

# 1 The next steps in the study of missing individuals in networks: a comment on

## 2 Smith *et al.* (2017)

3 Matthew J Silk<sup>1</sup>

4 <sup>1</sup>Environment and Sustainability Institute, University of Exeter, Penryn, Cornwall, UK

5 Corresponding author address (MJS): Environment and Sustainability Institute, University of Exeter,  
6 Penryn, Cornwall, UK [matthewsilk@outlook.com](mailto:matthewsilk@outlook.com)

7

8

### 9 Abstract

10 Social network analysis is now used widely to study social behaviour in humans and non-human  
11 animals, and missing individuals can represent a problem for network studies. This problem is  
12 becoming especially frequent in studies using bio-logging to collect interaction data, especially in  
13 animals. This therefore represents an important audience for Smith *et al.* (2017) who investigate  
14 how sub-sampling from networks impacts the outcome of subsequent analysis. Here I take  
15 advantage of the progress made by this paper to outline key issues that still require addressing to  
16 understand the effect of missing individuals on social network analysis.

17 **Keywords:** network sampling, precision, bias, accuracy, statistical modelling

18

### 19 Introduction

20 As a consequence of being relational data, the sampling of networks might inherently be  
21 expected to result in greater bias than other types of data (Alba, 1982; Silk *et al.*, 2015; Smith *et al.*,  
22 2017; Smith and Moody, 2013). This could apply both to missing individuals (nodes) and missing  
23 relationships (edges), but the former is easier to quantify and address. A number of studies have  
24 explored the impact of missing individuals on network analysis, the most recent of which is Smith *et*

25 *al.* (2017). The authors made substantial progress on a number of key issues, in particular in: a)  
26 assessing how non-random missingness of individuals might change the effect of sub-sampling from  
27 a network, and b) in providing a tool to allow researchers to determine the likely impact of missing  
28 individuals in a range of network structures and sizes.

29         Smith *et al.* (2017) stated that “*By looking at a wide range of networks, measures and types*  
30 *of missing data, we can offer recommendations and best practices for applied network*  
31 *practitioners*”. A major application of network analysis away from the social sciences is in the study  
32 of animal behaviour (Croft *et al.*, 2008; Krause *et al.*, 2014). Missing individuals are a frequent  
33 problem in animal network studies, when it is often necessary to capture and mark individuals to  
34 gather data. However, as the use of bio-logging technology becomes more widespread to study  
35 human social networks (e.g. Isella *et al.*, 2011; Kiti *et al.*, 2016; Mastrandrea *et al.*, 2015), similar  
36 problems often arise. I will provide a perspective as an applied network practitioner working on  
37 animal social behaviour as to the utility of their new findings, and then build on this to highlight  
38 important outstanding questions relating to missing individuals in networks. Finally, I will present  
39 some R code designed to test how missing individuals affect the calculation of network metrics in  
40 animal social networks that I hope will complement the java applet provided in that paper.

41

## 42 **An applied perspective on the implications of Smith *et al.* (2017)**

43         Smith *et al.* (2017) built upon previous work by the same authors (Smith and Moody, 2013)  
44 in exploring the consequences of missing individuals on the calculation of a range of network  
45 metrics. Together, especially when taken alongside complementary findings in other fields (e.g. (Silk  
46 *et al.*, 2015)), the results of this work have revealed a set of important considerations when  
47 analysing networks with missing data, of which the general rules are already very useful to applied  
48 network practitioners. In particular, knowledge of how network structure and size influence the  
49 impact of sub-sampling from a network is paramount, as is an understanding that more global

50 metrics (such as Betweenness) are less resilient to the presence of missing individuals. Finally, the  
51 exploration of biases in missing individuals addressed by Smith *et al.* (2017) is especially valuable  
52 from the perspective of animal behaviour research. This is partly as an aid in determining which  
53 individuals to collect data on when resources are limited, but also because methods of capture to  
54 make individual animals identifiable for network studies may place implicit biases on which  
55 individuals are most likely to remain unsampled. However, one opportunity that is missed here is the  
56 opportunity to discuss the importance of different types of “bias” caused by missing network data. I  
57 would like to highlight here the nomenclature/approach used by Silk *et al.* (2015) which looked at  
58 the effect of sub-sampling on three distinct properties of network metrics. In this paper the authors  
59 looked at *precision*, *accuracy* and *bias* of metric values in sub-sampled networks. As defined by these  
60 authors, *precision* is the correlation between values calculated in the partial (observed) network and  
61 those in the equivalent true network, *accuracy* is the value of the metric obtained from the observed  
62 partial network relative to its true value, and *bias* is systematic variation in the precision of metric  
63 values in the partial (observed) network. The impact of sub-sampling on each of these properties can  
64 depend on network structure and the type of network metric being investigated. Considering each  
65 independently is important in wider applications of network analysis, as these different properties  
66 can be used to address different questions. For example, in animal network studies it is the *precision*  
67 of metric values in sub-sampled networks that is important if a researcher asks a question about  
68 whether network position and a personality trait are linked, but it is *accuracy* that is important if a  
69 researcher seeks to compare the true values of network metrics between different contexts.

70 Another major step forward in Smith *et al.* (2017) is the development of a tool (a java  
71 applet) that can provide an idea of the impact of missing individuals (incorporating any biases in  
72 their centrality) according to network size for a range of network structures. For an applied network  
73 researcher that has worked on study systems with substantial proportions of missing individuals this  
74 is an exciting development, and has the potential to be useful as a guide to researchers designing  
75 network studies. However, from this perspective I also feel that it is essential to see a tool designed

76 in this way just as a starting point. The use of network analysis in animals is now highly question-  
77 driven and such a fixed tool has only restricted possibilities for use. It would be great to see a more  
78 modular set of functions that were able to use pilot empirical data or researcher knowledge to  
79 simulate a realistic network structure, and then sample from this structure, before determining how  
80 it might affect the outcome of a range of network analytical perspective. Such a package of functions  
81 would be best developed as a community, preferably with researchers from a range of fields so that  
82 the generation of networks, and the types of metrics to calculate (or statistical models to assess) was  
83 relevant to as wide a range of studies as possible. The advantages of taking this approach is that as a  
84 user assesses a new sampling framework and/or question, the code they used can be added to the  
85 system and shared with researchers who might be faced with a similar problem in another field.

86

### 87 **Three outstanding missing network data problems**

88 In this section I will highlight some important gaps in our understanding on the impact of  
89 sub-sampling from social networks that simulation-modelling could easily test and greatly aid the  
90 design of empirical studies. While these ideas come from a background of employing network  
91 analysis to study animal behaviour, I feel all are widely applicable in the use of social networks more  
92 generally.

93

### 94 **What is the best way of sub-sampling social networks?**

95 In many animal network studies time, cost and effort is required to capture individuals and  
96 make them individually identifiable for social network studies. This trade-off is now becoming more  
97 frequent for all types of network study, including in humans, as the use of bio-logging approaches to  
98 produce reality mining data on social behaviour is increasing (Barrat and Cattuto, 2015; Isella et al.,  
99 2011). Often these approaches are costly, and it is possible to use only relatively small sample sizes.

100 Therefore, deciding how best to deploy bio-logging devices for network studies remains an open  
101 question. For example, is it best to intensively sample a small part of a network or sample a larger  
102 part of the network more sparsely? Or similarly, how would studying replicate networks in multiple  
103 populations trade-off against the intensity of collaring individuals within each population? It is likely  
104 that the type of question being asked is important to making this decision. For example, work  
105 focussing on fine-scale behavioural interactions, such as dominance behaviour in animals (e.g. Dey et  
106 al., 2015), is likely to benefit from intensively sampling particular groups. In contrast, for the study of  
107 population-level processes such as disease transmission, a more even distribution of identifiable  
108 individuals throughout a population may be beneficial, especially when attempting to record  
109 infrequent interactions.

110 In order to assist empirical researchers making these decisions it will be important to move  
111 simulation models of non-random sampling beyond the missingness-centrality correlation  
112 investigated by Smith *et al.* (2017) to assess the impacts of the clustering of identifiable and  
113 unidentifiable individuals within sub-sampled networks. Exploring this in a range of social network  
114 structures will be an important step forward in aiding the study design of network studies in natural  
115 systems, especially for studies using bio-logging approaches.

116

### 117 **How do missing individuals affect individual-level hypothesis testing in networks?**

118 The relationship between social network position and other individual traits has been a  
119 major cross-disciplinary research focus (e.g. Aplin et al., 2013; Bollen et al., 2011; Goodreau et al.,  
120 2009; Rosenquist et al., 2011). The most popular methods to test these hypotheses has been  
121 different, for example the use of exponential random graph models (ERGMs; Lusher et al., 2013) and  
122 stochastic actor-oriented models (SAOMs; Snijders et al., 2010) in the social sciences, versus the  
123 development of randomisation-based generalised linear mixed model approaches in animal  
124 behaviour (Croft et al., 2011; Farine and Whitehead, 2015). Regardless, the impact of missing

125 individuals (and edges) on statistical inferences made using these all of these approaches remains an  
126 open and important question.

127         Smith *et al.* (2017) made a step towards addressing this, by looking at the consequences of  
128 missing individuals for tests of behavioural homophily within a network (and finding that it was  
129 possible to detect patterns of behavioural homophily when there was both a high proportion of  
130 missing individuals and a bias in which individuals were missing). However, a notion of the  
131 preponderance of type I and type II errors when different modelling frameworks are used to analyse  
132 networks with missing individuals would represent a considerable step forward in our understanding  
133 of the consequences of sub-sampling social networks. For example, while Shalizi and Rinaldo (2013)  
134 have suggested that ERGMs estimated on a sampled network are unlikely to reflect population-level  
135 parameters (*accuracy* in the framework outlined previously), this may not affect their ability to test  
136 hypotheses related to individual differences.

137         It would seem fairly simple to build on previous simulation-modelling work to examine how  
138 hypothesis testing using any of the statistical models mentioned above might be affected by the sub-  
139 sampling of networks. For example, the addition of a response variable that depended on network  
140 structure to the R function outlined in the next section would enable the impact on inference from  
141 generalised linear models to be addressed. In the case of models relating individual traits to  
142 individual-level network models, such as those suggested above, there are two main considerations  
143 to make. Firstly, the structure of the network, specifically the distribution of metrics used as an  
144 explanatory variable, is likely to influence the problems generated by sub-sampling. As a case in  
145 point, networks with higher modularity will have highly-skewed distributions for some metrics and  
146 missing individuals may reduce power considerably, especially if missingness is non-random.  
147 Secondly, the distribution of the response variable will also be an important consideration. High  
148 levels of overdispersion or zero-inflation in the response variable may exacerbate any problems  
149 associated with missing individuals. For example, individual traits with a negative binomial

150 distribution (such as parasite infection load) are likely to be problematic. Again, there might be a  
151 particularly large impact on statistical power if there is a correlation between trait values and  
152 missingness.

153

#### 154 **How do missing individuals affect estimates of transmission processes?**

155 A further major application of network analysis is to study transmission through populations;  
156 for example of information (Allen et al., 2013; Bakshy et al., 2012) and disease (Reynolds et al., 2015;  
157 Rohani et al., 2010; Stehlé et al., 2011). A case study is the application of network-based diffusion  
158 analysis (NBDA) as a powerful approach to detect social transmission in animal populations (Allen et  
159 al., 2013; Aplin et al., 2015; Franz and Nunn, 2009). Sub-sampled networks would also be expected  
160 to reduce the power of these approaches and/or lead to systematic biases in the inferences made  
161 about transmission (Ghani et al., 1998). Further type II error may result if non-random missingness  
162 means that more central individuals, which may spread information to more new individuals on  
163 average, are not identifiable. This effect of this might be particularly substantial in networks with  
164 higher modularity in which certain individuals are likely to have high brokerage between  
165 communities.

166 Smith et al. (2017) investigated the effect of missingness on measures of network topology  
167 that will have important implications for transmission - component size, bicomponent size and  
168 distance. All measures were highly sensitive to missing individuals (in the majority of the networks  
169 they investigated) when there was a positive missingness-centrality correlation, and even a relatively  
170 small proportion of missing nodes in this case had a substantial impact on these measures of  
171 network topology. The authors highlighted that the impact of the missingness-centrality correlation  
172 was expected in this case as more central individuals are more likely to be important to the cohesion  
173 of the network. They also suggested implications for diffusion or transmission within the network,  
174 although stopped short of using simulations to test this.

175           There are important applications of studying spreading processes in networks, including in  
176 systems prone to missing individuals such as animal populations (e.g. Hamede et al., 2012; Reynolds  
177 et al., 2015). As a result, simulation-studies that explore the impact of sub-sampling networks on  
178 conclusions drawn about transmission represent an important area of future work. It would not be  
179 too great an extension to simulate the spread of a trait across a static network and then  
180 progressively remove individuals (either at random or in a biased manner) and monitor the change  
181 in power of NBDA or similar approaches. In many cases, especially for disease transmission, the  
182 acquisition of a trait can result in potentially important changes to social network position (Ezenwa  
183 et al. 2016, Silk et al. 2017). Therefore, a further more complex extension would be to extend  
184 simulation-modelling approaches to test the impact of missing data in situations where both the  
185 network and trait were dynamic.

186

### 187 **A tool for generating and sampling from varying network structures**

188           In the supplementary information I present R functions for generating and sub-sampling  
189 from weighted, undirected social networks in a range of structures. Similarly to java applet provided  
190 by Smith et al. (2017) this is intended to encourage researchers to consider the potential impacts of  
191 missing nodes on their analyses (and in many situations this tool will be sufficient). However, unlike  
192 the java applet the code enables readers to generate user-defined networks and sample from them.  
193 An added advantage is that the code is designed in such a way that researchers familiar with R could  
194 add their own functions to address sampling-related problems of interest to them.

195           The code draws tie strengths from a zero-inflated negative binomial distribution. In its most  
196 basic form, this negative binomial distribution is equivalent for all individuals. However, I present it  
197 with two possible extensions, which will be of general interest to many network researchers. First,  
198 the tool includes code that enables generated networks to be “spatially structured” by placing  
199 individuals in 2d space and basing the parameters of the negative binomial distribution on the



200 distance matrix of their locations. Many social network studies take place in populations where  
201 social interactions are structured by space use, and in these contexts being able to develop  
202 simulated networks that account for this is important. Second the tool includes the ability to assign  
203 individuals into social groups of fixed or varying sizes. Different distributions of tie strengths can be  
204 set for within-group and out-of-group ties in the network allowing the generation of networks with  
205 the strong social group structure frequently found in many human and animal populations. It also  
206 provides a highly tractable way to test network sub-sampling questions in networks differing in  
207 modularity. The capability of this tool to produce a whole range of potentially realistic network  
208 structures is demonstrated in Figure 1. If preferred then networks could alternatively be generated  
209 by using an exponential random graph model to create networks with a set of desired target  
210 statistics, and the networks generated using this method could be used in the subsequent R  
211 functions in the same manner.

212         The supplementary information also includes code that is able to sub-sample from these  
213 networks, as a simple illustration of how it can be used to address some of the issues discussed in  
214 this paper. This calculates and outputs the precision, accuracy and bias of four centrality metrics  
215 (degree, strength, betweenness and eigenvector centrality) in partial networks of pre-defined size  
216 (i.e. the proportion of the population made identifiable). The precision, accuracy and bias of these  
217 metrics calculated in the networks depicted in Fig. 1 is provided in the supplementary material. In  
218 this example networks of 50 nodes were generated using the three models specified in the legend of  
219 Figure 1. Network a) contained no (imposed) modular and limited spatial structure, network b)  
220 contained strong spatial structure and no additionally imposed modular structure, and network c)  
221 consisted of 10 modules of five nodes with strong spatial structure between groups. Therefore, the  
222 examples provide a clear indication of a range of sensible parameter values. The second R function  
223 was then used to randomly sub-sample these networks to contain 60% and 80% of the original  
224 nodes once (it could be applied multiple times to generate repeat sampling events and facilitate

225 statistical comparisons). Finally, the third R function calculated the precision, accuracy and bias (as  
226 defined above) of metric values in the sub-sampled networks (Table S1).

227           Currently the code is limited to random sub-sampling of nodes, and to tests of the precision,  
228 bias and accuracy of centrality metrics. However, it provides a valuable basis for developing  
229 simulation models to address some of the outstanding issues highlighted above. For example, the  
230 ability to have both spatial and modular distribution of the nodes in the network facilitates  
231 adjustments to the sampling regime that would easily enable the impact of clustered versus random  
232 sampling at different scales to be tested. Similarly, the replacement of the final R function (which  
233 currently calculates the precision, accuracy and bias of sampled networks) with an alternative  
234 function that fits statistical models (e.g. ERGMs) and compares parameter estimates at different  
235 levels of sampling would make it easily possible to determine how sub-sampling nodes affects  
236 statistical inference and hypothesis testing.

237           Together these R functions provide a tool to estimate the impact of sub-sampling networks  
238 in a whole-range of user-defined biologically realistic network structures. By providing modular code  
239 in GNU software I hope to provide a toll that other researchers can use in a system-specific manner  
240 according to their needs, thus complementing the java applet provided by Smith *et al.* (2017) which  
241 is predominantly targeted as an easy-to-use and quick guide applicable to general types of network  
242 structure. Together, with the developments in that paper, I anticipate that these R functions will  
243 trigger further work towards addressing some of the outstanding questions about sub-sampling  
244 networks highlighted previously.

245

## 246 **Conclusions**

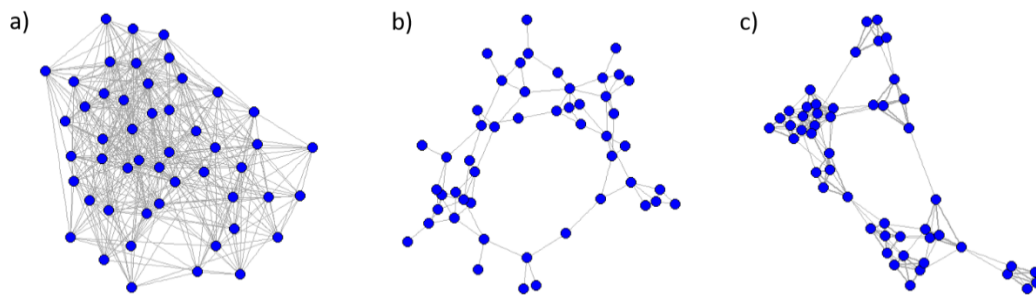
247           The publication of Smith *et al.* (2017) advances our understanding of the impact of missing  
248 individuals on network analysis, and offers great potential to applied network practitioners outside

249 of the social sciences. However, I advocate the use of approaches that move beyond calculating  
250 biases in descriptive network metrics. In particular, a move towards determining how sub-sampling  
251 networks in empirical systems affects our power to use statistical models of network structure and  
252 individual behaviour will be especially beneficial. This will be facilitated by more modular and  
253 adaptable approaches to examining the effects of missingness in network data that can be shared  
254 among researchers and will be applicable to networks in a range of fields.

255

256

257 **Figures**



258  
259 Figure 1. Three networks generated using the `network.generator()` R function provided in the  
260 supplementary material. For a) `groups=50`, `mean.group.size=max.group.size=1`, `d.eff=2`, `o.dens=0.7`  
261 and `i.dens=6` (not that this parameter is used in this case). For b) `groups=50`,  
262 `mean.group.size=max.group.size=1`, `d.eff=10`, `o.dens=0.7` and `i.dens=6` (not that this parameter is  
263 used in this case). For c) `groups=10`, `mean.group.size=max.group.size=5`, `d.eff=10`, `o.dens=1` and  
264 `i.dens=6`.

265

266 **References**

- 267 Alba, R.D., 1982. Taking stock of network analysis: A decade's results. *Res. Sociol. Organ.* 1, 39–74.
- 268 Allen, J., Weinrich, M., Hoppitt, W., Rendell, L., 2013. Network-based diffusion analysis reveals  
269 cultural transmission of lobtail feeding in humpback whales. *Science* (80-. ). 340, 485–488.
- 270 Aplin, L.M., Farine, D.R., Morand-Ferron, J., Cockburn, A., Thornton, A., Sheldon, B.C., 2015.  
271 Experimentally induced innovations lead to persistent culture via conformity in wild birds.  
272 *Nature* 518, 538–541.
- 273 Aplin, L.M., Farine, D.R., Morand-Ferron, J., Cole, E.F., Cockburn, A., Sheldon, B.C., 2013. Individual  
274 personalities predict social behaviour in wild networks of great tits (*Parus major*). *Ecol. Lett.* 16,  
275 1365–1372.
- 276 Bakshy, E., Rosenn, I., Marlow, C., Adamic, L., 2012. The role of social networks in information  
277 diffusion, in: *Proceedings of the 21st International Conference on World Wide Web. ACM*, pp.  
278 519–528.
- 279 Barrat, A., Cattuto, C., 2015. Face-to-face interactions, in: *Social Phenomena. Springer*, pp. 37–57.
- 280 Bollen, J., Gonçalves, B., Ruan, G., Mao, H., 2011. Happiness is assortative in online social networks.  
281 *Artif. Life* 17, 237–251.
- 282 Croft, D.P., James, R., Krause, J., 2008. *Exploring animal social networks. Princeton University Press.*
- 283 Croft, D.P., Madden, J.R., Franks, D.W., James, R., 2011. Hypothesis testing in animal social networks.  
284 *Trends Ecol. Evol.* 26, 502–507.
- 285 Dey, C.J., Tan, Q.Y.J., O'Connor, C.M., Reddon, A.R., Caldwell, J.R., Balshine, S., 2015. Dominance  
286 network structure across reproductive contexts in the cooperatively breeding cichlid fish  
287 *Neolamprologus pulcher*. *Curr. Zool.* 61, 45–54.
- 288 Ezenwa, V.O., Archie, E.A., Craft, M.E., Hawley, D.M., Martin, L.B., Moore, J. & White, L. (2016) Host  
289 behaviour–parasite feedback: an essential link between animal behaviour and disease ecology.  
290 *Proc. R. Soc. B*, p. 20153078. The Royal Society. Farine, D.R. (2013). Animal social network  
291 inference and permutations for ecologists in R using *asnipe*. *Methods in Ecology and Evolution*,  
292 4, 1187–1194.
- 293 Farine, D.R., Whitehead, H., 2015. Constructing, conducting and interpreting animal social network  
294 analysis. *J. Anim. Ecol.* 84, 1144–1163.
- 295 Franz, M., Nunn, C.L., 2009. Network-based diffusion analysis: a new method for detecting social  
296 learning. *Proc. R. Soc. London B Biol. Sci.* 276, 1829–1836.
- 297 Ghani, A.C., Donnelly, C.A., Garnett, G.P., 1998. Sampling biases and missing data in explorations of  
298 sexual partner networks for the spread of sexually transmitted diseases. *Stat. Med.* 17, 2079–  
299 2097.
- 300 Goodreau, S.M., Kitts, J.A., Morris, M., 2009. Birds of a feather, or friend of a friend? Using  
301 exponential random graph models to investigate adolescent social networks. *Demography* 46,  
302 103–125.
- 303 Hamede, R., Bashford, J., Jones, M., McCallum, H., 2012. Simulating devil facial tumour disease  
304 outbreaks across empirically derived contact networks. *J. Appl. Ecol.* 49, 447–456.
- 305 Isella, L., Stehlé, J., Barrat, A., Cattuto, C., Pinton, J.-F., Van den Broeck, W., 2011. What's in a crowd?

306 Analysis of face-to-face behavioral networks. *J. Theor. Biol.* 271, 166–180.

307 Kiti, M.C., Tizzoni, M., Kinyanjui, T.M., Koech, D.C., Munywoki, P.K., Meriac, M., Cappa, L., Panisson,  
308 A., Barrat, A., Cattuto, C., 2016. Quantifying social contacts in a household setting of rural  
309 Kenya using wearable proximity sensors. *EPJ Data Sci.* 5, 21.

310 Krause, J., James, R., Franks, D.W., Croft, D.P., 2014. *Animal social networks*. Oxford University Press.

311 Lopes, P.C., Block, P., König, B., 2016. Infection-induced behavioural changes reduce connectivity  
312 and the potential for disease spread in wild mice contact networks. *Sci. Rep.* 6.

313 Lusher, D., Koskinen, J., Robins, G., Lusher, D., Koskinen, J., Robins, G., 2013. Exponential random  
314 graph models for social networks. *Structural analysis in the social sciences*.

315 Mastrandrea, R., Fournet, J., Barrat, A., 2015. Contact patterns in a high school: a comparison  
316 between data collected using wearable sensors, contact diaries and friendship surveys. *PLoS*  
317 *One* 10, e0136497.

318 Reynolds, J.J.H., Hirsch, B.T., Gehrt, S.D., Craft, M.E., 2015. Raccoon contact networks predict  
319 seasonal susceptibility to rabies outbreaks and limitations of vaccination. *J. Anim. Ecol.* 84,  
320 1720–1731.

321 Rohani, P., Zhong, X., King, A.A., 2010. Contact network structure explains the changing  
322 epidemiology of pertussis. *Science* (80-. ). 330, 982–985.

323 Rosenquist, J.N., Fowler, J.H., Christakis, N.A., 2011. Social network determinants of depression. *Mol.*  
324 *Psychiatry* 16, 273–281.

325 Shalizi, C.R. & Rinaldo A. (2013). Consistency under sampling of exponential random graph models.  
326 *Annals of Statistics*, **41**, 508-535.

327 Silk, M.J., Jackson, A.L., Croft, D.P., Colhoun, K., Bearhop, S., 2015. The consequences of  
328 unidentifiable individuals for the analysis of an animal social network. *Anim. Behav.* 104, 1–11.  
329 doi:10.1016/j.anbehav.2015.03.005

330 Silk, M.J., Croft, D.P., Delahay R.J., Hodgson, D.J., Boots M., Weber N. and McDonald R.A. (2017).  
331 Using Social Network Measures in Wildlife Disease Ecology, Epidemiology, and Management.  
332 *BioScience* doi: 10.1093/biosci/biw175

333 Smith, J.A., Moody, J., 2013. Structural effects of network sampling coverage I: Nodes missing at  
334 random. *Soc. Networks* 35, 652–668.

335 Smith, J.A., Moody, J., Morgan, J.H., 2017. Network sampling coverage II: The effect of non-random  
336 missing data on network measurement. *Soc. Networks* 48, 78–99.

337 Snijders, T.A.B., Van de Bunt, G.G., Steglich, C.E.G., 2010. Introduction to stochastic actor-based  
338 models for network dynamics. *Soc. Networks* 32, 44–60.

339 Stehlé, J., Voirin, N., Barrat, A., Cattuto, C., Colizza, V., Isella, L., Régis, C., Pinton, J.-F., Khanafer, N.,  
340 Van den Broeck, W., 2011. Simulation of an SEIR infectious disease model on the dynamic  
341 contact network of conference attendees. *BMC Med.* 9, 1.

342