Bayesian spatial analysis of the vaccine registry data

How does vaccine coverage change over space & time?

- (Spatial regression) Relationship between vaccine coverage and covariates
  - Beware of potential ecological bias
- (Disease mapping) Temporal-spatial predicted maps at commune/district/province level of
  - Expected vaccine coverage rate
  - Ratio of observed and expected vaccine coverage rate
  - Probability of ratio of observed and expected vaccine coverage rate below a certain threshold

- Scales: depending on computational cost
  - Time: starting with month (2017-2022)
  - Space: starting with commune level
- Pathogens (of recent interest, due to recent/ongoing/potential outbreaks)
  - Measles: also evidence about the impact of allowing private clinics to provide MMR vaccination at 12m is needed (WHO want to push MoH to force measles vaccination starting at 9m for both public/private vaccination)
  - Diphtheria

Data

- Lattice/discrete spatial variation/spatial interaction data (small area data) but will be treated as continuous spatial variation data (geostatistical data)
- Outcome:
  - Vaccine coverage:
    - Number of children at age a, time t, commune $s = N(a, t, s)$
      - Option 1: Number of children identified from the vaccine registry data (starting with this, age as continuous)
      - Option 2: Number of children estimated from other data sources (WorldPop, age as group)
    - Number of vaccinated children at age a, time t, commune $s = V(a, t, s)$
- Exposures (at commune level)
  - Demographics: population density, age structure, gender ratio
  - Geographics: elevation, proximity to land border (to proxy remoteness, use smallest distance to city/urban area instead)
  - *Interventions: outbreak detection, supplementary immunisation activities?*
    - *Vaccination type [routine/campaign] is available but did not correct in the early period*
    - *Public/Private vaccination: may differ in schedule, type of vaccine (eg. Measles: MMR at 12m [private], MR at 9m-18m [public])*

Data

- By commune, year, age
  - Total number of children (identified from the vaccine registry)
  - Number of vaccinated children (by pathogen/by shot/by public/private)
- By commune, year
  - Total population -> population density
  - Age structure
  - Gender ratio
- By commune:
  - Land area
  - Average elevation
  - ~~Proximity to land border~~ Remoteness (Closest distance to a city/urban area)
  - District, province
  - Centroid: geometric centroid, population centroid

Analysis approach

- Geostatistical approach

$$V(a,t,s) \sim Bin(N(a,t,s), p(a,t,s))$$

$$log(\frac{p(a,t,s)}{1-p(a,t,s)}) \sim \beta \times X_{t,s} + U(t,s)$$

  - Age will be modelled using a flexible smooth function, may also interact with time and space
  - The lattice approach does not feasible in this situation (require huge neighbourhood matrix)
  - Integrated nested Laplace approximation—stochastic partial differential equation (INLA-SPDE)
    - R inla, inlabru

Potential problems (to think ahead/sensitivity analysis/discussion)

- Data quality
  - Generally low at the beginning but was improved over time
  - Analysis with and without removing commune(s) with “strange” data
- Computational cost
  - Use geostatistical approach instead of lattice approach
  - Start with the finest scale but could change to bigger scale (so less computational cost) later on
- Using geostatistical approach to analyse lattice data
  - Prediction of outcome between communes with large distance could be problematic
  - Use different centroids (geometry centroid, population centroid)
- Locations near border with Cambodia, Laos, China
  - Could be affected by low vaccine coverage of the neighbour communities in other countries
- Possibility of ecological bias when interpreting relationships between covariates and vaccine coverage
  - Discussion
- Potential mismatches when merging map data and geospatial data
  - Communes were split and merged during the data collection period
  - We will use the most recent status of administration structure. Lam documented communes with split and merge over time

Plan

- Lam prepare data to share by next week (22-26/07)
- After receive data from Lam, Quynh Anh could help to identify anomaly commune(s) to remove (criteria still need to be defined)
- Olivier will help with the remoteness indicator