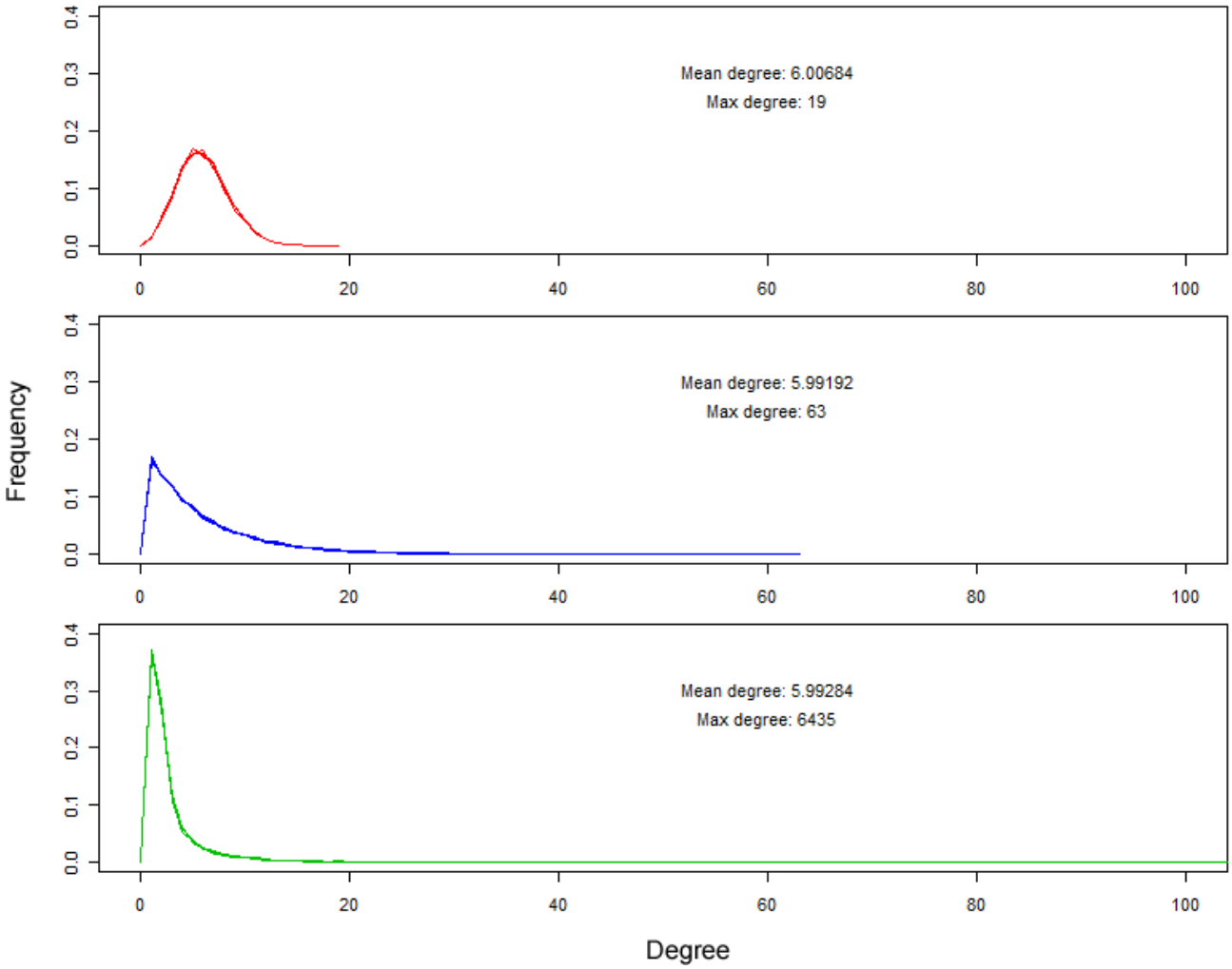
Supplementary Material
Inferring population-level contact heterogeneity from common epidemic data

Synthetic data

Network generation

Test networks were simulated for this study using the networkx module (version 1.1) for Python 2.6. We use the configuration model (Molloy & Reed, 1995) to generate uncorrelated, random networks. This algorithm assigns i_d "half-edges" to each node i where i_d is the degree of node i, and then randomly connects pairs half-edges to create edges until there are no half-edges left. Self-loops (an edge from a node to itself) and duplicate edges are removed subsequently by randomly rewiring edges using Taylor's algorithm (Taylor, 1981). As stated in the text, the degree distribution of each class reflects the heterogeneity in contacts in the population. Below the degree distributions for each network are plotted:

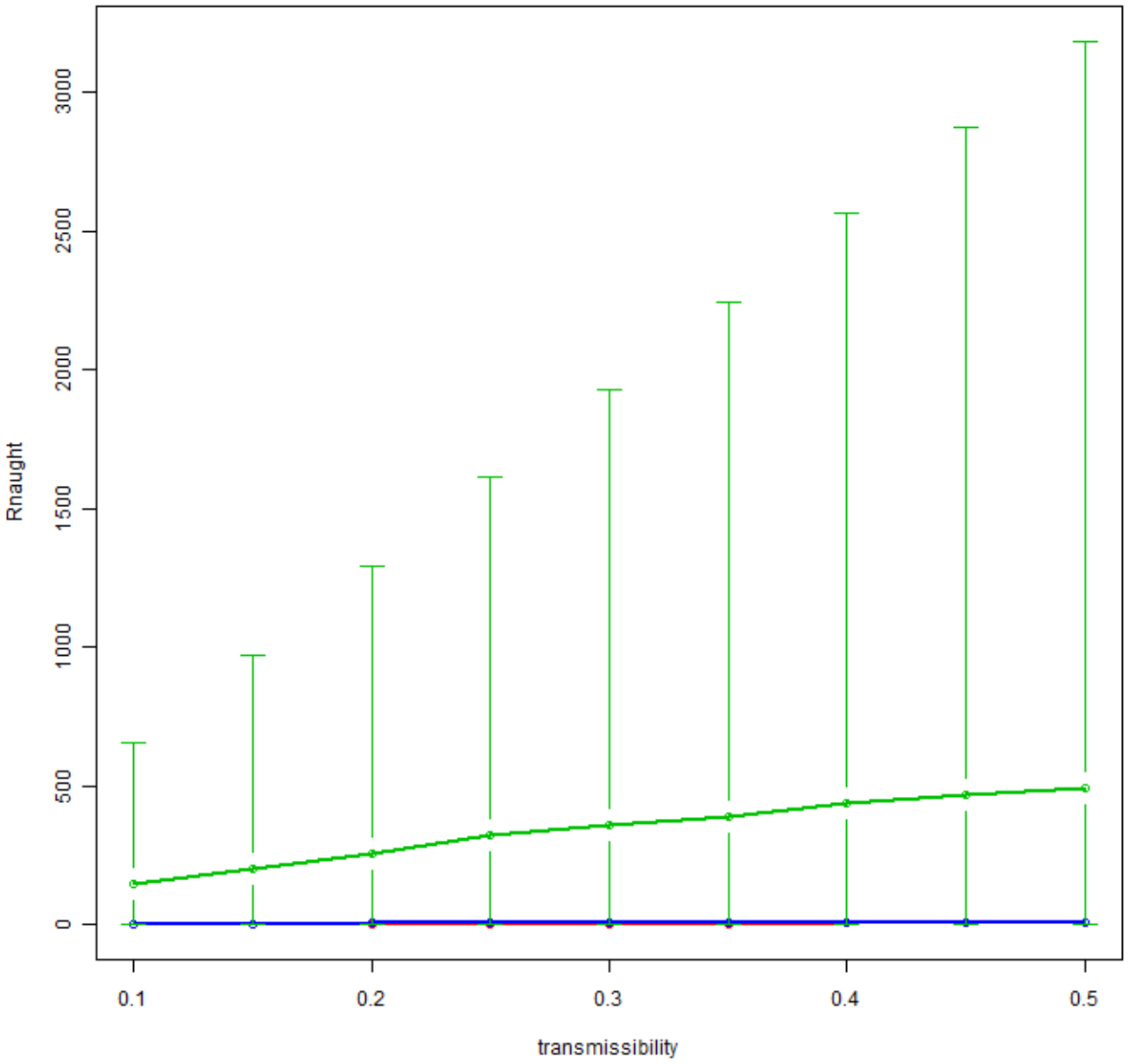


Supplementary Figure 1: The degree distributions for each instance of each network class used in the simulation study (mean degree = 6, N = 10000). Y-axis shows the relative frequency of a degree and the x-axis shows the degrees themselves. Red shows networks with Poisson distribution (top), blue shows networks with exponential distributions (middle), and green shows networks with scale-free distributions (bottom).

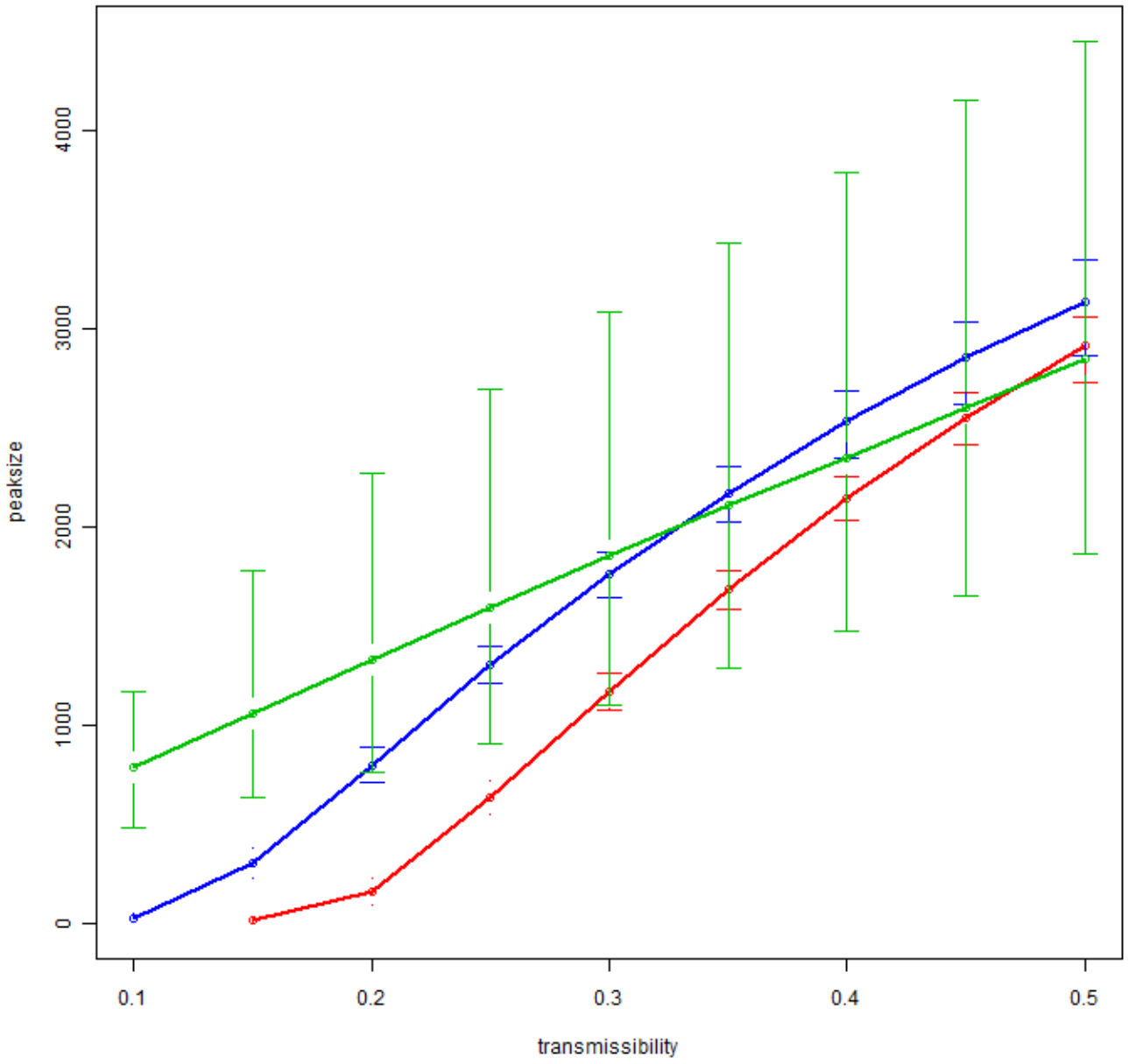
Epidemic simulation

Under the epidemic algorithm we used, one node is randomly chosen as the index case and "infected" and all other nodes are considered susceptible. The infection process then proceeds in discrete time. During a time step, all edges connecting an infected and susceptible node are discovered and the susceptible node is infected

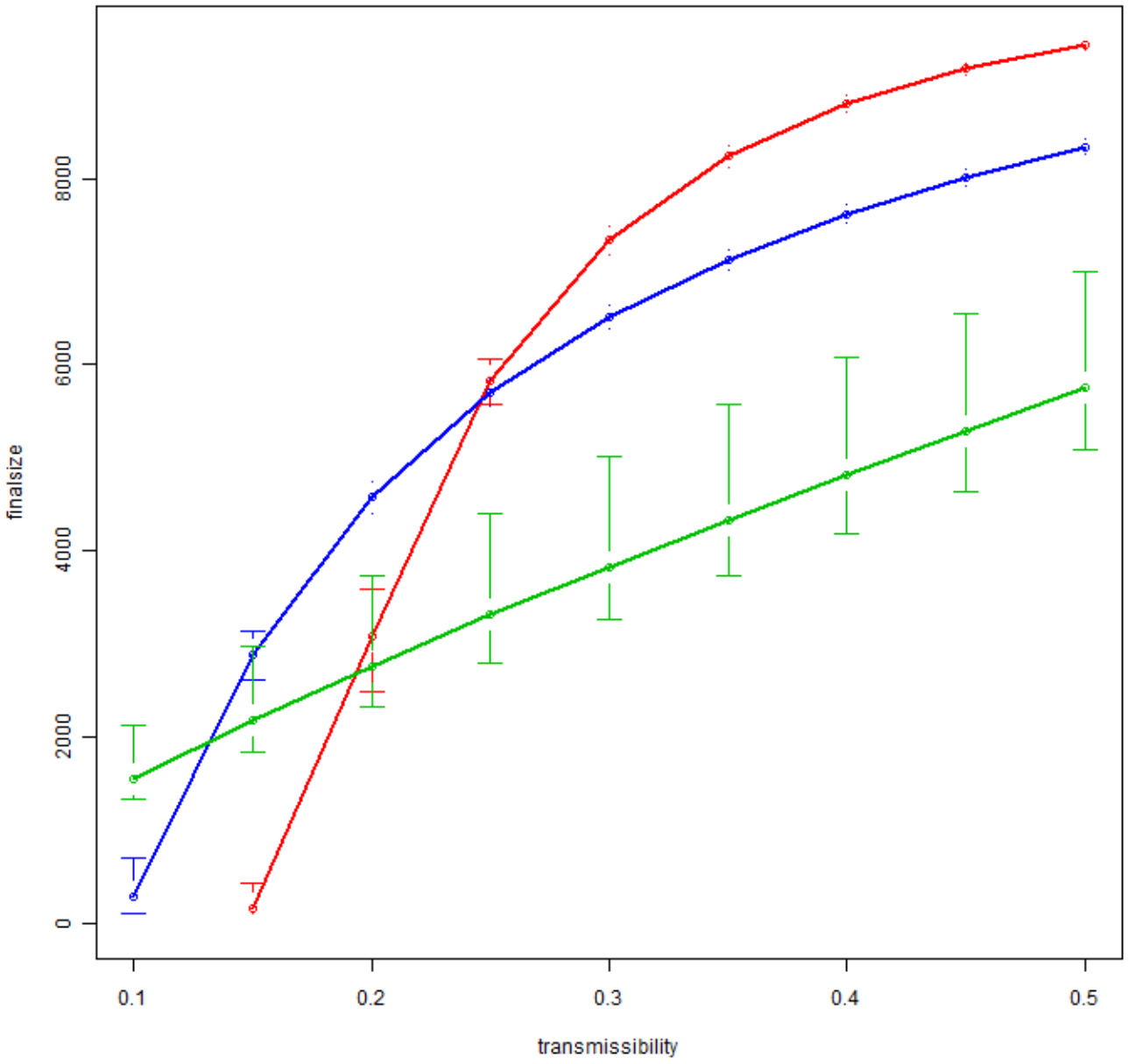
with probability *T*. Once this is completed, all nodes which were not infected during this time step are considered to have "recovered" and are effectively removed from the network. The time step then ends and a new one starts only if there are infected nodes in the network. There were then broken down into the four epidemiological measures and stored in a database. From here, the epidemiological measures were queried and used as distribution functions. Simulated distributions of these four measures are summarized below, for all transmission probabilities and network types.



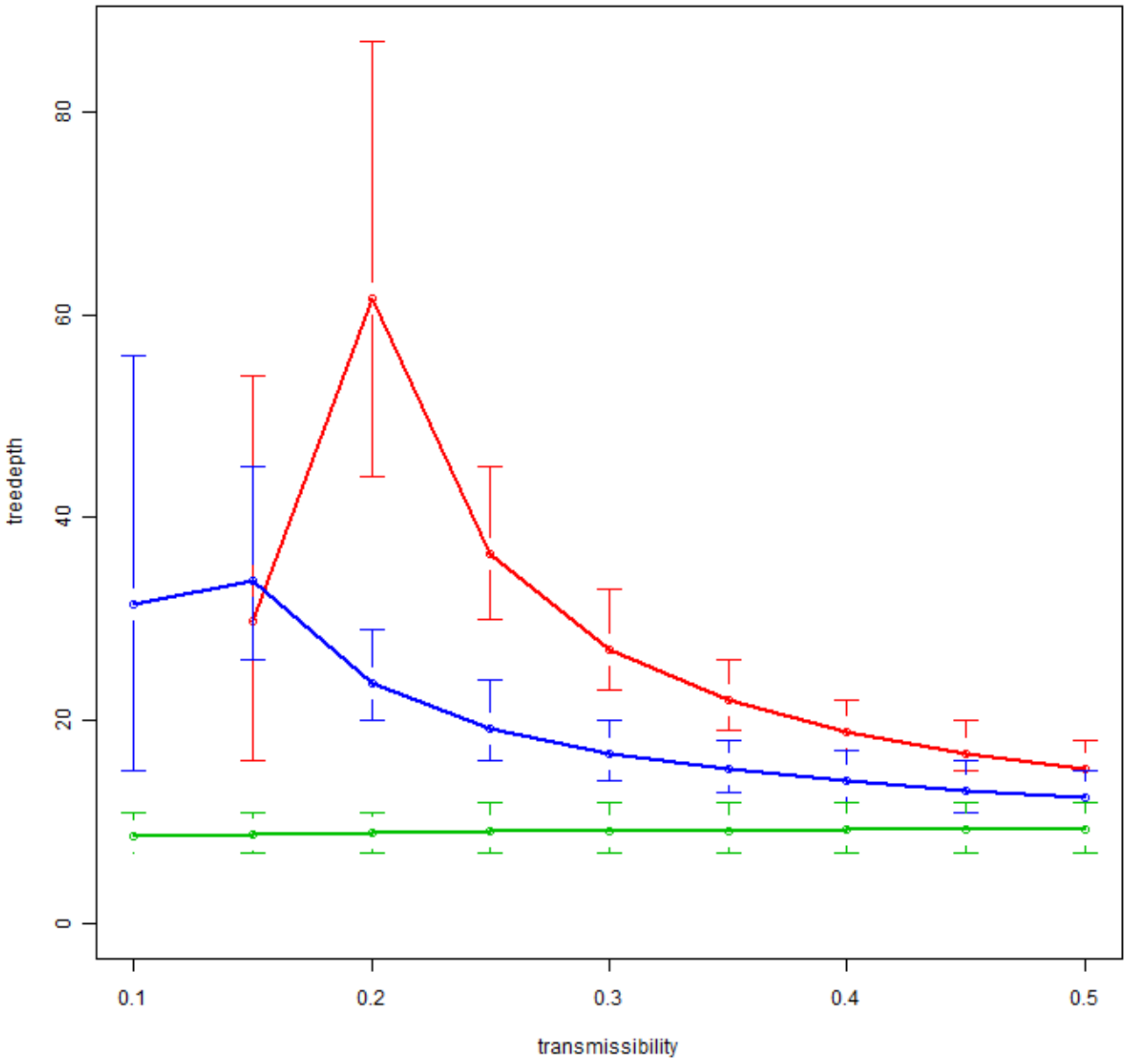
Supplementary Figure 2: R0 value distributions (y-axis) for all three random graph models over all the transmission probabilities analyzed (x-axis). Thick, horizontal lines represent the median value and the terminals of associated vertical lines show the 0.025% and 97.5% quantiles. The red, blue, and green lines indicate that the data were generated from Poisson, exponential, and scale-free graphs, respectively.



Supplementary Figure 3: Epidemic peak size distributions (y-axis) for all three random graph models over all the transmission probabilities analyzed (x-axis). Thick, horizontal lines represent the median value and the terminals of associated vertical lines show the 0.025% and 97.5% quantiles. The red, blue, and green lines indicate that the data were generated from Poisson, exponential, and scale-free graphs, respectively.



Supplementary Figure 4: Epidemic final size distributions (y-axis) for all three random graph models over all the transmission probabilities analyzed (x-axis). Thick, horizontal lines represent the median value and the terminals of associated vertical lines show the 0.025% and 97.5% quantiles. The red, blue, and green lines indicate that the data were generated from Poisson, exponential, and scale-free graphs, respectively.



Supplementary Figure 5: Epidemic duration distributions (y-axis) for all three random graph models over all the transmission probabilities analyzed (x-axis). Thick, horizontal lines represent the median value and the terminals of associated vertical lines show the 0.025% and 97.5% quantiles. The red, blue, and green lines indicate that the data were generated from Poisson, exponential, and scale-free graphs, respectively.

Model comparisons

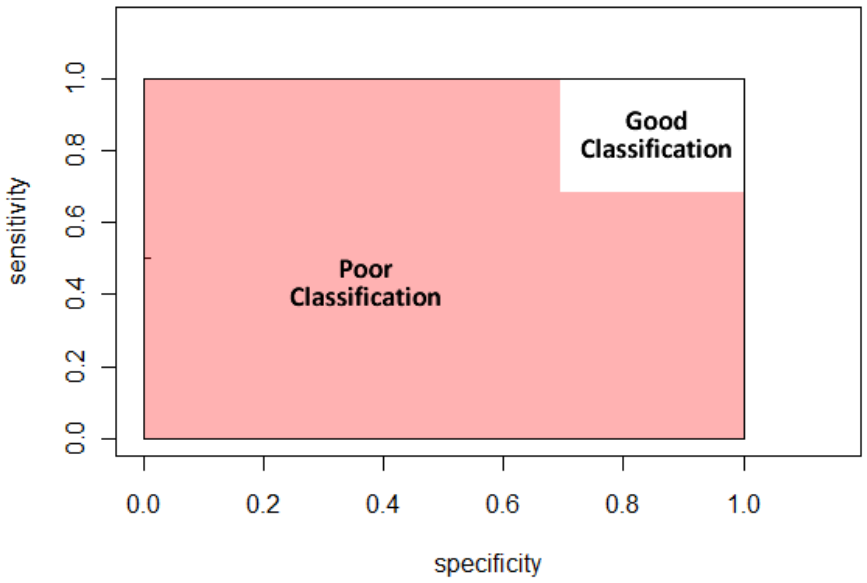
This section contains more details on how the three random graph models were compared and cross-validated, helping to interpret Figure 3-1 from the main text.

Sensitivity vs. Specificity

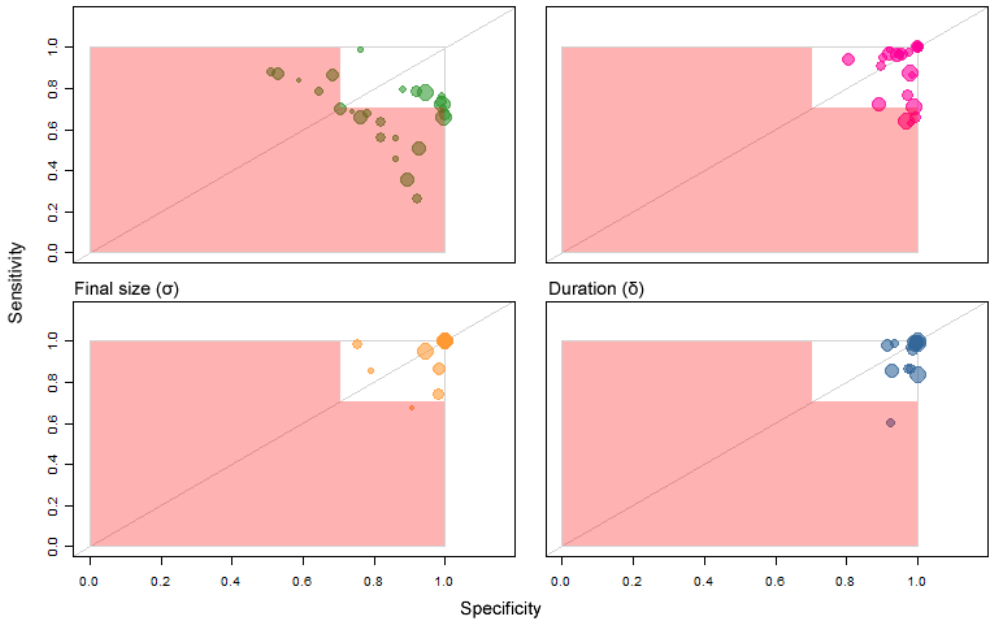
This analysis measures the results from classification functions which return binary data. In our study, the two binary outcomes are (0) test *did not* accurately and precisely associate the test datum with the random graph model it was generated on and (1) the test did do this. As we are comparing the relative likelihoods of three models (M_1,M_2,M_3), our classification function was defined as:

$$f(L_{actual}, L_{other.1}, L_{other.2}) = \begin{cases} 0, & L_{actual} < max(L_{actual}, L_{other.1}, L_{other.2}) \\ 1, & L_{actual} = max(L_{actual}, L_{other.1}, L_{other.2}) \end{cases}$$

where L_{actual} is the likelihood of the random graph model, M_{actual}, which was actually used to generate the datum and $L_{other.1}, L_{other.2}$ are likelihoods of the other two models, $M_{other.1}$, $M_{other.2}$. The schematic below describes how the sensitivity and specificity plots of these binary outcomes can be roughly interpreted:



Overlaid on Figure 3-1 of the main text, that figure becomes:

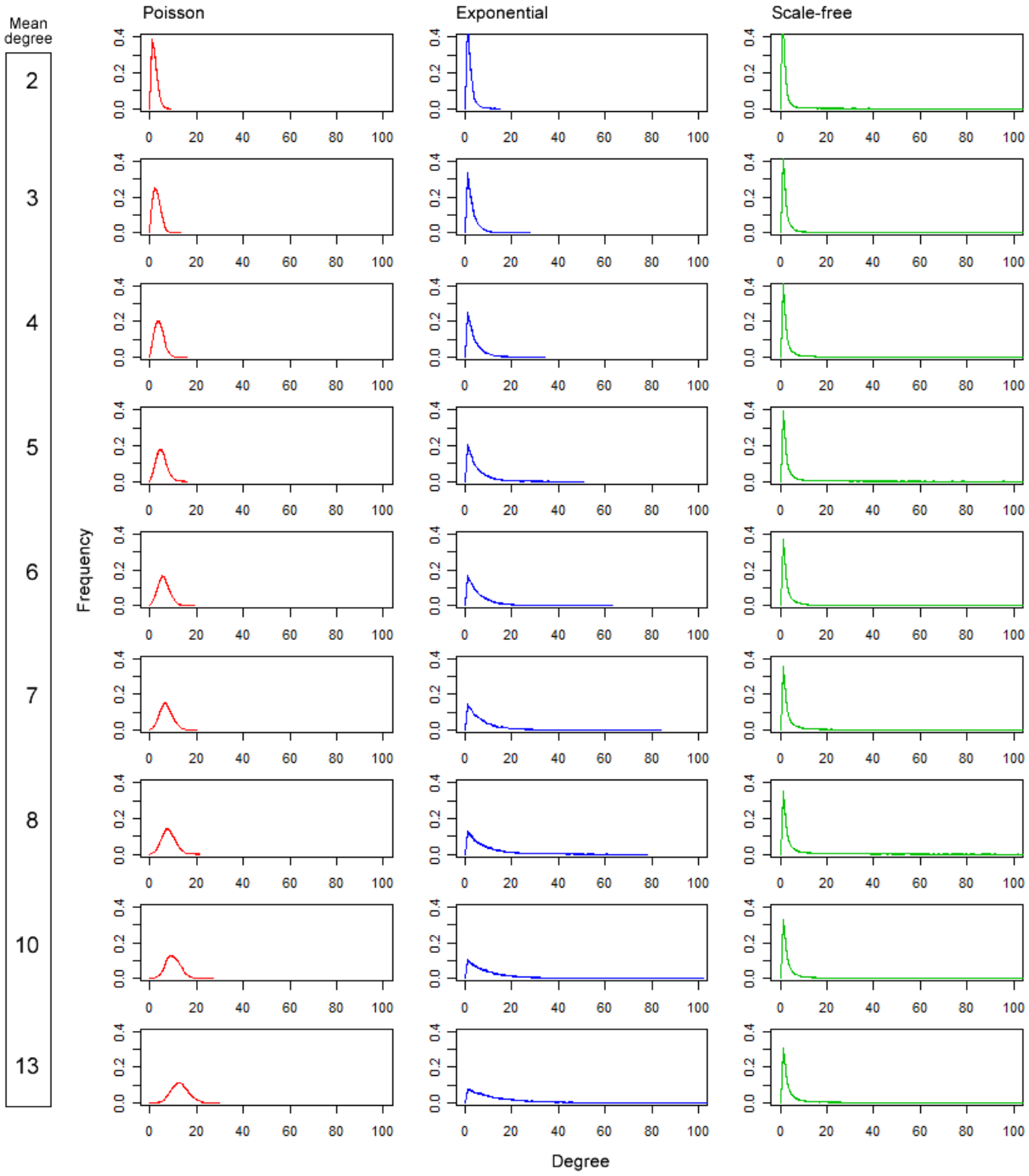


Sensitivity Analysis

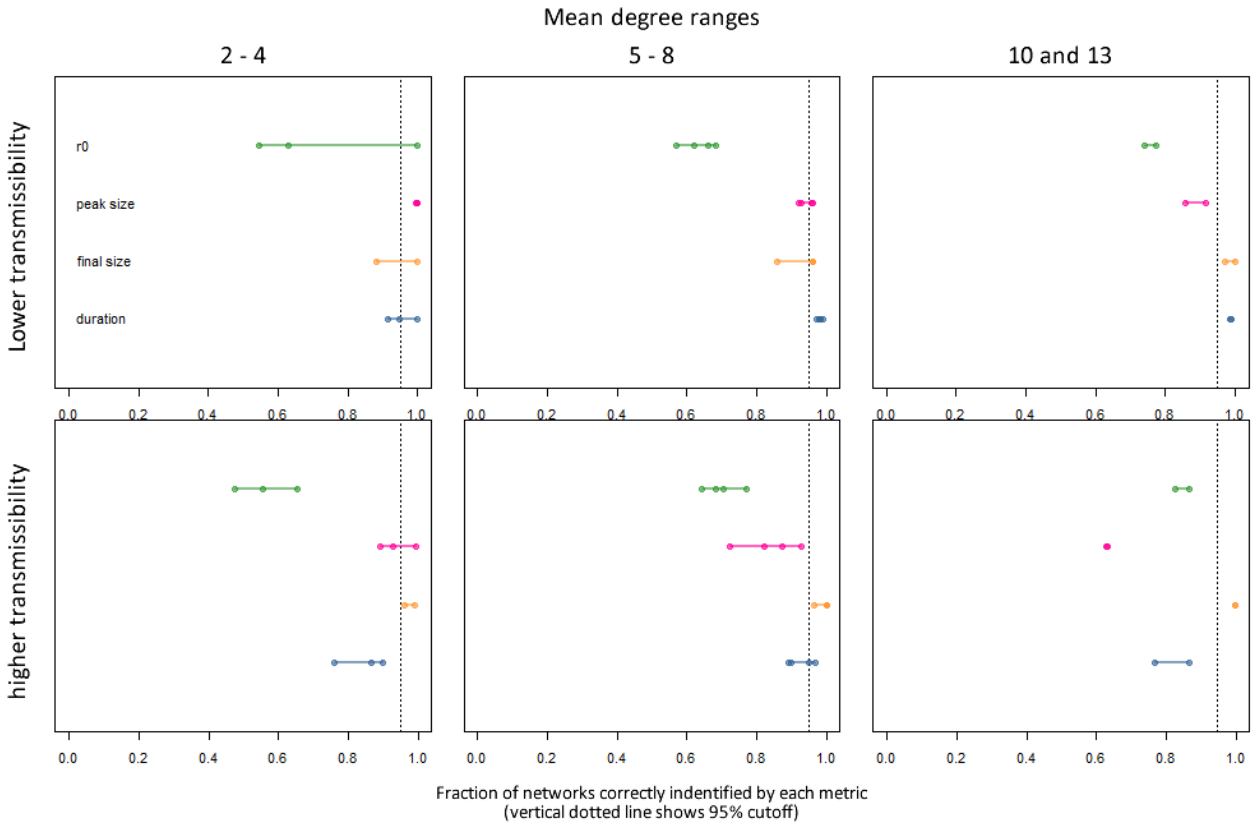
Sensitivity analyses were carried out in order to determine how our simulation results might change with changing network sizes and mean degrees. Data and plots were produced in the same way as presented in the main text.

N = 10,000

We simulated networks over instances of each network class which had mean degrees of 2–8, 10, and 13 keeping the number of nodes = 10,000 to be consistent with the analysis done in the main text. The range of transmission probabilities stayed the same (0.1 – 0.5, by 0.05) except in the mean degree = [2–4] cases, where a larger range was used due to the high epidemic threshold. Below we present two plots summarizing these data.



Supplementary Figure 6: The degree distributions for all the networks (N = 10000) divided up by mean degree (outer row) and by network class (outer column). The max x-axis value of 100 in each plot does not reflect the maximum degree, but was used for comparison purposes.

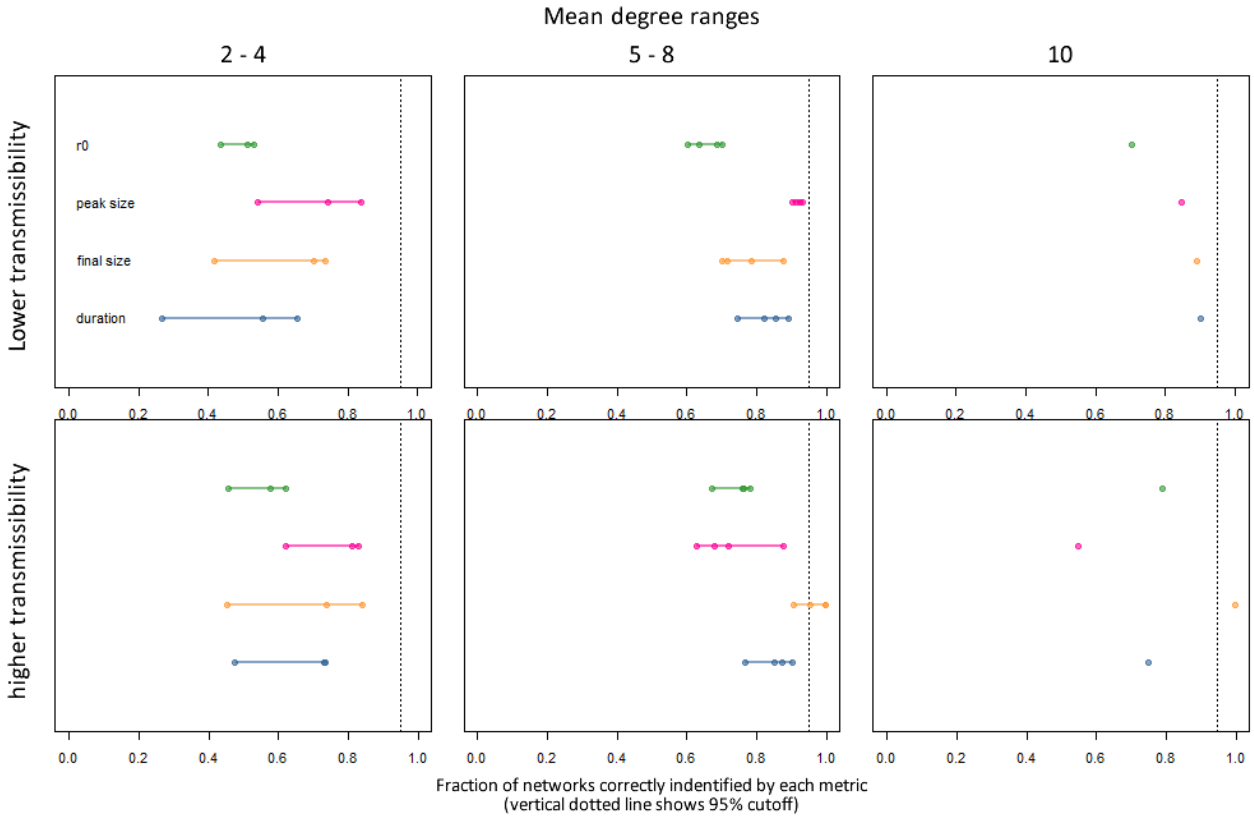


Supplementary Figure 7: This plot shows roughly how well each metric works as a network type classifier. There are low, medium, and high mean degree ranges on the x-axis and lower (0.1 – 0.25) and higher (0.3+) transmissibilities on the y-axis. The fractions of networks correctly classified by each data type are shown in color and are labeled in the top left hand plot.

N = 500

Using N = 500, we re-ran epidemic simulations for all mean degrees (excluding 13, because a scale-free network with those properties could not be found). For each network and transmission probability pair, 2000 simulations were completed. Their degree distributions were similar to those in Figure 6 above, but on a smaller scale. The range of transmission probabilities stayed the same (0.1 – 0.5, by 0.05) except when the mean degree

was between 2–4, where a larger range was used due to the high epidemic threshold. It can be seen in the plot immediately below, which summarizes all of these analyses, and in subsequent mean degree-specific plots that there seems to be a significant small network size effect:



Supplementary Figure 8: This plot shows roughly how well each metric works as a network type classifier. There are low, medium, and high mean degree ranges on the x-axis and lower (0.1 – 0.25) and higher (0.3+) transmissibilities on the y-axis. The fractions of networks correctly classified by each data type are shown in color and are labeled in the top left hand plot

Empirical results

Known network summary statistics

This table displays some statistics about the four known (empirical) networks used in the study:

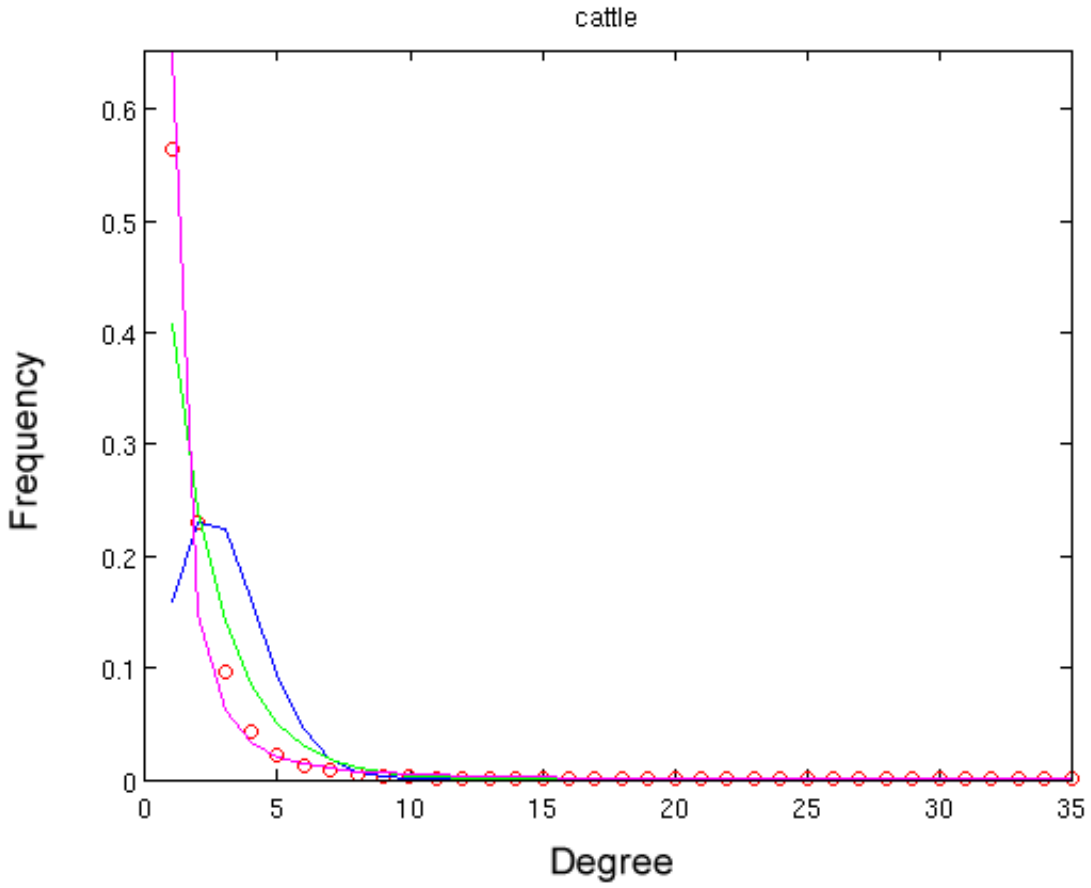
Network name	Number of nodes	Mean degree	Transitivity	Assortativity
Cattle	37787	3.04	0.002	-0.29
Adolescent Sexual	278	2.04	0.006	-0.24
Urban	12729	16.04	0.07	0.19
School	661	3.76	0.08	0.02

Maximum likelihood estimation

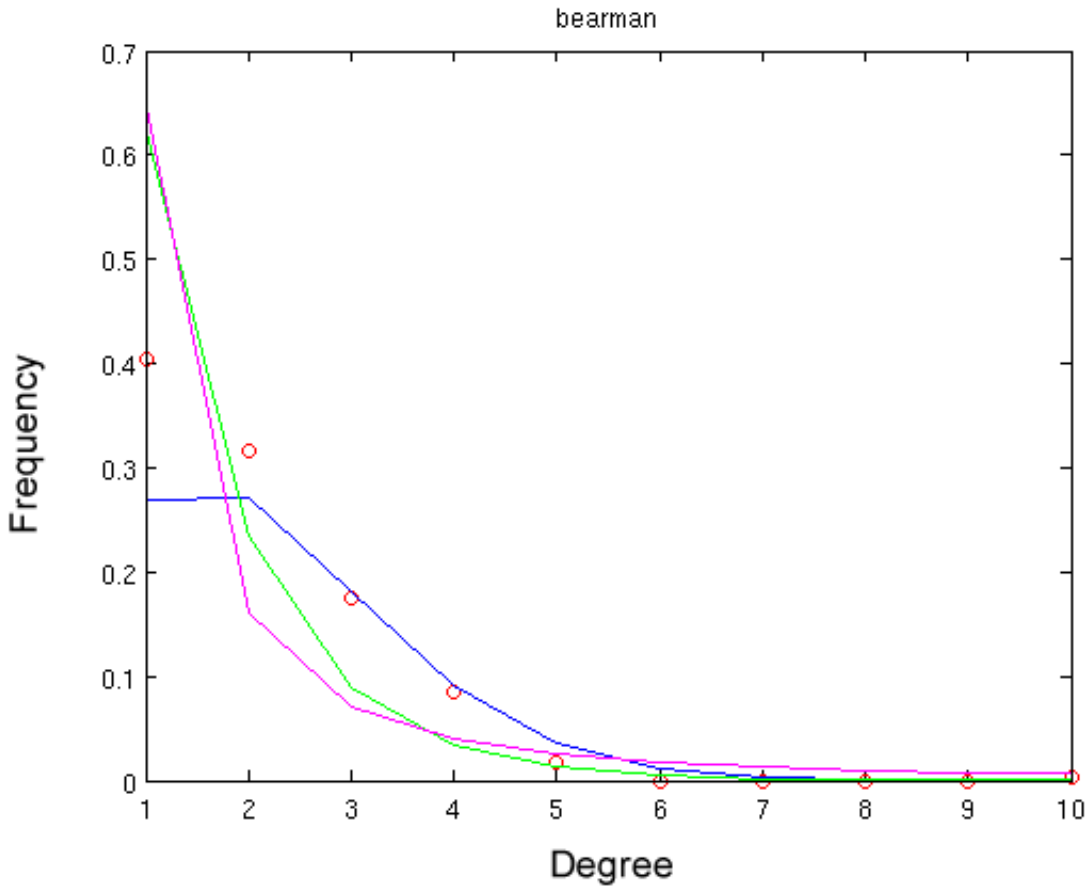
Parameters for Poisson, exponential, and power-law distributions were fit using maximum likelihood to the degree distribution of the each Cattle, Adolescent Sexual, School, and Urban network (detailed in the main text) in order to get an idea of how the network ought to be classifier using our framework. Goodness-of-fit was measured using the Kullback-Leibler divergence between the actual degree distribution and the Poisson, exponential, and power-law distributions parameterized using the maximum likelihood estimates.

Table 1: KL-divergences between best-fit distributions and actual degree distributions of each empirical network

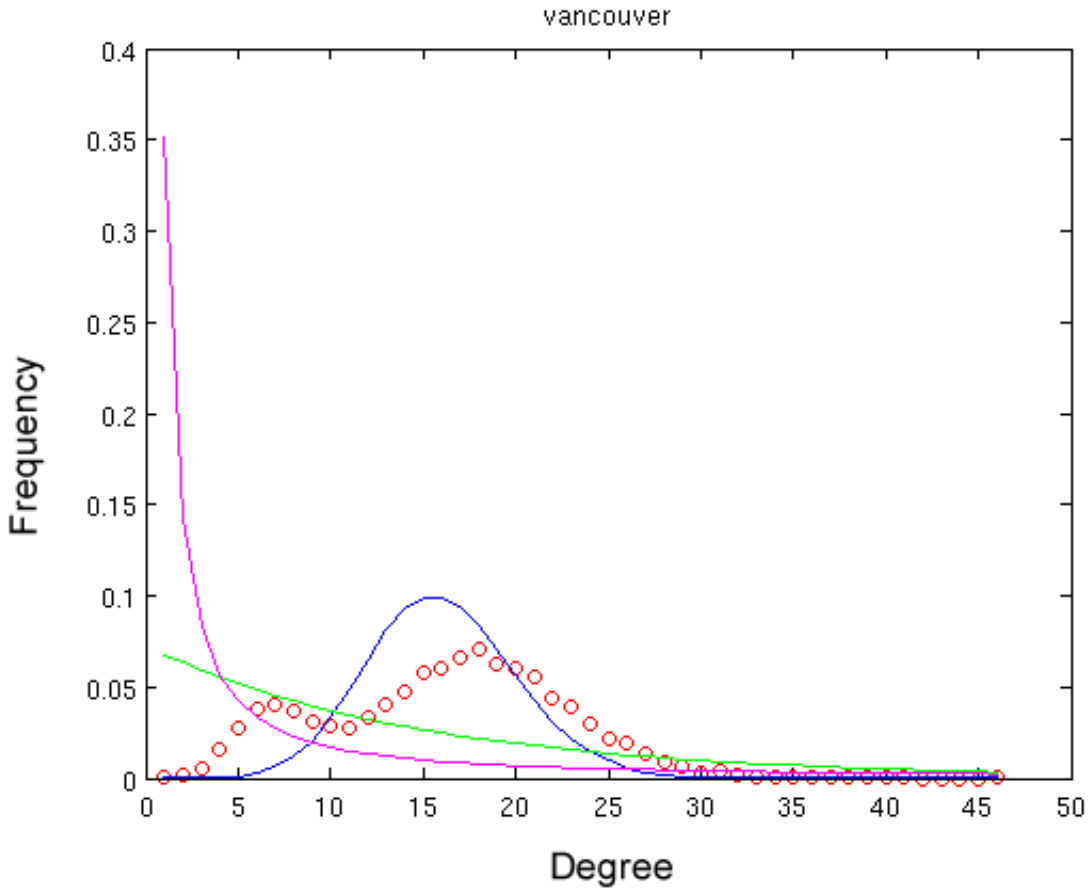
	KL divergence		
	Poisson	exponential	scale-free
Cattle	1.1586	0.5347	**0.0586**
Adolescent Sexual	0.2022	**0.1361**	0.2362
Urban	**0.406**	0.4849	1.3383
School	**0.0841**	0.1002	0.4068



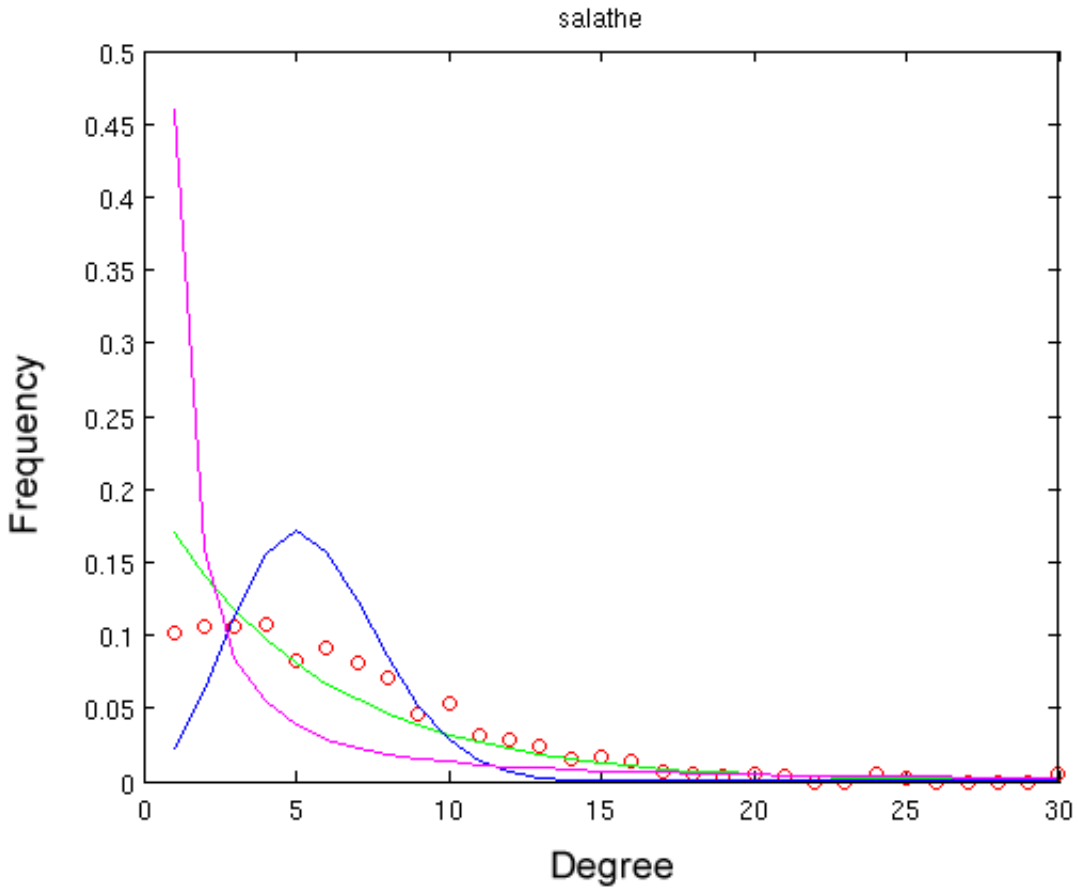
Supplementary Figure 9: MLE fit for the cattle network. The red dots are indicate the actual degree distribution. Best fit Poisson, exponential, and scale-free distributions are indicated by the blue, green, and pink lines, respectively.



Supplementary Figure 10: MLE fit for the Adolescent Sexual network. The red dots are indicate the actual degree distribution. Best fit Poisson, exponential, and scale-free distributions are indicated by the blue, green, and pink lines, respectively.



Supplementary Figure 11: MLE fit for the Urban network. The red dots are indicate the actual degree distribution. Best fit Poisson, exponential, and scale-free distributions are indicated by the blue, green, and pink lines, respectively.



Supplementary Figure 12: MLE fit for the school network. The red dots are indicate the actual degree distribution. Best fit Poisson, exponential, and scale-free distributions are indicated by the blue, green, and pink lines, respectively.

Network model comparisons

The log-Bayes Factors (log-BF) are shown for each of the four analyses, comparing the likelihoods of the different network classes. The numbers in each cell represent the 50th (5th, 95th) quantiles. The tables below should complement Figure 3-3 from the main text.

The log-Bayes factor comparisons for Cattle network simulations:

Final size

	Poisson	Expo.	S.F.
Poisson	0	0.00 (0.00, 0.00)	-20.72 (-20.72, -20.72)
Expo.	0.00 (0.00, 0.00)	0	-20.72 (-20.72, -20.72)
S.F.	20.72 (20.72, 20.72)	20.72 (20.72, 20.72)	0

Peak size

	Poisson	Expo.	S.F.
Poisson	0	0.00 (0.00, 0.00)	-20.72 (-20.72, -20.72)
Expo.	0.00 (0.00, 0.00)	0	-20.72 (-20.72, -20.72)
S.F.	20.72 (20.72, 20.72)	20.72 (20.72, 20.72)	0

Duration

	Poisson	Expo.	S.F.
Poisson	0	-0.75 (-1.45, -0.17)	0.00 (-1.32, 19.52)
Expo.	0.75 (0.17, 1.45)	0	0.61 (-1.15, 20.51)
S.F.	0.00 (-19.52, 1.32)	-0.61 (-20.51, 1.15)	0

The log-Bayes factor comparisons for Adolescent sexual network simulations:

Final size

	Poisson	Expo.	S.F.
Poisson	0	-20.03 (-20.30, -0.54)	-19.90 (-20.28, -0.46)
Expo.	20.03 (0.54, 20.30)	0	0.08 (-0.65, 0.64)
S.F.	19.90 (0.46, 20.28)	-0.08 (-0.64, 0.65)	0

Peak size

	Poisson	Expo.	S.F.
Poisson	0	-19.92 (-20.14, -0.48)	-19.98 (-20.32, -0.41)
Expo.	19.92 (0.48, 20.14)	0	-0.05 (-0.70, 0.23)
S.F.	19.98 (0.41, 20.32)	0.05 (-0.23, 0.70)	0

Duration

	Poisson	Expo.	S.F.
Poisson	0	-4.23 (-20.10, 1.73)	-3.95 (-20.30, 1.33)
Expo.	4.23 (-1.73, 20.10)	0	-0.01 (-0.63, 0.37)
S.F.	3.95 (-1.33, 20.30)	0.01 (-0.37, 0.63)	0

The log-Bayes factor comparisons for School network simulations:

Final size

	Poisson	Expo.	S.F.
Poisson	0	0.17 (-0.21, 20.72)	0.89 (-0.54, 20.72)
Expo.	-0.17 (-20.72, 0.21)	0	0.00 (-0.33, 19.20)
S.F.	-0.89 (-20.72, 0.54)	0.00 (-19.20, 0.33)	0

Peak size

	Poisson	Expo.	S.F.
Poisson	0	-0.04 (-0.04, 4.11)	-0.15 (-0.15, 20.71)
Expo.	0.04 (-4.11, 0.04)	0	-0.11 (-0.11, 18.70)
S.F.	0.15 (-20.71, 0.15)	0.11 (-18.70, 0.11)	0

Duration

	Poisson	Expo.	S.F.
Poisson	0	0.09 (-1.24, 2.42)	1.04 (-0.88, 20.21)
Expo.	-0.09 (-2.42, 1.24)	0	0.80 (-2.74, 20.47)
S.F.	-1.04 (-20.21, 0.88)	-0.80 (-20.47, 2.74)	0

The log-Bayes factor comparisons for Urban network simulations:

Final size

	Poisson	Expo.	S.F.
Poisson	0	0.53 (-0.37, 20.72)	0.48 (-0.44, 20.72)
Expo.	-0.53 (-20.72, 0.37)	0	-0.04 (-0.06, 0.14)
S.F.	-0.48 (-20.72, 0.44)	0.04 (-0.14, 0.06)	0

Peak size

	Poisson	Expo.	S.F.
Poisson	0	-0.09 (-0.09, 20.72)	-0.14 (-0.14, 20.72)
Expo.	0.09 (-20.72, 0.09)	0	-0.06 (-0.06, 0.38)
S.F.	0.14 (-20.72, 0.14)	0.06 (-0.38, 0.06)	0

Duration

	Poisson	Expo.	S.F.
Poisson	0	1.09 (-0.37, 20.72)	1.19 (-0.44, 20.72)
Expo.	-1.09 (-20.72, 0.37)	0	-0.01 (-0.08, 0.25)
S.F.	-1.19 (-20.72, 0.44)	0.01 (-0.25, 0.08)	0

Analytical likelihood method

Recently, Andre-Noel et al. (Noël, Davoudi, Brunham, Dubé, & Pourbohloul, 2009) proposed an analytical percolation-based framework to predict the time progression of disease on finite-size networks. With knowledge of the degree distribution of the contact network, a transmission probability, *T*, and assuming a constant generation time, this method generates $\Psi(m,s,g)$, the probability that there are *m* currently infected individuals and *s* cumulatively infected individuals at generation *g*.

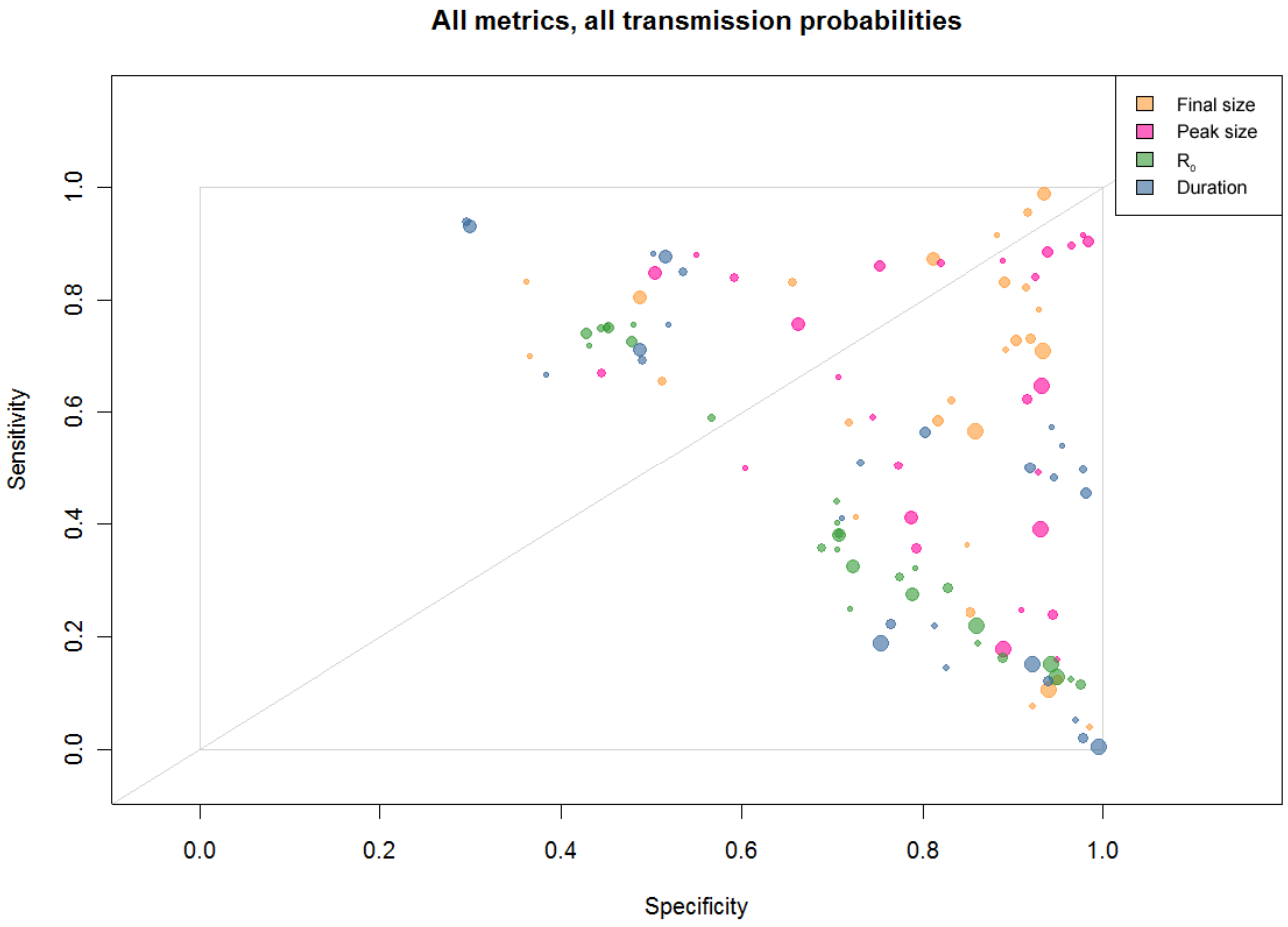
The distribution for the number of new infections at generation *g* is thus $\nu(m,g) = \sum_{\{s\}}\Psi(m,s,g)$. The likelihood functions for the epidemiological measures can then be calculated as:

$$P\left(R_{\{0\}} = r \middle| M_{\{i\}}\right) = \sum_{\{m\}} \nu(rm,3)\nu(m,2)$$

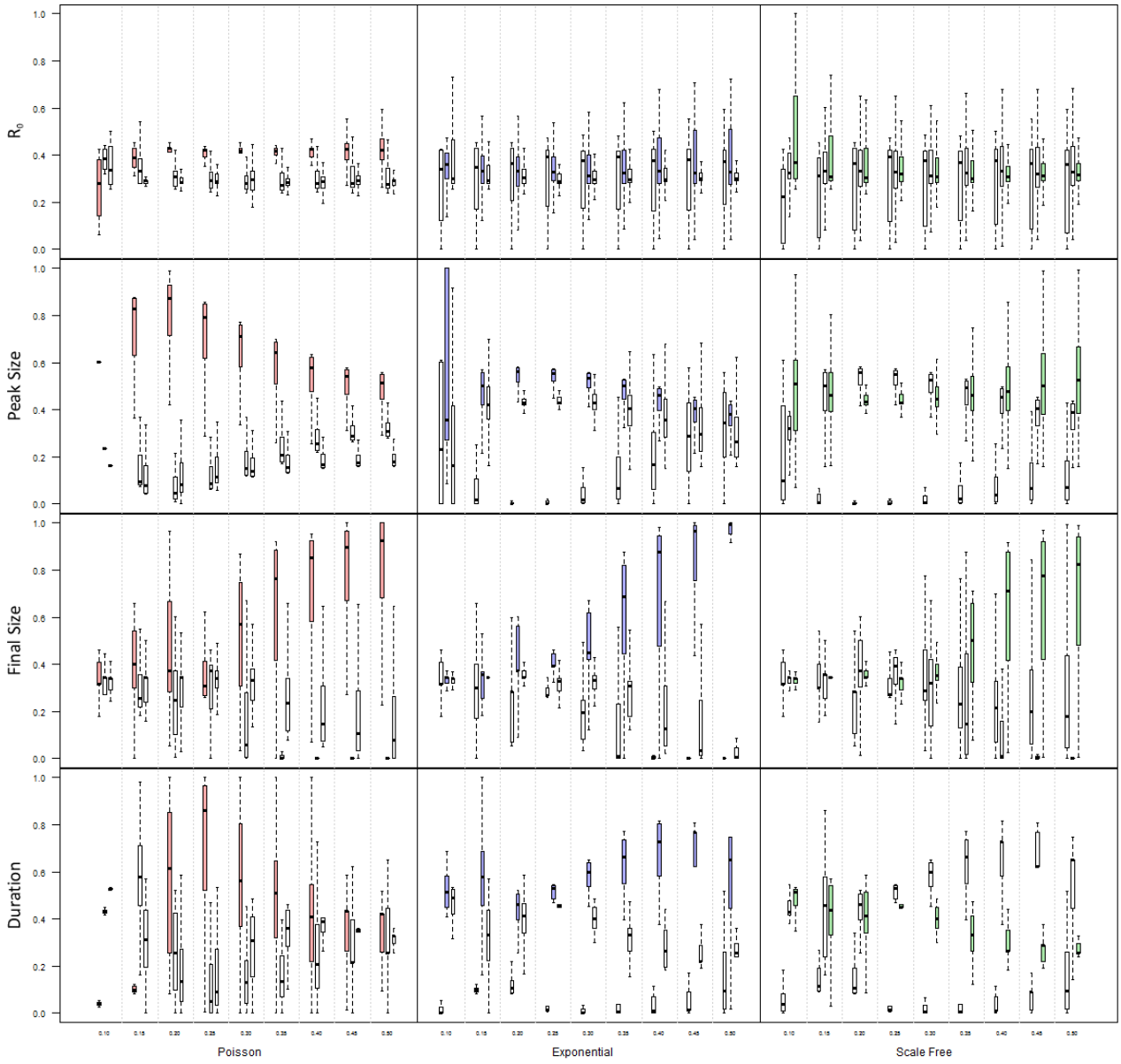
$$P\left(\sigma = S \middle| M_{\{i\}}\right) = \sum_{\{m\}} \Psi(m,s = S, g = g*)$$

$$P\left(\delta \leq d \middle| M_{\{i\}}\right) = \sum_{\{s\}} \Psi(m = 0, s, g = d)$$

where *g** is the final generation. ρ is not shown above as its probability distribution is approximated numerically from Ψ. In the two figures immediately below, we show the results of the Bayesian analysis using these likelihoods and find that they are also comparable to the results from simulations found in main text Figures 3-1 and 3-2.



Supplementary Figure 13: Sensitivity vs. Specificity results from simulated networks and epidemic data and model likelihood calculated using the aforementioned Andre-Noel framework. This is comparable to Figure 3-1 from the main text.

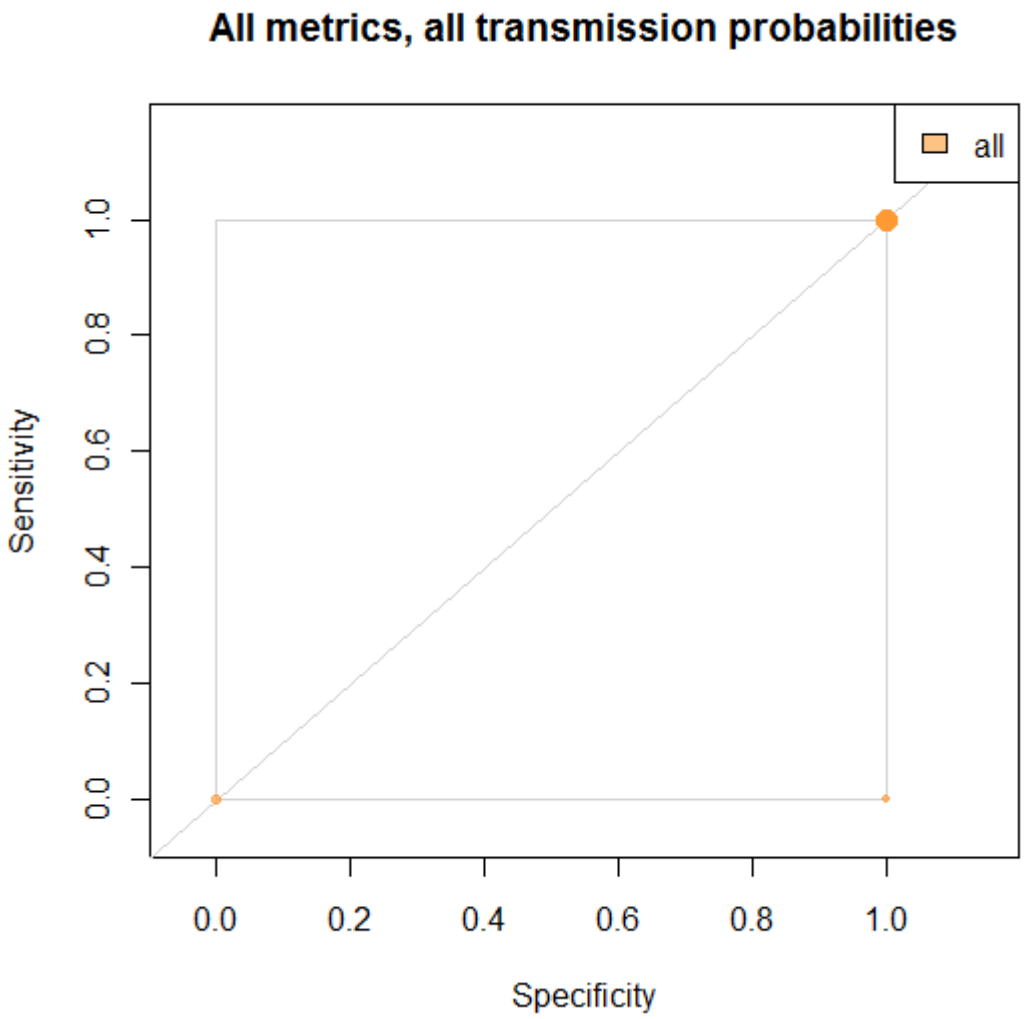


Supplementary Figure 14: Likelihood results using probabilities generated by the Andre-Noel framework for the same data used to construct the previous figure. This plot should be compared to Figure 3-2 from the main text.

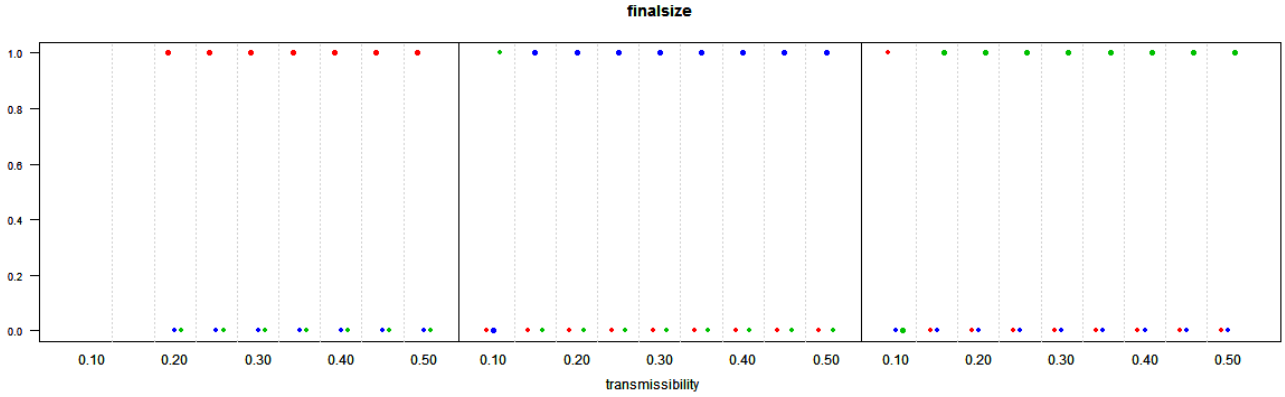
Preliminary Analysis for Combined Metrics

The goal of the framework described in the main text is to classify heterogeneity in population contact structures based on epidemiological measures available in public health settings. Thus, it was important in this context, to explore individual measures independently, in the event that all data are not available or are uncertain. Based on our analysis, we now understand that not all epidemiological measures are reliable for detecting population contact structure. In the event that multiple data types are available, however, we consider an alternative analysis which combines all four metrics (R_0, ρ, σ, δ) to make network classification predictions. This preliminary analysis is based on classification and regression trees (CART) (Breiman, Friedman, Olshen, & Stone, 1984), a recursive partitioning method which builds classification trees for predictor variables (i.e., R_0, ρ, σ, δ) which are used to predict network type. The R package 'Rpart' was used to create decision trees using half of the synthetic data set (same one from the main text) for each transmission probability using a complexity parameter of 0.01 and a maximum tree depth of 30. These trees were then used to get probabilities for each network class given all four epidemiological measures. Interestingly, the CART analyses indirectly confirm that R_0 is a relatively weak measure, while the others are not.

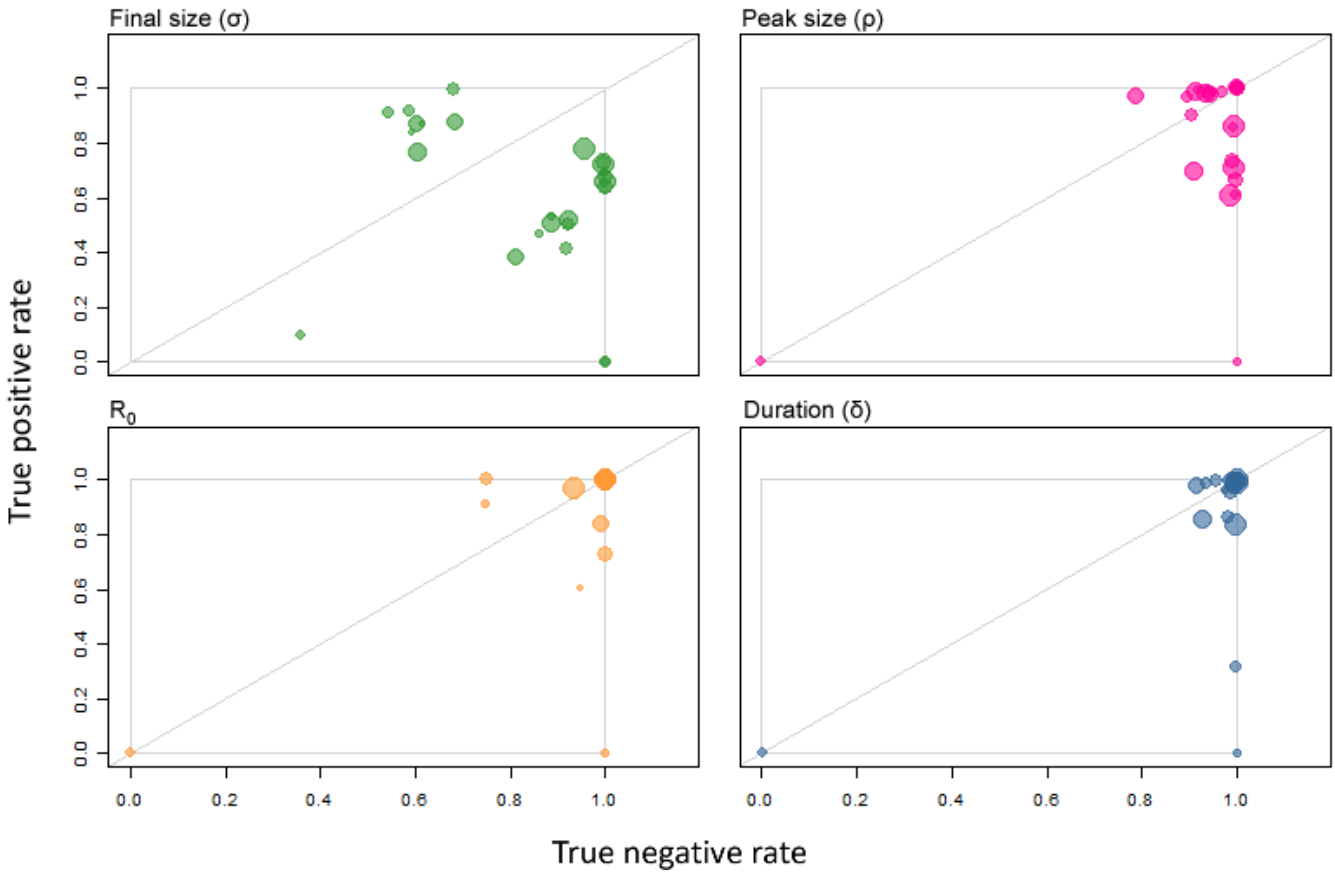
The results from the CART analyses are presented in the four figures below which are comparable to Figures 3-1 and 3-2 from the main text. Using Rpart, an implementation of CART in the R software environment, we analyzed how well all metrics, when combined, predict network type. We find that using them all together results in excellent predictions, even when R_0 is excluded altogether (results not shown). We also show the results when each measure is analyzed individually, as we did in the main text. The results for this show a similar pattern to those in the main text.



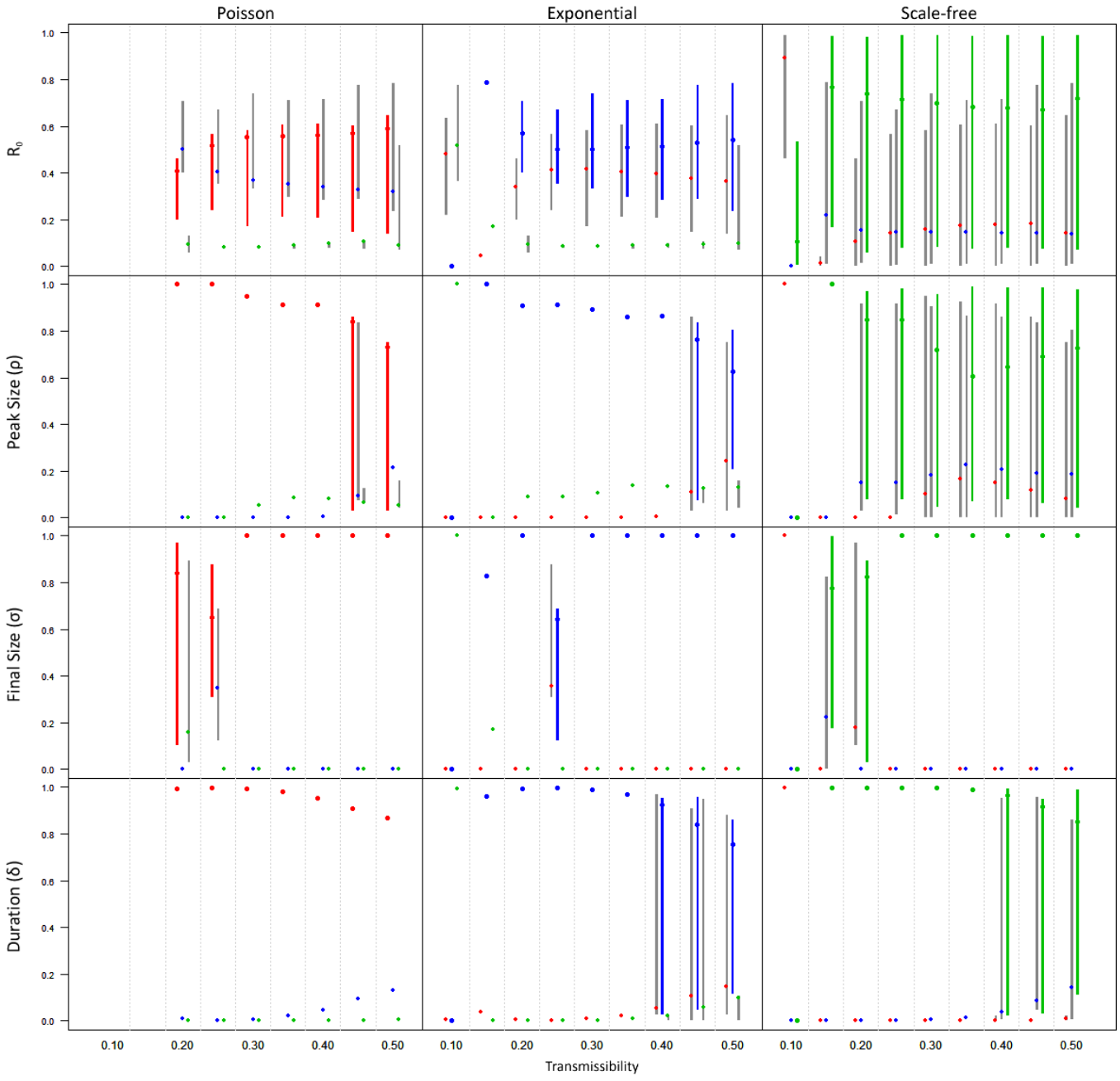
Supplementary Figure 15: Sensitivity vs. Specificity using CART analysis. Results are broken down by transmission probability, where larger points represent larger transmission probabilities.



Supplementary Figure 16: Likelihood results for all metrics taken together over transmission probabilities ranging from 0.1 to 0.5 using CART. This figure is comparable to main text Figure 3-2 and similar Figures above.



Supplemental Figure 17: Sensitivity (y-axis) versus Specificity (x-axis) results when each metric is analyzed independently using CART. This figure is comparable to Figure 3-1 from the main text.



Supplemental Figure 18: Likelihood results using divided up by metric where probabilities were again determined by CART analysis using each metric individually. This figure is in a similar format as Figure 3-2 from the main text.

References

Breiman, L., Friedman, J., Olshen, R., & Stone, C. (1984). *CART: Classification and Regression Trees* (1st ed., p. 368). Belmont, CA: Chapman and Hall/CRC.

Molloy, M., & Reed, B. (1995). A critical point for random graphs with a given degree sequence. *Random Structures & Algorithms*, *6*(2-3), 161-180. doi:10.1002/rsa.3240060204

Noël, P.-A., Davoudi, B., Brunham, R., Dubé, L., & Pourbohloul, B. (2009). Time evolution of epidemic disease on finite and infinite networks. *Physical Review E*, *79*(2), 1-14. doi:10.1103/PhysRevE.79.026101

Taylor, R. (1981). Contrained switchings in graphs. In K. McAvaney (Ed.), *Combinatorial Mathematics VIII* (pp. 314-336). Springer Berlin Heidelberg. doi:10.1007/BFb0091828