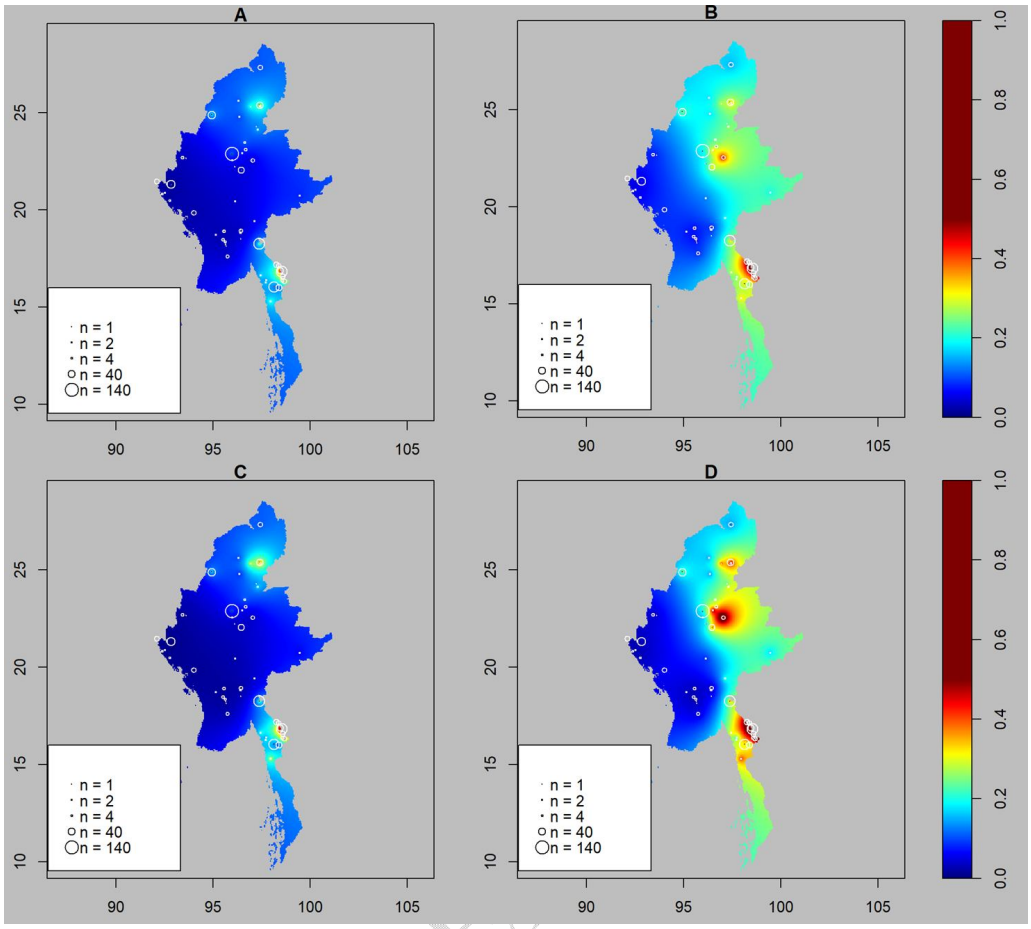

Figure 5. Geostatistical maps of estimated AR prevalence in Myanmar as generated by R-INLA with the BYM model specification using Metric 3 or Metric 4, again as defined in Table 1 respectively of either K13 mutant allele *confirmed* or *confirmed and candidate* codons based on the WHO (2018) list ((WHO 2018)), as presented in the Tun et al (2015) data. The maps are shown for an underlying Poisson distribution in: (A) with Metric 3, (B) with Metric 4 and for an underlying binomial distribution, in (C) with Metric 3 and (D) Metric 4. The colour scale in all the maps shown, ranges from dark blue (0%) to dark red (>=50%) as scaled in Tun et al (2015). Site sample sizes are superimposed as circles.

In summary, the R-INLA geostatistical maps in Figures 4 and (Figure) 5 demonstrate the disparity between prevalence of AR when estimated by the g440 metric when compared with the other four Metrics based on the WHO 2015 and 2018 published lists. The maps demonstrate that the inclusion (or not) of candidate codons known (in) as listed in 2015 and (or) 2018 (listed) by the WHO has a profound influence on the assessment of AR spread. They also demonstrate a wide range for the estimated AR prevalence but, which broadly, increases in spatial coverage as the number of mutant codons included into the AR metric are increased. This can be seen to be more pronounced in 2018 with the inclusion of more candidates in the larger WHO list of 2018, by a direct comparison of all the respective maps shown in Figures 4ABCD with those of Figures 5ABCD. MOVE TO DISCUSSION SECTION

A parallel comparison of results was obtained for the g440 metric and the four WHO-based Metrics in terms of either the spatial sites coverage or population proportion-spread for mutant allele frequency distributions. This is shown by the bar plots in Figures 6 and 7. Repeating the respective comparison with bar charts of corresponding frequency distributions of either individuals or sites, namely either those shown in Figures 6(A and C) or 7(A and C) with those shown in Figure 2 (A and B), have also revealed that the estimated prevalence of AR by Metric 3 is driven by just 5 codons (namely F446I, N458Y, M476I, R561H, C580Y) out of the 24 in total as were incorporated into the g440 metric model by Tun et al (2015).

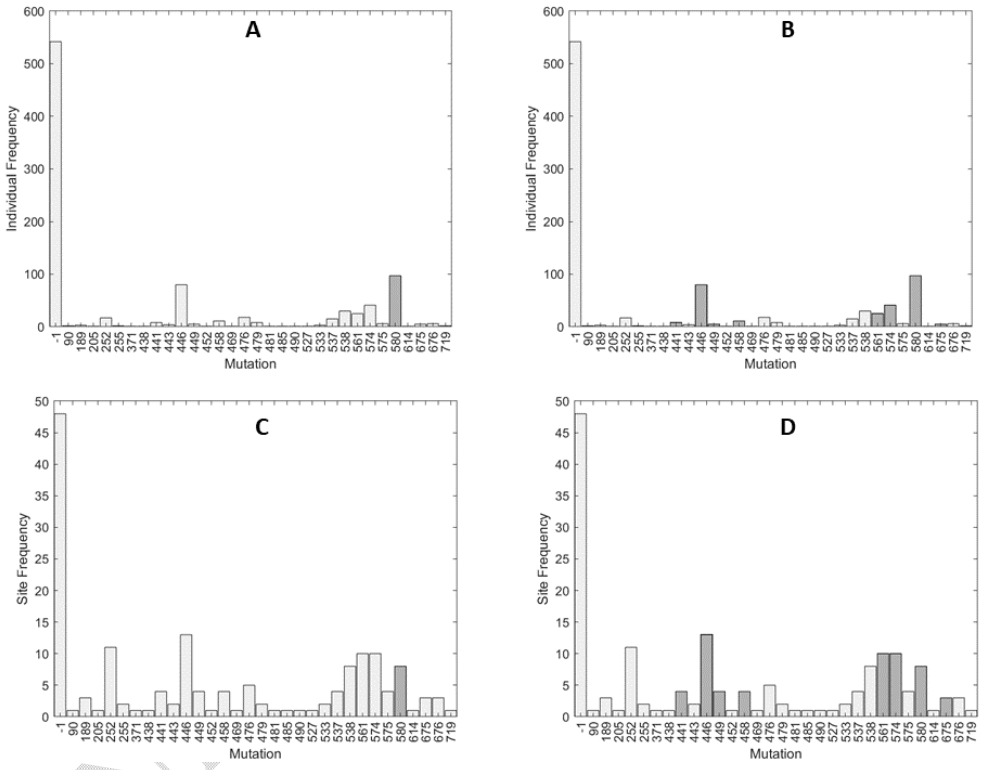

Figure 6. Bar plots of K13 mutant allele codon frequencies defined by Metric 1 and Metric 2 in Table 1, respectively based on the WHO 2015 list (WHO 2015) of either K13 *confirmed* or *confirmed and candidate* codons, present in the Tun et al (2015) data, as shown for a number of individuals (total *n*=940) in (A) for Metric 1 and (B) for Metric 2, for number of sites (total *n*=54) in (C) for Metric 1 and (D) for Metric 2.

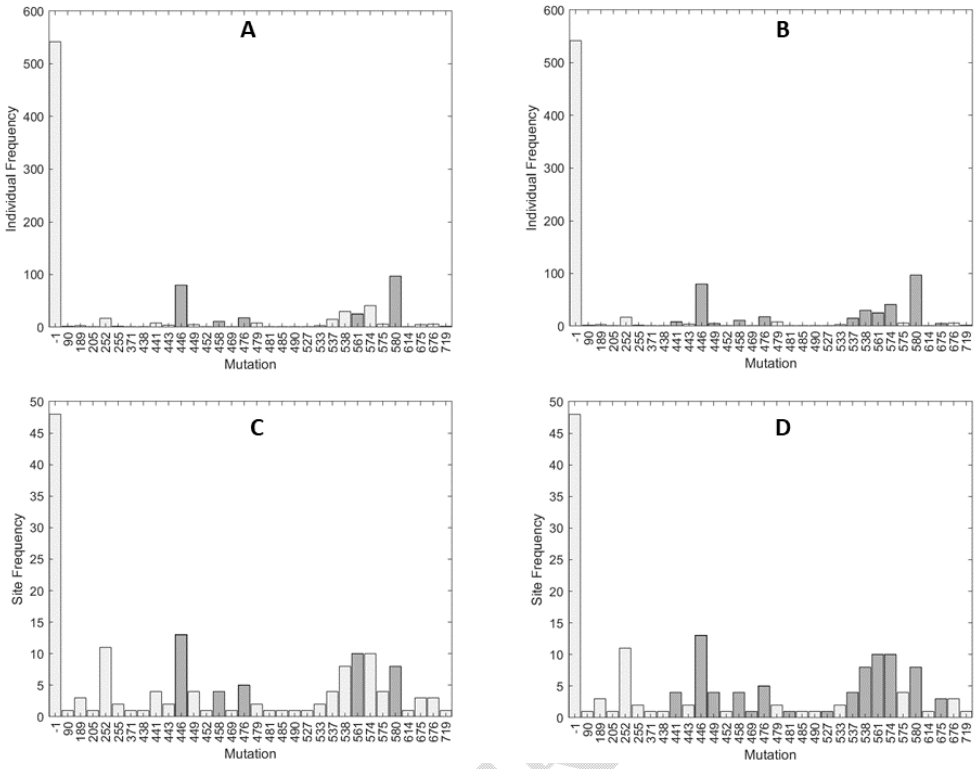

Figure 7. Bar plots of mutant allele codon frequencies as defined in (by) Metric 3 and Metric 4 in Table 1, respectively were based on the WHO 2018 ((WHO 2018)) list of either K13 mutant allele *confirmed* or? unconfirmed *and candidate* codons, as presented in the Tun et al (2015) data. The data shown represent a (for) number of individuals (total *n*=940) in (A) for Metric 3, (B) for Metric 4 and for number of sites (total *n*=54) in (C) for Metric 3 and (D) for Metric 4.

Table 2 compares all the summary statistics for the g440 metric of Tun et al (2015) with those of the four WHO based Metrics that were used for assessing the spread of AR by either population proportion, spatial site coverage or R-INLA geostatistical map. These data demonstrate that by using the g440 metric, consistent overestimation of the AR prevalence could occur when compared to all four of the WHO based metrics.

Metric 1 produces a clear underestimate of the prevalence of AR in the light of the Metric 3 benchmark. This implies that the WHO perception of AR prevalence in Myanmar in 2015, based on the codons, then known to confer AR (only codon C580Y), was a clear underestimation. By the same benchmark of Metric 3, if all the candidate codons in the WHO (2015) list are included to give Metric 2, an overestimate (is) could then be produced. The overestimation could necessarily be increased by extending the Metric 3 to Metric 4 through inclusion of all the candidate codons in the WHO (2018) list. However, this overestimate is still well

below that produced by the g440 metric, as can be seen by a direct comparison with any of the respective summary statistics (Table 2).

Table 2: LABEL THE TABLE

Summary statistic of AR spread	g440	Metric 1	Metric 2	Metric 3	Metric 4
Geostatistical map					
Poisson mean prevalence (%)	22	1.0	15	8.0	18
Binomial mean prevalence (%)	24	0.0	14	7.0	18
Population					
Population proportion (%)	39	10	29	25	36
Number of individuals (n=940)	*371*	*97*	*272*	*231*	*338*
Sites					
Sites proportion (%)	57	15	50	43	52
Number of sites (n=54)	*31*	*8*	*27*	*23*	*28*

Table 2. Measures of 'spread' of AR estimated by the g440 metric of Tun et al (2015) and by the 4 Metrics defined in Table 1 of this paper. Metric 1 and Metric 2 are respectively defined as the K13 mutant allele *confirmed codons* or *confirmed and candidate codons* to confer AR in the WHO (2015) list, Metric 3 and Metric 4 are respectively defined as the K13 mutant allele *confirmed codons* or *confirmed and candidate codons* to confer AR in the WHO (2018) list. The values are shown in percentages to 2 significant figures and evaluated by using the 3 different estimation approaches of: mean prevalence in the geostatistical map generated by R-INLA with the BYM model specification with either (top row) a Poisson or Binomial underlying distribution; (middle row) proportion of sampled population; (bottom row) proportion of sampled sites sampled.

Finally, a comparison of the mean estimated prevalence in Table 2, obtained with the R-INLA geostatistical maps with either underlying distribution (Poisson or binomial)., re-enforces the point again. The mean prevalence of AR estimated by the g440 metric (22% or 24%) and by the benchmark Metric 3 (8% or 7%) differ by a factor of 3. The margin is lower when a comparison with Metric 4 is made through inclusion of all the candidate codons as published by the WHO in 2018 (18% or 18%) and the overestimation of the mean prevalence of AR is then reduced to around 33%??.

4. Discussion

The Tun et al (2015) assessment of AR 'spread' in Myanmar differs radically from the AR prevalence estimated in this paper based on WHO (2015) data of codons associated with AR. A perspective based on the later updated WHO (2018) data,

although achieving a higher estimation of AR prevalence, also gives a prevalence level which is far below that estimated by Tun et al (2015).

If the AR metric is restricted to mutant alleles as *confirmed to confer* AR in the WHO (2018) list by the benchmark Metric 3, only two small localised hot spots of artemisinin resistance at disparate locations are identified in the south and north of Myanmar. It is impossible to view this map as 'spread' of AR as purveyed by Tun et al (2015). Even by including all the known *candidate* mutant alleles in WHO (2018) by Metric 4, the degree of contiguity in the estimated prevalence map remains far below that of the extensive geographic coverage conveyed by the g440 metric.

In the light of increasing scientific knowledge on genetic markers, the WHO (2015, 2018) data may not have been necessarily definitive. Indeed, the efficacy of the role of the WHO in tackling global health matters has often been unclear and viewed with scepticism, as was re-iterated by the recent failure of many countries to heed warnings of the covid pandemic issued by the WHO in January 2020 (Maxmen 2021). However, the WHO (2018) data are still the most current officially published data on codons known to be associated with conferring AR. The analysis presented in this paper shows that by asserting all codons above codon number 440 in the propeller region, of the K13 conferred artemisinin resistance, as was done by Tun et al (2015) could lead to an excessive misrepresentation of the estimation of 'spread' of AR in Myanmar.

This analysis also shows that prevalence of AR's estimates (ed) from WHO's (2015) and (WHO) (2018) published data, seemed to have followed a logical progression supportable by the geostatistical maps, which were produced from either the confirmed or candidate codons to confer AR (which were) as listed. In addition, the (se) WHO data could provide a clear benchmark to calibrate the importance of ongoing future progress of K13 mutant allele codon identification on AR.

While it is prudent to adopt a precautionary approach by evaluating worst case scenarios for the spread of AR, it is necessary that the robustness of such assertions is meticulously examined to ensure that they are scientifically rigorous. This principle was notably absent in the Tun et al (2015) estimation of AR by the g440 metric in Myanmar.

Some leading authors have warned that the controversial pursuit of 'AR spread' may deflect the focus away from other key research issues such as partner-drug efficacy when used in artemisinin combination therapies (ACTs) or the importance of initial parasite densities in tackling malaria endemicity (e.g. Ferreira et al., 2013; Krishna & Kremsner, 2013).

The danger of 'crying wolf' (Meshnick, 2012) is (of) very invoking 'once bitten, twice shy'. That will mean estimations of true emerging artemisinin resistance, wherever they might actually occur, are not given the attention they deserve in the future research work.

References

Ariey, F., Witkowski, B., Amaratunga, C., Beghain, J., Langlois, A.-C., Khim, N., Kim, S., Duru, V., Bouchier, C., Ma, L., Lim, P., Leang, R., Duong, S., Sreng, S., Suon, S., Chuor, C. M., Bout, D. M., Ménard, S., Rogers, W. O., … Ménard, D. (2014). A molecular marker of artemisinin-resistant Plasmodium falciparum malaria. *Nature*, *505*(7481), 50–55. https://doi.org/10.1038/nature12876

Besag, J., York, J., & Mollié, A. (1991). A Bayesian image restoration with two applications in spatial statistics Ann Inst Statist Math 43: 1–59. *Find This Article Online*, *43*(1), 1–20.

Chookajorn, T. (2018). How to combat emerging artemisinin resistance: Lessons from "The Three Little Pigs." *PLoS Pathogens*, *14*(4), 4–11. https://doi.org/10.1371/journal.ppat.1006923

Fairhurst, R. M. (2015). Understanding artemisinin-resistant malaria. *Current Opinion in Infectious Diseases*, *28*(5), 417–425. https://doi.org/10.1097/qco.0000000000000199

Ferreira, P. E., Culleton, R., Gil, J. P., & Meshnick, S. R. (2013). Artemisinin resistance in Plasmodium falciparum: What is it really? *Trends in Parasitology*, *29*(7), 318–320. https://doi.org/10.1016/j.pt.2013.05.002

Grist, E. P. M., Flegg, J. A., Humphreys, G., Mas, I. S., Anderson, T. J. C., Ashley, E. A., Day, N. P. J., Dhorda, M., Dondorp, A. M., Faiz, M. A., Gething, P. W., Hien, T. T., Hlaing, T. M., Imwong, M., Kindermans, J. M., Maude, R. J., Mayxay, M., McDew-White, M., Menard, D., … Guerin, P. J. (2016). Optimal health and disease management using spatial uncertainty: A geographic characterization of emergent artemisinin-resistant Plasmodium falciparum distributions in Southeast Asia. *International Journal of Health Geographics*, *15*(1), 1–10. https://doi.org/10.1186/s12942-016-0064-6

Kaestli, M., Grist, E. P. M., Ward, L., Hill, A., Mayo, M., & Currie, B. J. (2016). The association of melioidosis with climatic factors in Darwin, Australia: A 23-year time-series analysis. *Journal of Infection*, *72*(6), 687–697. https://doi.org/10.1016/j.jinf.2016.02.015

Krishna, S., & Kremsner, P. G. (2013). Artemisinin resistance needs to be defined rigorously to be understood: Response to dondorp and ringwald. *Trends in Parasitology*, *29*(8), 361–362. https://doi.org/10.1016/j.pt.2013.05.006

Lawson, A. B. (2008). Bayesian disease mapping: Hierarchical modeling in spatial epidemiology. In *Bayesian Disease Mapping: Hierarchical Modeling in Spatial Epidemiology*. https://doi.org/10.1111/j.1467-985x.2010.00681_11.x

Lindgren, F., & Rue, H. (2015). Bayesian spatial modelling with R-INLA. *Journal of Statistical Software*, *63*(19), 1–25. https://doi.org/10.18637/jss.v063.i19

Maxmen A. (2021) Why did the world's pandemic warning fail for covid? *Nature* 589, 499-500.

Ménard, D., Khim, N., Beghain, J., Adegnika, A. A., Shafiul-Alam, M., Amodu, O., Rahim-Awab, G., Barnadas, C., Berry, A., Boum, Y., Bustos, M. D., Cao, J., Chen, J.-H., Collet, L., Cui, L., Thakur, G.-D., Dieye, A., Djallé, D., Dorkenoo, M. A., … Mercereau-Puijalon, O. (2016). A Worldwide Map of Plasmodium falciparum K13-Propeller Polymorphisms . *New England Journal of Medicine*, *374*(25), 2453–2464. https://doi.org/10.1056/nejmoa1513137

Meshnick, S. (2012). Perspective: Artemisinin-resistant malaria and the wolf. *American Journal of Tropical Medicine and Hygiene*, *87*(5), 783–784. https://doi.org/10.4269/ajtmh.2012.12-0388

Morgan O., Vreiheid M. and Dolk H. (2004) risk of low birth weight near eurohazcon hazardous waste landfilll sites in England. Archives of Environmental health 59, 149-151.

Takala-Harrison, S., Jacob, C. G., Arze, C., Cummings, M. P., Silva, J. C., Dondorp, A. M., Fukuda, M. M., Hien, T. T., Mayxay, M., Noedl, H., Nosten, F., Kyaw, M. P., Nhien, N. T. T., Imwong, M., Bethell, D., Se, Y., Lon, C., Tyner, S. D., Saunders, D. L., … Plowe, C. V. (2015). Independent emergence of artemisinin resistance mutations among Plasmodium falciparum in Southeast Asia. *The Journal of Infectious Diseases*, *211*(5), 670–679. https://doi.org/10.1093/infdis/jiu491

Tun, K. M., Imwong, M., Lwin, K. M., Win, A. a, Hlaing, T. M., Hlaing, T., Lin, K., Kyaw, M. P., Plewes, K., Faiz, M. A., Dhorda, M., Cheah, P. Y., Pukrittayakamee, S., Ashley, E. a, Anderson, T. J. C., Nair, S., McDew-White, M., Flegg, J. a, Grist, E. P. M., … Woodrow, C. J. (2015). Spread of artemisinin-resistant Plasmodium falciparum in Myanmar: a cross-sectional survey of the K13 molecular marker. *The Lancet Infectious Diseases*, 21–26. https://doi.org/10.1016/S1473-3099(15)70032-0

Walker L. and Gotway C. (2004) Applied spatial statistics for public health, New York, Wiley

Woodrow, C. J., & White, N. J. (2017). The clinical impact of artemisinin resistance in Southeast Asia and the potential for future spread. *FEMS Microbiology Reviews*, *41*(1), 34–48. https://doi.org/10.1093/femsre/fuw037

World Health Organization. Status Report on Artemisinin Resistance: September 2015. Geneva, Switzerland: World Health Organization, September 2015.

World Health Organization. (2018). Status report on artemisinin resistance and ACT

efficacy (August 2018). *World Health Organization*, *August*, 10. https://www.who.int/malaria/publications/atoz/artemisinin-resistance-august2018/en/

Zaw, M. T., Emran, N. A., & Lin, Z. (2018). Updates on k13 mutant alleles for artemisinin resistance in Plasmodium falciparum. *Journal of Microbiology, Immunology and Infection*, *51*(2), 159–165. https://doi.org/10.1016/j.jmii.2017.06.009

Appendix

Poisson model

The case count y_i is assumed to have a mean μ_i within each small area *i* of the gridded 5km x 5km domain [covering the whole of the Myanmar map] and to be independently distributed as

$$y_i \sim Poisson(\mu_i)$$

so the likelihood is

$$L(\boldsymbol{y}|\boldsymbol{\mu}) = \prod_{i=1}^{m} \mu_i^{y_i} \exp(-\mu_i)/\, y_i!$$

The usual assumption is made that data are independently distributed with expectation

$$E(y_i) = \mu_i = e_i \theta_i$$

where e_i is the expected rate for the *i* th area, θ_i is the relative risk for the *i* th area and $\{y_i\}$ are assumed to be conditionally independent given knowledge of $\{\theta_i\}$.

The relative risk θ_i is then modelled as a linear predictor by assuming the logarithmic link

$$log\theta_i = \eta_i$$

Binomial model

The probability of a case is assumed to be p_i and the case count y_i within each small area *i* of the gridded 5km x 5km domain [covering the whole of the Myanmar map] is assumed to be distributed independently as

$$y_i \sim bin(p_i, n_i)$$

so, the likelihood is

$$L(y_i|p_i, n_i) = \prod_{i=1}^{m} \binom{n_i}{y_i} p_i^{y_i}(1-p_i)\ (1-p_i)^{(n_i-y_i)}$$

and for p_i to be a linear predictor, a logit link is chosen so that

$$p_i = \frac{exp(\eta_i)}{1+exp(\eta_i)}$$

The BYM model is then specified by

$$\eta_i = \alpha + v_i\ + u_i$$

Where,

$$v_i \sim Normal(0, \tau_v^2)$$

is the unstructured residual and

$$u_i | u_{j \neq i} \sim Normal(\overline{u_i}, \tau_i^2)$$

is the spatially structured residual modelled as an intrinsic conditional autoregressive (iCAR) structure with

$$\overline{u_i} = \frac{\sum_{j \in N(i)} u_j}{\#N(i)}$$

and

$$\tau_i^2 = \frac{\tau_u^2}{\#N(i)}$$

where $N(i)$ denotes a neighbouring area to area *i* and $\#N(i)$ denotes the number of areas which have boundaries with the *i*th one. So τ_v^2 and τ_u^2 control the number of random effects respectively in v_i and u_i