Social contact dispersal and its relevance for the construction of epidemic models

Universiteit Antwerpen
UHASSELT

K. Van Kerckhove[1], L. Willem[2], C. Faes[1], P. Beutels[2], V. Colizza[3,4] and N. Hens[1,2]

1. Hasselt University, Belgium 2. University of Antwerp, Belgium 3. Univ. Paris 06, France 4. ISI Foundation, Italy

niel.hens@uhasselt.be

Introduction

- In the last decade, various authors found a power law to best describe the spatial spread of infectious diseases [1−5].
- Geilhufe *et al.* (2014)[5] suggested that using social contact data can offer further improvements and insights. Meyer and Held (2017) [6] could not identify age-dependent decays across distance for norovirus gastroenteritis and suggested using social contact matrices across distance.
- Using the social contact hypothesis [7], one can question whether this power law can be observed in social contact dispersal as well or whether transmission and social contact dispersal differ.
- A systematic review of social contact surveys [8] reveals a limited number of contact studies contain information about the relationship between contacts and distance [9−14] though more in depth analyses are lacking.
- We present results on the relationship between contacts and distance based on a social contact survey conducted between September 2010 and July 2011 in the Flemish geographical region.
- Our aim is twofold:
  1. Testing various underlying continuous degree distributions over distance, including the power law distribution.
  2. Estimating social contact matrices and the force of infection over distance.

Data and Methods

- We use data from a social contact survey conducted in 2010-2011 in Flanders [15,16,17] which had a similar set-up as the POLYMOD Survey [18] requesting participants or proxies to fill in a contact diary during one randomly assigned day (24 hours).
- In addition, the time use during the assigned day was questioned using pre-specified time blocks. Each time block participants indicated the one location at which they spent most of their time and the distance from home for this location (0-1 km, 2-9 km, 10-74 km, >75 km).
- Given the location in both contact and time use surveys, the contacts were mapped to distances. Since contacts were recorded only once and for possibly multiple locations, subcontacts were defined making up a distributed contact over different locations. We defined, for each participant t,
  - a 4-tuple $\mathbf{Y}_t = (Y_{1t}, Y_{2t}, Y_{3t}, Y_{4t})'$ with Y_{dt} the number of contacts made at the ordered distance category d
  - a 4-tuple $\mathbf{D}_t$ with D_{dt} the indicator variable taking value 1 if participant t has travelled distance d and 0 otherwise
  - weights w_{it}^d for handling distributed contacts over different locations
- We model Y_{dt} as

$$Y_{dt} \sim \text{NegBin}(m_{dt}, \phi_d),$$

where $\log(m_{dt}) = \log(\bar{y}_{\cdot\cdot}) + \log(f^*(d, \boldsymbol{\theta}_t))$ with $f^*(d, \boldsymbol{\theta}_t)$ a distance function for which $\sum_{d=1}^{4} f^*(d, \boldsymbol{\theta}_t) = 1$ and $\log(\phi_d) = \eta_d$.

  - Denote $f(d^*, \boldsymbol{\theta}_t)$ for a continuous (parametric) distance density, with $\boldsymbol{\theta}_t$ representing the parameters of the distance function, which we regress on participant characteristics.
  - The link between $f(d^*, \boldsymbol{\theta}_t)$ and $f^*(d, \boldsymbol{\theta}_t)$ can be achieved by integrating $f(d^*, \boldsymbol{\theta}_t)$ over the distance categories, e.g.

$$f^*(1, \boldsymbol{\theta}_t) = \int_0^2 f(d^*, \boldsymbol{\theta}_t) dd^*.$$

  - Different parametric forms for $f(d^*, \boldsymbol{\theta}_t)$ and different participant characteristics are compared using AIC, BIC and MSPE.
- Contact matrices and forces of infection over distance were also calculated as follows:

	Conditional on displacement to distance d	Unconditional (independent travelled to distance d)
M	$\mathbf{M}_d$ with elements $m_{ijd} = E(Y_{ijdt} \vert D_{dt} = 1)$ $\hat{m}_{ijd} = \frac{\sum_{t=1}^{T_{id}} w_{it}^d y_{ijdt} D_{idt}}{\sum_{t=1}^{T_{id}} w_{it}^d D_{idt}}$	$\mathbf{M}_{d^*}$ with elements $m_{ijd^*} = E(Y_{ijdt})$ $\hat{m}_{ijd^*} = \frac{\sum_{t=1}^{T_{id}} w_{it}^d y_{ijdt}}{\sum_{t=1}^{T_{id}} w_{it}^d}$
FOI	λ_{id} as cFOI $\lambda_{id} = \sum_j q m_{ijd} \iota_{jd}$	λ_{id^*} as uFOI $\lambda_{id^*} = \sum_j q m_{ijd^*} \iota_{jd}$

with ι_{jd} representing the assumed age-specific relative incidence at distance d.

- Results for the FOI not shown here.

References

[1] Held *et al.* (2005); [2] Held *et al.* (2006); [3] Paul and Held (2011); [4] Held and Paul (2012); [5] Geilhufe *et al.* (2014); [6] Meyer and Held (2017); [7] Wallinga *et al.* (2007); [8] Hoang *et al.* (In preparation); [9] Danon *et al.* (2012); [10] Danon *et al.* (2013); [11] Fu *et al.* (2012); [12] Stein *et al.* (2012); [13] Hoek *et al.* (2013); [14] Melegaro *et al.* (2017); [15] Willem *et al.* (2012); [16] Kifle *et al.* (2016); [17] Bilcke *et al.* (2017); [18] Mossong *et al.* (2008)

Results

- A total of 1560 participants for which linking was possible reported an overall median number of contacts of 10 (IQR: 6-19).
- Figure 1 clearly shows the differences in distance-contact behaviour for regular/holiday and weekday/weekend settings.

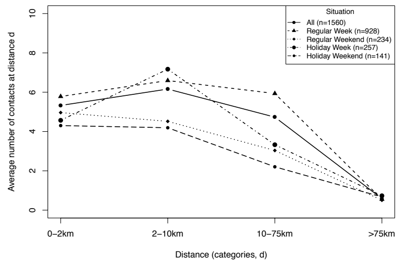

Figure 1: The average number of contacts at distance d ($\bar{Y}_d$) are connected by (full) lines for all participants.

- The lognormal model with age and week provided the best continuous degree distribution $f(d^*, \boldsymbol{\theta}_t)$ based on AIC and BIC whereas a saturated model with age, weekday and holiday provided the best distribution based on MSPE.
- For the conditional social contact matrices, yielding an individual based perspective, an increase in contacts amongst adults across distance is observed; an increase (decrease) in contacts amongst children is present at distance 2-10 km (beyond 10 km).

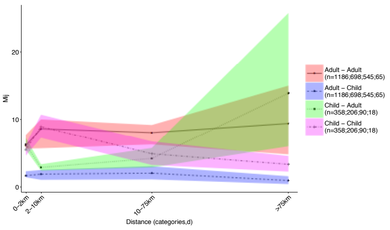

Figure 2: Lines connect the dots which represent the cells of the social contact matrix over distance $\mathbf{M}_d$ conditional upon travelling the distance with 95 % CI bands.

- The unconditional social contact matrices, yielding a population based perspective, show a decline with distance. Discrepancies exist for regular and holiday periods (not shown).

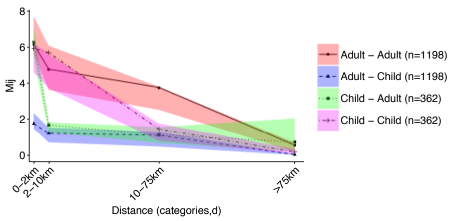

Figure 3: Lines connect the dots which represent the cells of the social contact matrix over distance $\mathbf{M}_{d*}$ unconditional upon travelling the distance with 95 % CI bands.

Summary

This study shows that

1. the power law behaviour observed in disease spread is not reflected by the contact degree distribution over distance,
2. considering constant contact matrices across distance fails to capture the dispersal of infection across distances.

Further research includes investigating gender-related differences; integrating these results in different epidemic models.

Acknowledgments

This project was funded by the European Research Council (ERC) under the European Unions Horizon 2020 research and innovation programme (grant agreement 682540 - TransMID); the University of Antwerp scientific chair in Evidence-Based Vaccinology and the PHC Tournesol Flanders program no.35686NE.

UHASSELT · I-BioStat Interuniversity Institute for Biostatistics and statistical Bioinformatics · KU LEUVEN · erc European Research Council · TransMID · European Commission · Health Economics & Modelling Infectious Diseases, Vaccine & Infectious Disease Institute, University of Antwerp