Estimating Location of Clusters in a Surveillance Setting

C. Jeffery [1], A. Ozonoff [1, 2], L. White [1, 2], M. Pagano [1]

[1] Harvard School of Public Health, [2] Boston University School of Public Health, Departments of Biostatistics

We focus on the problem of comparing the distribution of observations in a planar region to a prespecified null distribution. Our motivation is a surveillance setting where we map locations of incident disease, aiming to monitor these data over time, to locate potential spatial disturbances and to direct public health actions.

We propose a non-parametric approach to distance-based disease mapping inspired by tomographic imaging. We consider several univariate projections via the observed distribution of distances to a chosen fixed point; we then compare this distribution to that expected under the null and average this measure across projections to compute a relative-risk-like score at each point in the region. The null distribution is typically determined by historical data. Scores are displayed on the map using a color scale.

To assess the performance of this method, we compare it to the widely used ratio of kernel density estimates. As a performance metric, we evaluate the power to detect simulated clusters superimposed on a uniform distribution in the unit disk. Results suggest that both methods can adequately locate spatial disturbances but each rely on an appropriate choice of parameters. However, our proposed method can generalize to arbitrary metric spaces and/or high-dimensional data.