



British Ecological Society

---

Comparison of Social Networks Derived from Ecological Data: Implications for Inferring Infectious Disease Dynamics

Author(s): Sarah E. Perkins, Francesca Cagnacci, Anna Stradiotto, Daniele Arnoldi and Peter J. Hudson

Source: *Journal of Animal Ecology*, Vol. 78, No. 5 (Sep., 2009), pp. 1015-1022

Published by: [British Ecological Society](#)

Stable URL: <http://www.jstor.org/stable/40405837>

Accessed: 24-04-2015 19:14 UTC

## REFERENCES

Linked references are available on JSTOR for this article:

[http://www.jstor.org/stable/40405837?seq=1&cid=pdf-reference#references\\_tab\\_contents](http://www.jstor.org/stable/40405837?seq=1&cid=pdf-reference#references_tab_contents)

You may need to log in to JSTOR to access the linked references.

---

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at

<http://www.jstor.org/page/info/about/policies/terms.jsp>

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.



British Ecological Society is collaborating with JSTOR to digitize, preserve and extend access to *Journal of Animal Ecology*.

<http://www.jstor.org>

# Comparison of social networks derived from ecological data: implications for inferring infectious disease dynamics

Sarah E. Perkins<sup>1\*</sup>, Francesca Cagnacci<sup>2</sup>, Anna Stradiotto<sup>2</sup>, Daniele Arnoldi<sup>2</sup> and Peter J. Hudson<sup>1</sup>

<sup>1</sup>Center for Infectious Disease Dynamics, 208 Mueller Laboratory, Penn State University, State College, PA 16803, USA; and <sup>2</sup>Edmund Mach Foundation, Research and Innovation Centre, Via Mach 1, 38010 San Michele all'Adige, Trentino, Italy

## Summary

1. Social network analyses tend to focus on human interactions. However, there is a burgeoning interest in applying graph theory to ecological data from animal populations. Here we show how radio-tracking and capture–mark–recapture data collated from wild rodent populations can be used to generate contact networks.
2. Both radio-tracking and capture–mark–recapture were undertaken simultaneously. Contact networks were derived and the following statistics estimated: mean-contact rate, edge distribution, connectance and centrality.
3. Capture–mark–recapture networks produced more informative and complete networks when the rodent density was high and radio-tracking produced more informative networks when the density was low. Different data collection methods provide more data when certain ecological characteristics of the population prevail.
4. Both sets of data produced networks with comparable edge (contact) distributions that were best described by a negative binomial distribution. Connectance and closeness were statistically different between the two data sets. Only betweenness was comparable. The differences between the networks have important consequences for the transmission of infectious diseases. Care should be taken when extrapolating social networks to transmission networks for inferring disease dynamics.

**Key-words:** capture–mark–recapture, parasites, radio-tracking, social network analysis, transmission

## Introduction

In the study of infectious disease dynamics, the most critical and intractable parameter that needs to be estimated for the prediction of disease dynamics is the transmission coefficient beta (McCallum, Barlow & Hone 2001). Transmission events can rarely be recorded directly and yet models are often highly sensitive to variation in transmission and how it scales from local interactions to population-level patterns. Transmission is a combination of both the likelihood of a host coming into contact with the infectious agent (which may equate to the infectious host) and the likelihood of the susceptible host becoming infected. Researchers attempt to estimate this by fitting models to age infection data or longitudinal data and then estimating the basic reproduction number  $R_0$  (Anderson & May 1991; Hudson & Dobson

1995). Transmission is often considered frequency dependent at the individual local level but density-dependent at the population level (Tompkins & Begon 1999; McCallum *et al.* 2001). A further complication is that the minority of individuals within a population are highly connected and consequently more likely to be involved in transmission events (Anderson & May 1991; Lloyd-Smith, Getz & Westerhoff 2004; Lloyd-Smith *et al.* 2005).

Heterogeneities in individual susceptibility and transmission, particularly when they covary, have been shown to have a significant effect on the temporal dynamics of infection and substantially increase the basic reproduction number,  $R_0$  (Anderson & May 1991; Keeling *et al.* 2003; Lloyd-Smith *et al.* 2005; May 2006; Smith *et al.* 2007). In such instances, models that use individual-based contact rates can be more appropriate and provide a better explanation than mean field models (Bansal, Grenfell & Meyers 2007). One way to capture the characteristics of these heterogeneities is to use social

\*Correspondence author. E-mail: sep18@psu.edu

network analyses, and these have been applied to frequency-dependent infections such as sexually transmitted pathogens in humans (Klov Dahl 1985; Anderson & May 1991). One method for recording potential transmission events is through contact tracing (Eames & Keeling 2002; Meyers, Newman & Pourbohloul 2006). The disadvantage of contact tracing is that the data may not be a true reflection of the actual process as it can rely on memory and truth of the infected individuals. This approach has worked in small outbreaks but even then workers can rarely record who was exposed and not infected (e.g. SARS). Within ecology, there exists a wealth of data on the identity of free-living animals and their potential social contacts that could be used to construct social networks and subsequently infer disease dynamics, as has been done for human pathogens (Krause, Croft & James 2007).

In population and behavioural ecology, social interactions are recorded by monitoring the specific location of individual animals and these data can allow us to estimate individual contact patterns. Radio telemetry has been used to infer social interactions by estimating patterns of spatial overlap (e.g. home range, Doncaster 1990; Minta 1992; Fieberg & Kochanny 2005). A good example of how this relates to disease ecology has been the study of the ecology of bovine tuberculosis in European badgers *Meles meles*. The telemetry data have shown that social interactions between groups increases after culling and this probably increases transmission events between groups (Tuytens *et al.* 2000) and also the risk of pathogen spill-over into cattle (Vicente *et al.* 2007). A second method used to record individual location is the capture-mark-recapture (CMR) technique (e.g. Otis *et al.* 1978; Nichols, Pollock & Hines 1984). An application of these data within disease ecology is provided by Carslake *et al.* (2005) who inferred mixing patterns in rodent populations and determined that localized transmission was important in the population-level dynamics of cowpox transmission. Additional methods are emerging specifically to record and describe social structure and individual contacts including passive-induced transponder (PIT) tags and proximity loggers (e.g. Garnett, Delahay & Roper 2002; Corner, Pfeiffer & Morris 2003; Ji, White & Clout 2005). These data have proved useful in inferring disease dynamics, particularly in populations where there is a high variance in contact rate between individuals (Porteous & Pankhurst 1998; Ji *et al.* 2005), although few studies have applied formal social network analyses (but see Cross *et al.* 2004). One obstacle limiting the use of these data to infer disease dynamics is the challenge in relating the data to the likelihood of transmission and producing meaningful transmission networks from social networks.

In this paper, we examine how ecological data can be used to estimate social networks in a small mammal population and, using network analyses, how this can provide insights into the transmission of infectious diseases. Specifically, we compare and contrast the properties of social networks derived from a rodent population that we simultaneously monitored using both radio-tracking and CMR over a 2-year

period during a change in density. We posit that the temporal resolution of the data and its utility for inferring transmission dynamics depends upon the infectious period of the parasite or pathogen. In particular, we propose that the dynamics of parasites with short infectious periods are better captured with data of high temporal resolution, such as those collected by radio-tracking, whilst CMR data provides an adequate network for understanding the dynamics of parasites with longer infectious periods, or for those that have a free-living stage, for example certain nematode species.

## Materials and methods

### STUDY AREA AND LIVE-TRAPPING

Fieldwork was carried out in the north-eastern Italian Alps (10°57'47" E 45°58'50" N), where *Apodemus flavicollis*, yellow-necked mouse, was the predominant rodent species. From July to October 2005 and 2006, we carried out a CMR-trapping session over five nights, every 3 weeks (2005: 8–13 July; 7–12 August; 4–9 September; 3–9 October and 2006: 16–21 July; 13–18 August; 10–15 September; 8–13 October). Live-trapping occurred on a large trapping grid, of square lattice design, with 18 by 18 traps (324 traps in total). Traps used were multiple capture live-traps (Special Mouse 2; Ugglan, Gråhns, Sweden) and set with a 15-m inter-trap interval such that the entire trapping grid covered 6.5 ha. The location of each trap was accurately recorded and mapped using a Global Positioning System (GPS; GeoExplorer3 Trimble; Crisel, Roma, Italy) using post-processed differential correction (accuracy of 1–5 m).

The trapping design was decided according to previous estimations of optimal population sampling. We chose 5 days of consecutive trapping as this has been recommended for a robust and precise estimation of population size, survival rate, recruitment and for good performance of closed population models (Otis *et al.* 1978; Nichols *et al.* 1984). Trapping grid size has been recommended to include a grid of  $r$  (number of rows) and  $c$  (number of columns) between  $9 \times 9$  and  $15 \times 15$  (Otis *et al.* 1978), while White *et al.* (1982) suggested  $r + c > 25$  and Jones *et al.* (1996) a square grid of at least  $10 \times 10$  traps. The size of grid required to achieve a representative network has not been previously investigated and, with no a priori information available, we used a trapping grid larger than those recommended, measuring  $18 \times 18$  live-traps.

Each individual in the study was identified with a PIT tag (Trovan ID 100; Ghislandi & Ghislandi, Covo, Italy) and subsequent captures provided data for estimating population density using a closed population model in the program Capture (White *et al.* 1978).

### RADIO-TRACKING

During live-trapping, we fitted resident adults with VHF radio-transmitters (BD-2C; Holohil System Ltd., Carp, ON, Canada). Adults were defined as those with brown pelage and a mass of at least 29 g (Flowerdew 1984). To focus on residents and avoid transient mice, we radio-collared individuals only if they were trapped a minimum of three times in more than one live-trapping session (Rajski-Jurgiel 2001). The transmitter batteries had an average life span of 53 days and, when batteries were getting closer to the end of their life expectancy, we attempted to re-trap individuals to fit fresh radio-collars. In both years of the study, we followed a comparable number of animals with 20 individual males and 12 females radio-collared in 2005

(July: 11 mice; August: 15; September: 18; October: 10) and 19 males and 13 females in 2006 (July: 14 mice; August: 21; September: 10; October: 7). From July to October of both years, we completed four radio-tracking sessions in the period between the live-trapping sessions (2005: 14 July–5 August; 13 August–2 September; 10 September–1 October; 10 October–31 October and 2006: 22 July–11 August; 19 August–8 September; 16 September–6 October; 14 October–7 November), aiming at a sample of 50 fixes per animal per session. The location of each rodent was determined by ‘homing-in’, which involved following the signal’s increasing strength until the animal was observed or located within a small area (3.5 m radius) as determined by circling the area. Animal movements were recorded from dusk to dawn, with time intervals of 50 min or more between successive fixes, as this was considered sufficiently short to follow movements of each mouse (Wolton 1985), but long enough to avoid autocorrelation of the data (White & Garrott 1990).

## DEFINING CONTACTS

Social interactions between rodents are notably plastic, depending upon the habitat, and individuals tend to exhibit overlapping ranges with tolerance of closely related individuals (Wolton & Flowerdew 1985). As such, we have defined a social contact between two individuals to have occurred when they were observed in the same spatial location within the same live-trapping/radio-tracking time period, whilst a transmission contact depends upon the specifics of the parasite or pathogen, such as the infectious period. It is worth noting that depending upon the parasite or pathogen, a transmission network and a social network are not necessarily mutually exclusive. For live-trapping data, individual rodents that were caught in the same trap (traps are multiple capture) or in a directly adjacent trap (traps have a 15-m inter-trap interval) during the same trapping session were defined as a social contact. Individuals caught on the outermost traps of the grid were likely to have contacts with unknown individuals and so were excluded from the analyses.

For the purposes of constructing transmission networks, we have assumed a hypothetical pathogen/parasite with an infectious period of a few days and an infective stage that persists in the environment for the same time period and so can be transmitted by space-sharing. This definition ensured that the social network was analogous with a transmission network. Although we did not specify a particular infectious agent we propose that faecal–orally transmitted parasites, common in wild rodents, could be transmitted on these networks. Examples include the free-living directly transmitted helminths, such as *Heligmosomoides polygyrus*, *Syphacia* spp. and faecal–oral transmitted pathogens such as *Salmonella* and *Listeria*. Rodents deposit faeces throughout their home range and susceptible individuals that overlap in space use have a high probability of contacting the infective stages that develop from an infected individual’s faeces; and as such are potential transmission contacts (Randolph 1973, 1977).

For the radio-tracking data, we defined a contact as any individual observed within a 15-m radius of another during a radio-tracking session and so provided a direct comparison with the live-trapping data. Only fixes recorded within the trapping grid were used, and again fixes observed on the boundary of the grid (i.e. 18th trap) were excluded.

## SOCIAL NETWORK ANALYSIS

Contact matrices were constructed for each of four observation sessions (July, August, September and October) over 2 years (2005 and 2006) using both live-trapping and radio-tracking data. We produced

a set of adjacency matrices where in a network of  $n$  nodes, the adjacency matrix is an  $n \times n$  matrix with binary entries indicating whether there is a contact between two nodes ( $i, j$ ) and a node’s degree ( $k$ ) is the number of edges (or contacts). Each network consisted of nodes (individual mice) that were connected by edges, defined by space-sharing, to one or more other individuals. The networks were non-directed networks such that the adjacency matrix was symmetric, i.e. contact  $i, j$  is equal to  $j, i$ .

The edge distribution (PCK) provided a point of comparison between networks estimated from the live-trapping data with those of radio-tracking over time (July–October 2005 and 2006). The edge distribution is a frequency distribution of contacts and is an aggregate statistic that is a common descriptor of networks, particularly useful in disease dynamics (Newman 2003; Bansal *et al.* 2007). For each network, we determined using maximum likelihood estimation whether a Poisson or negative-binomial distribution better fit the edge distribution (the latter being shown to be a common distribution of contacts for transmission networks, Lloyd-Smith *et al.* 2005). The use of networks within animal ecology is relatively new and methods used to quantitatively compare networks are beginning to be addressed in the ecology literature (Krause *et al.* 2007; Croft, James & Krause 2008). One statistical issue is that the data points (nodes) are not independent, therefore linear models cannot be used to compare distributions at the node level as this violates the underlying assumption of independent data. Randomization techniques including matrix tests, such as the Mantel test, have been used to compare networks as these do not assume independent data points (see Croft *et al.* 2008 for examples). However, the sample sizes must be equal. The radio-tracking networks were smaller than that of the live-trapping, therefore we carried out a qualitative comparison of the networks using quantile–quantile plots (q–q plots).

The q–q plot is a graphical technique that provides a visual goodness-of-fit for determining if the two data sets (networks) have a common distribution. If the degree distributions of the two networks are the same, the points should be approximately linear along a 45° reference line and have no curvature. The greater the departure from this reference line, the greater the evidence that the two networks have different distributions.

The topology (i.e. geometric structure) of the network is critical in affecting the pattern of parasite spread (Strogatz 2001). Two networks with the same edge distribution can have very different network properties. Therefore, we used social network statistics to examine differences in network topology between the two data-collection methods. We measured four different network properties, relevant to disease dynamics.

1. *The average contact rate ( $c$ )*, is the mean square of the degree distribution divided by the mean [ $c = \langle k^2 \rangle / \langle k \rangle$ ], after May (2006). We also approximated the *basic reproduction number* ( $R_0$ ) following May (2006), where  $R_0$  is a measure of the contact rate of the population, coupled with the duration of infectiousness of the pathogen. Here,  $R_0 = \rho_0 [1 + (CV)^2]$ , where CV is the coefficient of variation of the degree distribution and  $\rho_0$  is the transmission probability ( $\beta$ ) multiplied by the duration of infectiousness ( $D$ ). To compare between the networks, we assumed that  $\rho_0$  was a constant and so  $R_0$  scaled as  $R_0 = 1 + (CV)^2$ .

2. *Centrality (closeness and betweenness)*. Centrality is a measure of an individual nodes’ importance in the network and indices of centralization provide insight into the heterogeneity of a network. The two centrality indices measured were closeness and betweenness.

- (i) *Closeness* is the mean geodesic distance (i.e. the shortest path) between an individual to all other individuals. Intuitively,



closeness provides an index of the extent to which an individual is in the 'middle' of a given structure. The more central the individual, the greater potential role it has in facilitating pathogen transmission, as this describes how readily infectious stages from that individual can reach all others (Corner *et al.* 2003). We computed a closeness index, which quantified closeness for the entire network. At its maximum (1.0), the index represents a 'super-spreader' network, where one central individual is connected to all others, with no connections between those others, forming a 'star network'. A low closeness index (minimum = 0.0) implies a homogenous population where all individuals are equally connected to one another in a 'circle network' (Wasserman & Faust 1998).

- (ii) *Betweenness* is a measure of the number of paths that pass through an individual along the shortest path between all other individuals. Conceptually, betweenness measures the flow of a parasite or pathogen through the network and an individual with high betweenness can be thought of as a 'fire-break' in terms of transmission. The contrast with closeness is that an individual with high betweenness does not necessarily have a high number of connections, more it can link motifs (regular patterns, such as triangles) within the network. When 'betweenness' is zero all individuals are equal, when it is 1 then one individual is the 'hub' that links to all others. Both scores were normalized to allow comparison between the networks.

Where networks were only partially connected (i.e. not all nodes are connected), it was not possible to compute a centrality index. The number of unconnected nodes accounted for a small proportion of the population, and these individuals were treated as 'transient' outliers and a centrality index was computed by excluding these individuals. These nodes were only excluded for the purposes of computing centrality indices.

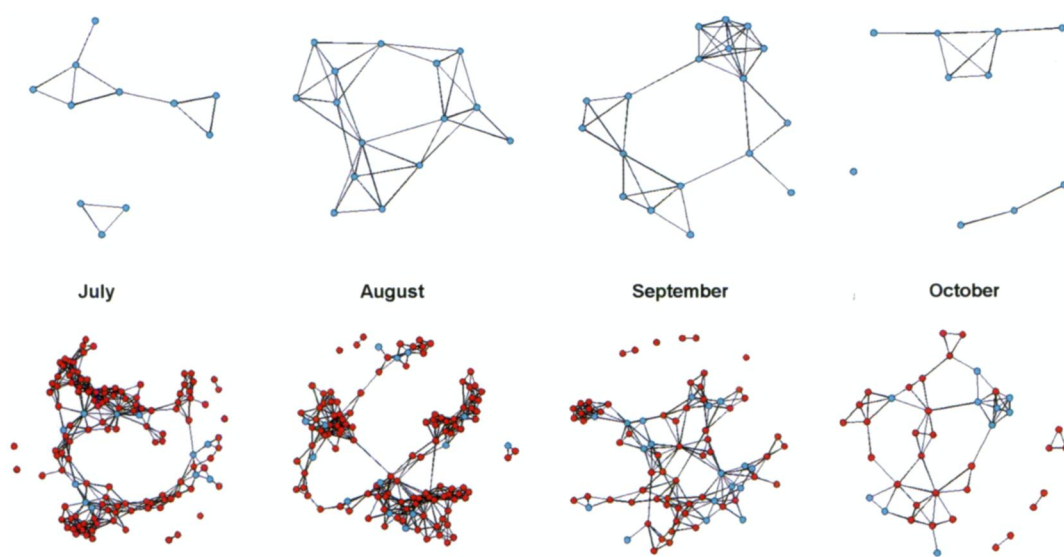
**3. Connectedness.** Connectedness scores can range from 0 (for a disconnected graph) to 1 (for connected graphs). Connectance is the proportion of all possible links that are realized within a network and represents the probability that any two individuals will interact, thereby giving an 'interaction' strength (Proulx, Promislow & Phillips 2005). In terms of disease, this represents how fast an epidemic would

move through the population, assuming that the network is unchanged by an epidemic.

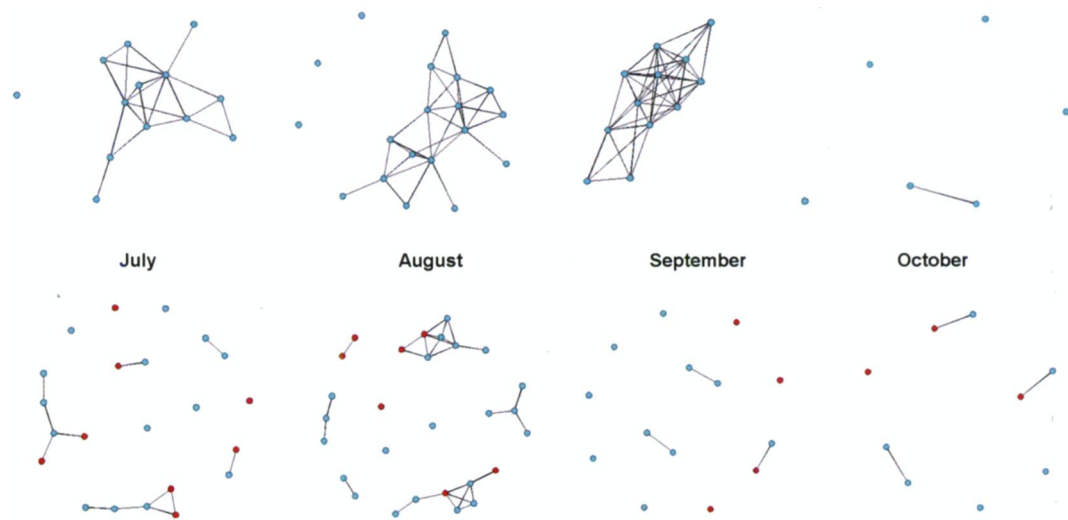
We tested for statistically significant differences in the properties of the network by using betweenness, connectedness and closeness separately as response variables in a generalized linear model with an appropriate error distribution. The size of the networks varied over time, so the response variables were weighted by the number of nodes (or individuals) per network. The data-collection method (radio-tracking or CMR) and host density were used as explanatory variables as was the two-way interaction between these variables. We used backwards step-wise deletion to determine the minimal adequate model. Each network graph was drawn using the Fruchterman-Reingold layout (Fruchterman & Reingold 1991). This sorts nodes into a desirable layout for presentation purposes and does not imply spatial location of an individual. All analyses were undertaken in R (2007), with social network analyses carried out using Statnet and sna (Handcock *et al.* 2003; Butts 2007).

## Results

The radio-tracking data generated smaller networks with fewer nodes than CMR (Figs 1 and 2, Table 1). However, other network properties were density dependent with the radio-tracking networks from 2005 (high density) having a lower mean-contact rate than those generated from CMR data (Fig. 2), but the opposite was true when density was low in 2005 (Fig. 1). This pattern was also evident in the deviation of the data from a 1 : 1 reference line in q-q plots, where in 2005 the data consistently fell below the reference line, meaning the number of contacts observed were much higher using CMR data than radio-tracking data (Fig. 3). This was especially pronounced for the months of July and August. In 2006, the pattern was reversed and a higher number of contacts were observed in radio-tracking networks for all



**Fig. 1.** Contact networks for a wild rodent population, over time, for the year 2005, showing social contacts (lines) between individuals (circles), which were recorded by either radio-tracking (top row) or capture-mark-recapture in static traps (bottom row). In the capture-mark-recapture derived networks the animals that were radio-tracked are shown as blue nodes.



**Fig. 2.** Contact networks for a wild rodent population, over time, for the year 2006, showing contacts (lines) between individuals (circles), which were recorded by either radio-tracking (top row) or capture-mark-recapture in static traps (bottom row). In the capture-mark-recapture-derived networks, the animals that were radio-tracked are shown as blue nodes. Note that the density was exceptionally low in comparison to 2005 (Fig. 1), thus changing the network considerably.

**Table 1.** Network properties of contact networks derived from a rodent population, over time, for the years 2005 and 2006

Month	Year	Mean number of contacts		Basic reproduction number ( $R_0$ )		Connectedness		Closeness		Betweenness		Size (number of nodes)	
		CMR	RT	CMR	RT	CMR	RT	CMR	RT	CMR	RT	CMR	RT
July	2005	7.15	2.36	13.7	2.40	0.88	0.56	0.07	0.19	0.22	0.22	131	11
August	2005	7.29	4.67	14.29	6.44	0.91	1.00	0.12	0.22	0.47	0.27	127	15
September	2005	5.20	4.44	7.76	5.90	0.86	1.00	0.13	0.11	0.21	0.24	80	18
October	2005	3.50	2.00	4.06	2.00	0.68	0.40	0.14	0.22	0.25	0.10	40	10
July	2006	1.09	3.38	1.30	3.86	0.10	0.85	0	0.23	0.02	0.25	22	13
August	2006	2.00	3.47	2.00	4.02	0.14	0.70	0	0.18	0.02	0.22	28	19
September	2006	0.43	6.62	1.05	11.94	0.03	0.85	0	0.20	0	0.12	14	13
October	2006	0.67	0.4	1.11	1.04	0.08	0.10	0	0	0	0	19	5

Networks were derived from either radio-tracking (RT) data or capture-mark-recapture (CMR) data.

months, with the exception of October where data were too sparse to make useful conclusions.

The distribution of contacts was best fit by a negative binomial probability distribution, although for the radio-tracking data in July 2005 and October 2005 and the CMR data from September and October 2006 neither distribution fitted, possibly due to small sample sizes. The similarity of the degree distributions between the methods was further corroborated by the q-q plots where the data fitted a straight line, implying the two networks had similar degree distributions with the exception of October 2006 where the data were too sparse to yield conclusive results (Fig. 3).

The differences between the networks in terms of their aggregate statistics allowed us to carry out a direct comparison between the methods (Table 1). Betweenness did not differ significantly between the radio-tracking and CMR networks ( $\chi^2_{14} = 1.32$ ;  $P = 0.25$ ). However, there was a sig-

nificant increase in betweenness with increasing rodent density ( $\chi^2_{13} = 27.13$ ;  $P < 0.01$ ). Closeness was significantly higher in radio-tracking networks than CMR networks ( $\chi^2_{14} = 6.63$ ;  $P = 0.01$ ), but unaffected by rodent density ( $\chi^2_{13} = 0.65$ ;  $P = 0.42$ ). Finally connectedness increased significantly with increasing host density ( $\chi^2_{13} = 135.01$ ;  $P < 0.01$ ) and was significantly higher in radio-tracking networks compared to CMR ones ( $\chi^2_{13} = 15.79$ ;  $P < 0.01$ ).

## Discussion

We examined the properties of contact networks produced from observational data on the same rodent population with the same transmission contact assumptions and found that they differed significantly according to the data-collection method (Figs 1 and 2, Table 1). The degree distributions were best represented by a negative binomial distribution and the distribution of contacts was not qualitatively different

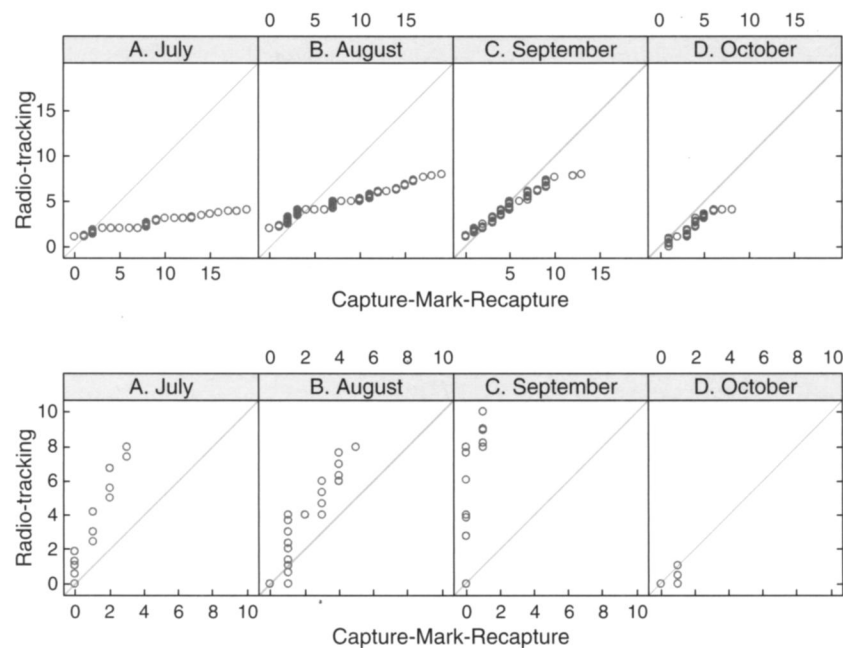


Fig. 3. Quantile–quantile plots comparing the degree (contact) distributions derived from capture–mark–recapture and radio-tracking networks from July to October 2005 (upper panel) and July to October 2006 (lower panel). A 1 : 1 reference line is plotted.

between radio-tracking and CMR data sets (Fig. 3). However, the mean number of contacts differed intra- and inter-annually. During 2005, the mean contacts as estimated using CMR data were considerably higher than those estimated from radio-tracking, whilst in 2006 the reverse was true and the mean number of contacts would be underestimated if only CMR networks were observed (Table 1, Fig. 3).

The disparity in the mean number of contacts between the data-collection methods were likely a function of the fluctuating rodent density between and within years. In 2005, the rodent density was comparatively high with an annual mean density of 9.2 mice  $\text{ha}^{-1}$ , compared with a lower mean density in 2006 of 1.7 mice  $\text{ha}^{-1}$ . The low mean-contact rates recorded in 2005 from radio-tracking was probably due to the relatively small proportion of the population being monitored in comparison with CMR methods. However, when rodent density was low, in 2006, the low capture success of live-traps under these ecological conditions (Mihok, Lawton & Schwartz 1988) resulted in comparable numbers of animals being monitored, and the increased temporal sampling of radio-tracking over CMR enhanced the probability of recording rare contacts. When either spatial or temporal sampling is low, very different social networks can be produced from the same data set (Cross *et al.* 2004). Given also that the degree distributions follow a negative-binomial distribution when the population is small, the estimation of contacts may be subject to a sampling issue associated with negative-binomial distributions, whereby the tail of the distribution is under-sampled (Wilson *et al.* 2002).

The observed number of contacts allowed us to estimate a value proportional to  $R_0$  (Table 1). As with contact rate, the  $R_0$  estimates were comparably higher for capture–mark–

recapture networks than radio-tracking when density was high. Conversely, the estimates were higher for radio-tracking than capture–mark–recapture networks when density was low. However, it is worth noting that the mean-contact rate and  $R_0$  estimate did not scale linearly between the two methods (Table 1). In particular, the estimates from July 2005 were disproportionately higher for CMR than radio-tracking and, in September 2006, were disproportionately higher for radio-tracking than CMR (Table 1). The differences in contact rates and  $R_0$  estimates have obvious implications for inferring disease dynamics; both methods would lead to very different conclusions. This observation is corroborated when we contrast the network statistics from each data-collection method (Table 1).

A parasite or pathogen could be expected to spread quickly in a population with high betweenness and whilst that was roughly equivalent in both networks when density was high, the speed of transmission (measured as betweenness) was close to zero for the CMR networks in the low density year (Table 1). Although the differences in betweenness did not differ statistically between the methods, it was positively associated with rodent density, which may be due to better resolution of the network with the increased sampling at high host density. A high closeness index signifies a propensity for ‘super-spreading’ events to occur because contacts within the network are heterogeneous. As such, where the index is high, an infection starting in a central individual could potentially infect a large proportion of individuals. In contrast, an infection in a network with a low closeness index would be unlikely to spread as extensively through the population. In general, both data-collection methods produced rather low closeness indices, indicating that there was low heterogeneity

in the networks. However, the indices were significantly different from each other, with closeness being significantly higher in the radio-tracking networks, indicating a higher level of 'super-spreading' especially when rodent density was low. This leads to different interpretations about the degree of super-spreading and ultimately to the application of optimal disease control. We observed a similar pattern with connectance, which measures the fraction of all possible links that are realized in a network. Connectance was significantly higher for radio-tracking networks, especially when rodent density was low, whereas the CMR data suggested very low connectance. This suggests that the CMR networks in the low density year were not well resolved compared to the radio-tracking networks. Overall, the differences in the centrality indices were much lower for CMR over radio-tracking networks and were especially pronounced when rodent density was low in 2006 because the CMR networks were not well resolved.

Ecological attributes such as sex, body mass and reproductive status are known to have important consequences for disease dynamics (Perkins *et al.* 2003). These data were recorded in our experiment, but they were not included in these analyses. Few network analyses have examined the interactions between different groups of the population and this along with additional methodological analyses of network sampling issues is a focus of future research for this group.

This work has demonstrated an important and well-cited point in network ecology – that the network structure is only an approximation of the contacts. Defining a true network would require collecting data on every contact between all individuals, a difficult if not an impossible task in ecological systems. This refers to a sampling issue that is prevalent throughout ecology. We rarely, if ever, sample the entire population, only a subsample. Addressing this issue in the context of collecting a representative network, not only is the size of the network, but also the methods by which the data are collected are important. Radio-tracking data generated smaller networks (fewer nodes) than CMR because the number of animals monitored was fewer, but when comparable numbers of animals were monitored radio-tracking furnished significantly more information than CMR, especially in terms of centrality indices. Therefore, the selection of the data used to approximate the network is extremely important. The two data-collection approaches compared in this paper should also be evaluated as a trade-off between costs and benefits in terms of data-collection effort. The live-trapping required *c.* 15 people day<sup>-1</sup> session<sup>-1</sup>, whilst radio-tracking required 45 people day<sup>-1</sup> session<sup>-1</sup>. In terms of material, radio-tracking involved purchasing of radio-transmitters and relied on live-trapping for collaring animals.

From this work, it is difficult to conclude that one network is a better representative than the other because we do not actually know what the real network looks like, but certainly it seems that the radio-tracking networks are 'better resolved' in terms of examining pathogens and parasites with short infectious periods (< 5 days) while, for infectious periods

> 1 week, CMR provides more useful data. When rodent density is low, the low mean-contact rate and centrality indices derived from CMR networks may suggest slow transmission of pathogens and parasites. Conversely, under the same conditions, the high contact rate and connectedness revealed from radio-tracking data suggest rapid pathogen transmission. Whilst social contacts can be inferred from either data-collection method, the transmission contacts are properly documented only by using data at an appropriate temporal resolution with respect to the infectious period of a specific parasite or pathogen. This brings us to an important point that there should be a clear distinction between social and transmission networks and the data collected for producing transmission networks should be tailored to capturing the dynamics of the particular parasite or pathogen. We conclude that data collated at high temporal resolution, such as radio-tracking, may be better suited to understanding the dynamics of short infectious-period pathogens, whilst the more temporally sparse data provided by CMR better suits pathogens and parasites with longer infectious periods.

## Acknowledgements

This work was funded by NSF grant EF-0520468 as part of the joint NSF-NIH Ecology of Infectious Disease program. FC was supported by the Autonomous Province of Trento under Grant N. 3479 (BECOCERWI – Behavioural Ecology of Cervids in Relation to Wildlife Infections). We are grateful to A. Aimi, D. Arnoldi, F. Bertola, G. Bocedi, V. Gonzalez, M. Peli and Y. Valerio for their contribution to fieldwork. G. Carpi provided veterinary assistance and S. Tioli data manipulation. We also thank Dave Hunter and David Welch for statistical advice and comments, which improved this paper. All animal-handling procedures were carried out in accordance with the protocols approved by the Scientific Committee of the Research Fund of the Autonomous Province of Trento.

## References

- Anderson, R.M. & May, R.M. (1991) *Infectious Diseases of Humans. Dynamics and Control*. Oxford University Press, New York.
- Bansal, S., Grenfell, B.T. & Meyers, L.A. (2007) When individual behaviour matters: homogeneous and network models in epidemiology. *Journal of the Royal Society Interface*, **4**, 879–891.
- Butts, C.T. (2007) *SNA: Tools for Social Network Analysis*, R package version 1.5. <http://erzuli.ss.uci.edu/R.stuff>.
- Carslake, D., Bennett, M., Bown, K., Hazel, S., Telfer, S. & Begon, M. (2005) Space-time clustering of cowpox virus infection in wild rodent populations. *Journal of Animal Ecology*, **74**, 647–655.
- Corner, L.A.L., Pfeiffer, D.U. & Morris, R.S. (2003) Social network analysis of *Mycobacterium bovis* transmission among captive brushtail possums (*Trichosurus vulpecula*). *Preventive Veterinary Medicine*, **59**, 147–167.
- Croft, D.P., James, R. & Krause, J. (2008) *Exploring Animal Social Networks*. Princeton University Press, Princeton, NJ, USA.
- Cross, P.C., Lloyd-Smith, J.O., Bowers, J.A., Hay, C.T., Hofmeyr, M. & Getz, W.M. (2004) Integrating association data and disease dynamics in a social ungulate: bovine tuberculosis in African buffalo. *Annales Zoologici Fennici*, **41**, 879–892.
- Doncaster, C.P. (1990) Non-parametric estimates of interaction from radio-tracking data. *Journal of Theoretical Biology*, **143**, 431–443.
- Eames, K.T.D. & Keeling, M.J. (2002) Modeling dynamic and network heterogeneities in the spread of sexually transmitted diseases. *Proceedings of the National Academy of Science of the United States of America*, **99**, 13330–13335.



- Fieberg, J. & Kochanny, C.O. (2005) Quantifying home-range overlap: the importance of the utilization distribution. *Journal of Wildlife Management*, **69**, 1346–1359.
- Flowerdew, J. (1984) *Woodmice and Yellow-Necked Mice*. Mammal Society, London, UK.
- Fruchterman, T.M.J. & Reingold, E.M. (1991) Graph drawing by force-directed placement. *Software – Practice and Experience*, **21**, 1129–1164.
- Garnett, B.T., Delahay, R.J. & Roper, T.J. (2002) Use of cattle farm resources by badgers (*Meles meles*) and risk of bovine tuberculosis (*Mycobacterium bovis*) transmission to cattle. *Proceedings of the Royal Society of London. Series B, Biological Sciences*, **269**, 1487–1491.
- Handcock, M.S., Hunter, D.R., Butts, C.T., Goodreau, S.M. & Morris, M. (2003) *Statnet: Software tools for the statistical modelling of network data*. <http://statnet.project.org>
- Hudson, P.J. & Dobson, A.P. (1995) Macroparasite: observed patterns. *Ecology of Infectious Disease in Natural Populations* (eds B.T. Grenfell & A.P. Dobson), pp. 144–176. Cambridge University Press, Cambridge.
- Ji, W., White, P.C.L. & Clout, M.N. (2005) Contact rates between possums revealed by novel proximity data loggers. *Journal of Applied Ecology*, **42**, 595–604.
- Jones, C., McShea, M.W., Conroy, M.J. & Kunz, T.H. (1996) Capturing mammals. *Measuring and Monitoring Biological Diversity. Standard Methods for Mammals* (eds D.E. Wilson, F.R. Cole, J.D. Nichols, R. Rudran & M.S. Foster), pp. 115–155. Smithsonian Institution Press, Washington and London.
- Keeling, M.J., Woolhouse, M.E.J., May, R.M., Davies, G. & Grenfell, B.T. (2003) Modelling vaccination strategies against foot-and-mouth disease. *Nature*, **421**, 136–142.
- Klov Dahl, A.S. (1985) Social networks and the spread of infectious diseases: the AIDS example. *Social Science and Medicine*, **21**, 1203–1216.
- Krause, J., Croft, D.P. & James, R. (2007) Social network theory in the behavioural sciences: potential applications. *Behavioral Ecology and Sociobiology*, **62**, 15–27.
- Lloyd-Smith, J.O., Getz, W.M. & Westerhoff, H.V. (2004) Frequency-dependent incidence in models of sexually transmitted diseases: portrayal of pair-based transmission and effects of illness on contact behaviour. *Proceedings of the Royal Society of London. Series B, Biological Sciences*, **271**, 625–634.
- Lloyd-Smith, J.O., Schreiber, S.J., Kopp, P.E. & Getz, W.M. (2005) Super-spreading and the impact of individual variation on disease emergence. *Nature*, **438**, 355–359.
- May, R.M. (2006) Network structure and the biology of populations. *Trends in Ecology and Evolution*, **21**, 394–399.
- McCallum, H., Barlow, N. & Hone, J. (2001) How should pathogen transmission be modelled? *Trends in Ecology & Evolution*, **16**, 295–300.
- Meyers, L.A., Newman, M.E.J. & Pourbohloul, B. (2006) Predicting epidemics on directed contact networks. *Journal of Theoretical Biology*, **240**, 400–418.
- Mihok, S., Lawton, T. & Schwartz, B. (1988) Fates and movements of meadow voles (*Microtus pennsylvanicus*) following a population decline. *Canadian Journal of Zoology*, **66**, 323–328.
- Minta, S.C. (1992) Test of spatial and temporal interaction among animals. *Ecological Applications*, **2**, 178–188.
- Newman, M.E.J. (2003) The structure and function of complex networks. *SIAM Review*, **45**, 167–256.
- Nichols, J.D., Pollock, K.H. & Hines, J.E. (1984) The use of a robust capture-recapture design in small mammal population studies: a field example with *Microtus pennsylvanicus*. *Acta Theriologica*, **29**, 357–365.
- Otis, D.L., Burnham, K.P., White, G.C. & Anderson, D.R. (1978) Statistical inference from capture data on closed animal populations. *Wildlife Monographs*, **62**, 1–135.
- Perkins, S.E., Cattadori, I.M., Tagliapietra, V., Rizzoli, A.P. & Hudson, P.J. (2003) Empirical evidence for key hosts in persistence of a tick-borne disease. *International Journal for Parasitology*, **33**, 909–917.
- Porteous, I.S. & Pankhurst, S.J. (1998) Social structure of the mara (*Dolichotis patagonum*) as a determinant of gastro-intestinal parasitism. *Parasitology*, **116**, 269–275.
- Proulx, S.R., Promislow, D.E.L. & Phillips, P.C. (2005) Network thinking in ecology and evolution. *Trends in Ecology and Evolution*, **20**, 345–353.
- R (2007) *R: A Language and Environment for Statistical Computing*, Version 2.6.1. R Foundation for Statistical Computing, Vienna, Austria.
- Rajska-Jurgiel, E. (2001) Movement behaviour of woodland rodents: looking from beyond small trapping grids. *Acta Theriologica*, **46**, 145–159.
- Randolph, S.E. (1973) Tracking technique for comparing individual home ranges of small mammals. *Journal of Zoology*, **170**, 509–520.
- Randolph, S.E. (1977) Changing spatial relationships in a population of *Apodemus sylvaticus* with onset of breeding. *Journal of Animal Ecology*, **46**, 653–676.
- Smith, D.L., McKenzie, F.E., Snow, R.W. & Hay, S.I. (2007) Revisiting the basic reproductive number for malaria and its implications for malaria control. *PLOS Biology*, **5**, 531–542.
- Strogatz, S.H. (2001) Exploring complex networks. *Nature*, **410**, 268–276.
- Tompkins, D.M. & Begon, M. (1999) Parasites can regulate wildlife population. *Parasitology Today*, **15**, 311–313.
- Tuytens, F.A.M., Delahay, R.J., MacDonald, D.W., Cheeseman, C.L., Long, B. & Donnelly, C.A. (2000) Spatial perturbation caused by a badger (*Meles meles*) culling operation: implications for the function of territoriality and the control of bovine tuberculosis (*Mycobacterium bovis*). *Journal of Animal Ecology*, **69**, 815–828.
- Vicente, J., Delahay, R.J., Walker, N.J. & Cheeseman, C.L. (2007) Social organization and movement influence the incidence of bovine tuberculosis in an undisturbed high-density badger *Meles meles* population. *Journal of Animal Ecology*, **76**, 348–360.
- Wasserman, S. & Faust, K. (1998) *Social Network Analysis: Methods and Applications*. Cambridge University Press, Cambridge.
- White, G.C. & Garrott, R.A. (1990) *Analysis of Wildlife Radio-Tracking Data*. Academic Press, San Diego, CA, USA.
- White, G.C., Anderson, D.R., Burnham, K.P. & Otis, D.L. (1982) Capture-mark-recapture and removal methods for sampling closed populations. Los Alamos National Laboratory Rep. LA-8787-NERP, Los Alamos, New Mexico, USA, p. 235.
- White, G.C., Burnham, K.P., Otis, D.L. & Anderson, D.R. (1978) *User's Manual for Program CAPTURE*. Utah State University Press, Logan, UT, USA.
- Wilson, K., Bjørnstad, O.N., Dobson, A.P., Merler, S., Pogliayen, G., Randolph, S.E., Read, A.F. & Skorpning, A. (2002) Heterogeneities in macroparasite infections: patterns and processes. *The Ecology of Wildlife Diseases* (eds P.J. Hudson, A. Rizzoli, B.T. Grenfell, H. Heesterbeek & A.P. Dobson), pp. 6–44. Oxford University Press, Oxford, UK.
- Wolton, R.J. (1985) The ranging and nesting behaviour of Wood mice, *Apodemus sylvaticus* (Rodentia: Muridae), as revealed by radio-tracking. *Journal of Zoology*, **206**, 203–222.
- Wolton, R.J. & Flowerdew, J.R. (1985) Spatial distribution and movements of woodmice, yellow necked mice and bank voles. *Symposium of Zoological Society of London*, **55**, 249–275.

Received 28 August 2008; accepted 6 April 2009

Handling Editor: Andy White