Use **global envelope tests** not **pointwise CIs**, for **graphical hypothesis tests** of **spatiotemporal clustering** and **tau statistic** τ.

Clustering range estimate more precise with modified **bootstrap sampling method**.

Spatiotemporal clustering using the tau statistic

@TimPollington, Tildesley, Hollingsworth & Chapman

INTRO.

- The *tau statistic* τ, uses geolocation and, usually, symptom onset time to assess global spatiotemporal clustering and estimate its extent, for epi. studies [1,2]. This is the first review into its use [3], and how computation/implementation could bias τ estimates [4] (Figs. 1 & 2).

METHODS

[3] Literature review of works citing the *foundation* papers [1,2] using Google Scholar. Papers tabulated by key characteristics.

[4] Compare baseline analysis of open access measles dataset [2] vs. modified implementation of graphical hypothesis test, and estimated clustering range by bootstrap sampling method (Figs. 6 & 8); number of bootstrap samples *N* (Figs. 6 & 7); CI type (bias-corrected & accelerated (BCa) vs percentile); distance band sets (Fig. 9).

RESULTS

[3]

- Only ½ of the 16 studies used τ as defined by [1,2].
- All papers using graphical hypothesis tests or parameter estimation, applied them incorrectly. Corrections were explained (Figs. 3-5).
- Unclear how to choose time relatedness interval (to relate pairs), or the distance band set, both required in the τ calculation.
- New untested τ statistic estimator for disease rates defined, for datasets with variable individual person–time-at-risk.

[4]

- Re-analysis of [2] still found evidence against no spatiotemporal clustering, p ∈ [0,0·014] (global envelope test) (Fig. 5).
- τ-specific modification of Loh & Stein bootstrap sampling method developed; its more precise bootstrapped τ estimates led to 20% higher clustering range than previously published (36·0m, 95% BCa CI (14·9,46·6), vs 30m) (Fig. 10).
- Estimated bias reduction led to increase in the clustering area of elevated disease odds by 44%. BCa CIs are essential for asymmetric sample bootstrap distributions of τ estimates.

CONCLUSIONS

- Bootstrap sampling method and CI type can bias the clustering range estimate. Moderate radial bias to the range estimate are more than doubled when considered on the areal scale, which public health resources are proportional to. We advocate proper implementation of this useful statistic, ultimately to reduce inaccuracies in control policy decisions made during disease clustering analysis.

REFERENCES

[1] Salje et al. *PNAS*. 2012

[2] Lessler et al. *PLoS One*. 2016

[3] Pollington et al. *arXiv/stat.AP*. 2019

[4] Pollington et al. *arXiv/stat.ME*. 2019

With thanks to …

NTD Modelling Consortium · MRC Medical Research Council · centre for complexity science · EPSRC Engineering and Physical Sciences Research Council · WARWICK The University of Warwick

and Tom Crellen & Will Probert for poster review. Poster design inspired by Mike Morrison, http://osf.io/ef53g

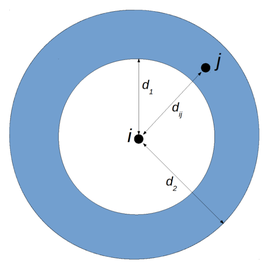
TAU STATISTIC DEFN.

$$\hat{\tau}_{\text{odds}}(d_1, d_2) = \frac{\hat{\theta}(d_1,d_2)}{\hat{\theta}(0,\infty)}$$

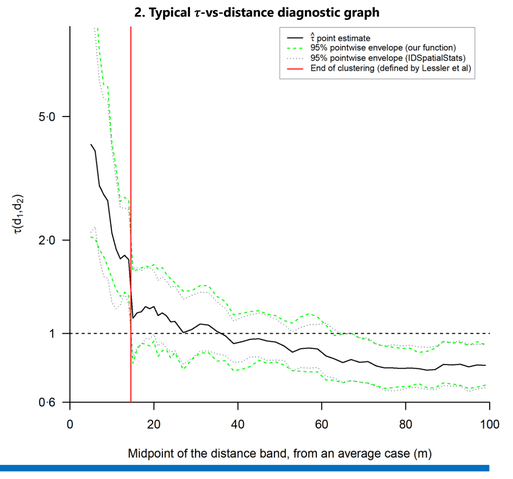
$$\hat{\theta}(d_1, d_2) = \frac{\sum_{i}^{N}\sum_{j,j\neq i}^{N} \mathbb{1}(z_{ij}=1,\ d_1 \leq d_{ij} < d_2)}{\sum_{i}^{N}\sum_{j,j\neq i}^{N} \mathbb{1}(z_{ij}=0,\ d_1 \leq d_{ij} < d_2)}$$

$$z_{ij} = \begin{cases} 1, & \text{if } |t_i - t_j| \leq \text{S.I.} \\ 0, & \text{otherwise} \end{cases}$$

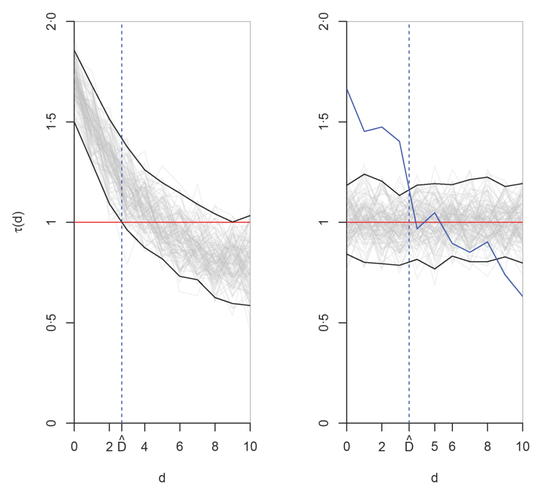
1. A distance band (d_1,d_2] around a case *i*



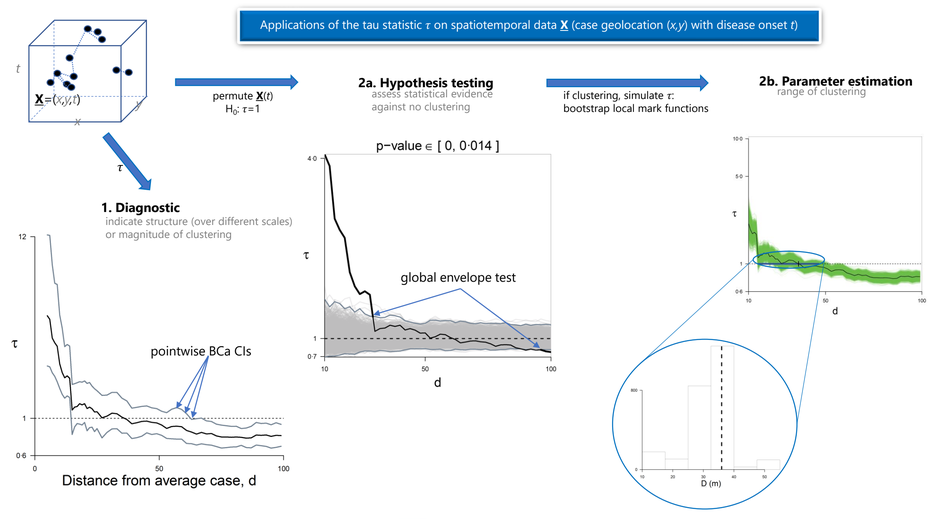
2. Typical τ-vs-distance diagnostic graph



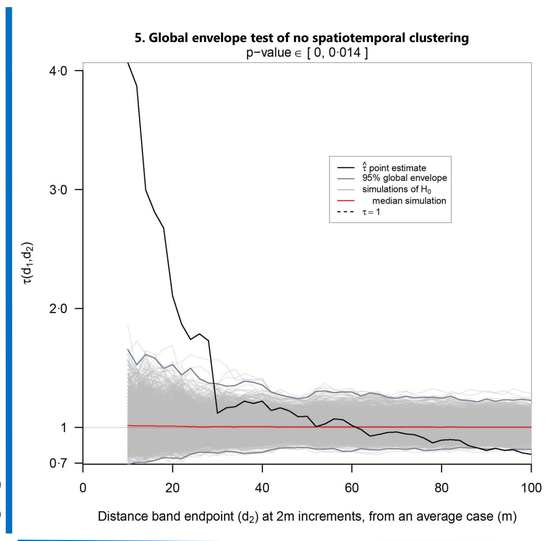
3. Incorrect method of using central (L) or null (R) envelopes and where they meet τ=1 or $\hat{\tau}$ point estimate , respectively, to estimate clustering range $\hat{D}$



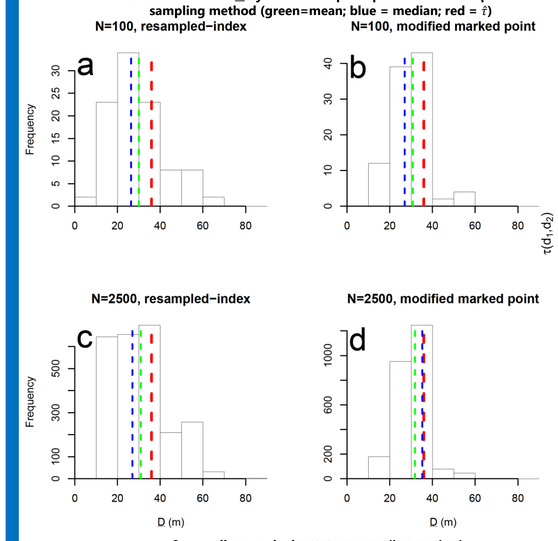
4. Graphical abstract of proper methods for diagnosis, hypothesis testing & parameter estimation



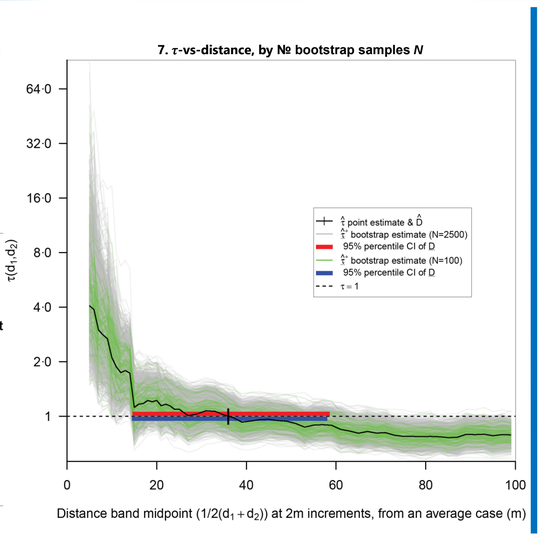
5. Global envelope test of no spatiotemporal clustering



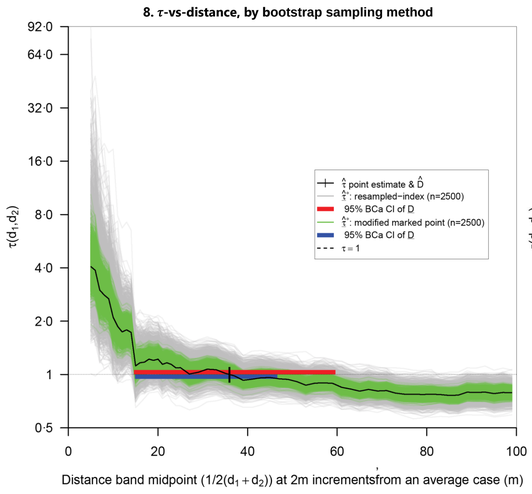
6. Distribution of D by № bootstrap samples *N* & bootstrap sampling method (green=mean; blue = median; red = $\hat{\tau}$)



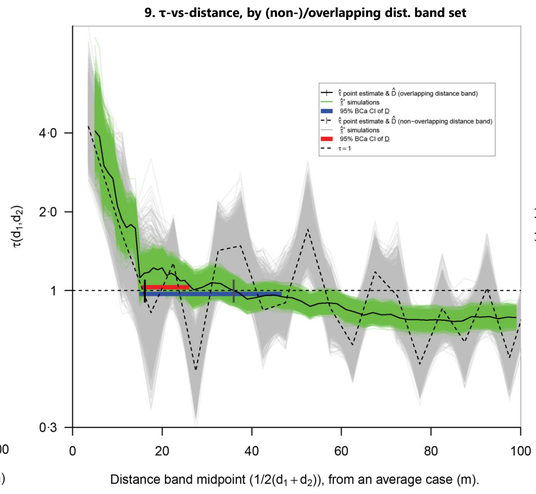
7. τ-vs-distance, by № bootstrap samples *N*



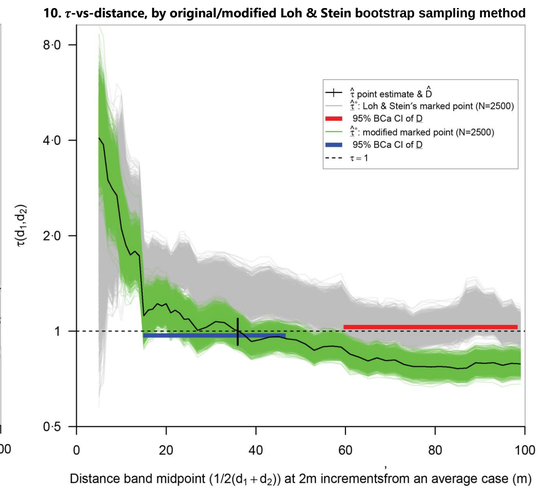
8. τ-vs-distance, by bootstrap sampling method



9. τ-vs-distance, by (non-)/overlapping dist. band set



10. τ-vs-distance, by original/modified Loh & Stein bootstrap sampling method



Download poster

[3] Literature review
***arXiv*: 1911.11476**

Download preprints

[4] Technical analysis
***arXiv*: 1911.08022**