**Main Manuscript for**

Hospital environmental reservoirs could shape nosocomial transmission and implications for bacterial communities coexistance

Jaime Cascante Vega1, Rami Yarii1, Sen Pei1, Jeffrey Shaman1,3

1. Department of Environmental Health Sciences, Mailman School of Public Health, Columbia University, New York, NY, USA.

3. Columbia Climate School, Columbia University, New York, NY, USA.

\* Sen Pei, Jeffrey Shaman

**Email:**  sp3449@cumc.columbia.edu, jls106@cumc.columbia.edu

**Author Contributions:** J.C.V. and J.S. conceived the study and managed the project. J.C.V. performed the analysis. J.C.V., R.Y., S.P. and J.S. analyzed the results. J.C.V drafted the manuscript. All authors revised and reviewed the manuscript.

**Competing Interest Statement:** J.S. and Columbia University disclose partial ownership of SK Analytics. J.S. discloses consulting for BNI. All other authors declare no competing interests.

**Classification:** Major: Biological Sciences. Minor: Biophysics and Computational Biology.

**Keywords:** epidemiology, micro-organisms, antimicrobial resistance, environmental transmission, pathogenic bacteria

**This PDF file includes:**

Main Text

Figures 1 to 6

Tables 1 to 2

**Abstract (244/250 words)**

Antimicrobial-resistant organisms (AMROs) are a major threat to public health. These organisms increase mortality, hospital length of stay, and healthcare associated costs. Many AMROs are present in both hospitals and the broader community; however, information on AMRO transmission in both settings is limited. In this work, we leverage electronic health records from a major New York City hospital system collected during 2020-2021 to support simulation of importation and transmission, and inference of critical epidemiological quantities for eight pathogens. We develop an agent-based model to simulate admission, transfer, and discharge of patients in the hospital system, AMRO importation from the community, and patient-to-patient transmission of AMROs. The model is coupled with a Bayesian inference algorithm to estimate importation and nosocomial transmission rates. We evaluate parameter identifiability for this model-inference system and then apply the framework to estimate both quantities for seven prevalent organisms: *Escherichia coli, Klebsiella pneumoniae, Pseudomonas aeruginosa, Staphylococcus aureus* (both sensitive, MSSA, and resistant, MRSA, phenotypes), *Candida albicans, Staphylococcus epidermidis*, and *Enterococcus faecalis*. The parameter estimates reveal substantially higher community prevalence (i.e., importation rate) for *E. coli* and *K. pneumoniae* than the other six organisms. Nosocomial transmission rates are found to be highest for *P. aeruginosa* and MSSA. This work highlights how fine-scale patient data can support dynamic modeling and estimation of epidemiological properties of microorganisms. Evaluation of the community prevalence and transmission potential for different pathogens could ultimately support the development of in-hospital control measures that limit the spread of these pathogens.

**Significance Statement (118/120 words)**

**Main Text**

**Introduction**

Pathogenic bacteria asymptomatically colonize individuals as they naturally inhabit the human body across different body sites, i.e. are commensals with specialized niches. This have been referred as biogeography of bacterial comunities in tbuhe human body. Researchers widely had recognize this and focussed research efforts towards elucidate diversity and distribution of the human microbiota. In hospital settings once bacteria colonize typically sterile body sites the human host cascade a series of immune response towards treating the infection which increase the length of stay, and are a major contributor to mortality. Despite this fact understanding in-hospital (nosocomial) transmission is tremendously challenging due to the different modes of transmission contributing to the burden, heterogeneous control in the hospital systems, heterogeneous patient traffic (admission, discharges, length of stay and transfers) across wards in different facilities comprising the hospital system, as well as imperfect observation of carriage and infection. Despite these difficulties, recent work suggests patient traffic determines nosocomial transmission, suggesting that bacterial pathogens have similar levels of transmission as well as possible similar modes.

Nosocomial transmission is thought to be a contribution of patient-to-patient, environmental, and health-care-workers associated transmission. In this study we compare different approaches to model transmission dynamics of patient-to-patient and environmental transmission. We consider first a model which tracks the environmental force of infection assuming carriers are shedding bacteria at a constant rate to the environment, increasing the numbers of bacteria in surfaces, and thus increasing the force of infection. This environmental force of infection is also cleared at a constant rate reflecting cleaning of surfaces in the hospital setting. Second, recognizing that a more specified model would require explicit modeling of the bacterial communities in the environment we use a logistic growth to model a community of a bacterium in hospital surfaces. Third, is likely that bacterial live in multispecies communities in hospital surfaces. In consequence we design a dynamic model that explictly include this mutispecies communities co-existance. At last we compared the transmission dynamics, i.e. the hospital prevalence as well as trade-off between the patient-to-patient and environmental transmission rates in the different models.

**Models of nosocomial transmission** **duet o person-to-person contact and environmental reservoirs**

asdads

**Results**

### *Empirical patterns and heterogeneity of microorganism burden*

### *Identifiability: synthetic simulations and inference*

### *Inference using real data*

**Discussion**

**Materials and Methods**

### *Overview*

To estimate key epidemiological characteristics of AMROs, we developed an ABM to simulate the dynamics of these organisms in hospital settings. The model was informed by patient hospitalization and culture data from electronic healthcare records collected between February 1 2020 and February 28 2021. We coupled the ABM with a Bayesian inference algorithm and, using simulated outbreaks, validated the ability of this ABM-inference system to identify importation and nosocomial transmission rates. We then assimilated the real-world lab-confirmed positive case data and estimated the importation and nosocomial transmission rates for eight co-circulating organisms. Using these estimated parameters and the ABM, we were able to reproduce the time series of positive cases for six clusters of wards in the hospital system. The estimated importation rates and nosocomial transmission rates were compared for the eight organisms.

### *Data*

### *The transmission models with environmental force of infection*

### *Hospital-level Observational model*

### *Inference*

### *Identifiability: Synthetic simulations and inference*

### *Inference using real-world data*

**Acknowledgments**

This study was supported by funding from the Centers for Disease Control and Prevention U01CK000592 and 75D30122C14289. We thank Matteo Perini and the CDC Modeling Infectious Diseases in Healthcare Network (MInD – Healthcare) for comments and discussions. We also thank Rami Yaari for partially reproducing the results.

**References**

1. de Kraker, M. E. A. & Lipsitch, M. Burden of Antimicrobial Resistance: Compared to What? *Epidemiologic Reviews* **43**, 53–64 (2022).

2. Murray, C. J. *et al.* Global burden of bacterial antimicrobial resistance in 2019: a systematic analysis. *The Lancet* **399**, 629–655 (2022).

3. Lehtinen, S. Co-colonisation and coexistence. *Nat Ecol Evol* **3**, 334–335 (2019).

4. Pham, T. M., Kretzschmar, M., Bertrand, X., Bootsma, M., & on behalf of COMBACTE-MAGNET Consortium. Tracking Pseudomonas aeruginosa transmissions due to environmental contamination after discharge in ICUs using mathematical models. *PLoS Comput Biol* **15**, e1006697 (2019).

5. Bergstrom, C. T., Lo, M. & Lipsitch, M. Ecological theory suggests that antimicrobial cycling will not reduce antimicrobial resistance in hospitals. *Proc Natl Acad Sci U S A* **101**, 13285 (2004).

6. Bonhoeffer, S., Lipsitch, M. & Levin, B. R. Evaluating treatment protocols to prevent antibiotic resistance. *Proc Natl Acad Sci USA* **94**, 12106 (1997).

7. Lipsitch, M., Bergstrom, C. T. & Levin, B. R. The epidemiology of antibiotic resistance in hospitals: Paradoxes and prescriptions. *Proc Natl Acad Sci USA* **97**, 1938 (2000).

8. Pei, S., Morone, F., Liljeros, F., Makse, H. & Shaman, J. L. Inference and control of the nosocomial transmission of methicillin-resistant Staphylococcus aureus. *eLife* **7**, e40977 (2018).

9. Paul, P., Slayton, R. B., Kallen, A. J., Walters, M. S. & Jernigan, J. A. Modeling Regional Transmission and Containment of a Healthcare-associated Multidrug-resistant Organism. *Clin Infect Dis* **70**, 388–394 (2020).

10. Cooper, B. S. *et al.* Methicillin-resistant Staphylococcus aureus in hospitals and the community: Stealth dynamics and control catastrophes. *Proceedings of the National Academy of Sciences* **101**, 10223–10228 (2004).

11. Pei, S., Liljeros, F. & Shaman, J. Identifying asymptomatic spreaders of antimicrobial-resistant pathogens in hospital settings. *Proc. Natl. Acad. Sci. U.S.A.* **118**, e2111190118 (2021).

12. Lipsitch, M. & Samore, M. H. Antimicrobial Use and Antimicrobial Resistance: A Population Perspective. *Emerg. Infect. Dis.* **8**, 347–354 (2002).

13. Austin, D. J., Kristinsson, K. G. & Anderson, R. M. The relationship between the volume of antimicrobial consumption in human communities and the frequency of resistance. *Proceedings of the National Academy of Sciences* **96**, 1152–1156 (1999).

14. Austin, D. J., Bonten, M. J. M., Weinstein, R. A., Slaughter, S. & Anderson, R. M. Vancomycin-resistant enterococci in intensive-care hospital settings: Transmission dynamics, persistence, and the impact of infection control programs. *Proceedings of the National Academy of Sciences* **96**, 6908–6913 (1999).

15. Uhlemann, A.-C., Otto, M., Lowy, F. D. & DeLeo, F. R. Evolution of community- and healthcare-associated methicillin-resistant Staphylococcus aureus. *Infection, Genetics and Evolution* **21**, 563–574 (2014).

16. Smith, D. L., Levin, S. A. & Laxminarayan, R. Strategic interactions in multi-institutional epidemics of antibiotic resistance. *Proc Natl Acad Sci U S A* **102**, 3153 (2005).

17. the EuSCAPE Working Group *et al.* Epidemic of carbapenem-resistant Klebsiella pneumoniae in Europe is driven by nosocomial spread. *Nat Microbiol* **4**, 1919–1929 (2019).

18. Rosvall, M. & Bergstrom, C. T. Maps of random walks on complex networks reveal community structure. *Proceedings of the National Academy of Sciences* **105**, 1118–1123 (2008).

19. Gelman, A., Simpson, D. & Betancourt, M. The Prior Can Often Only Be Understood in the Context of the Likelihood. *Entropy* **19**, 555 (2017).

20. Kennedy, L., Simpson, D. & Gelman, A. The experiment is just as important as the likelihood in understanding the prior: A cautionary note on robust cognitive modelling. 12.

21. Wieland, F.-G., Hauber, A. L., Rosenblatt, M., Tönsing, C. & Timmer, J. On structural and practical identifiability. *Current Opinion in Systems Biology* **25**, 60–69 (2021).

22. Larramendy, S. *et al.* Risk Factors of Extended-Spectrum Beta-Lactamases-Producing Escherichia coli Community Acquired Urinary Tract Infections: A Systematic Review. *IDR* **Volume 13**, 3945–3955 (2020).

23. Diggle, S. P. & Whiteley, M. Microbe Profile: Pseudomonas aeruginosa: opportunistic pathogen and lab rat: This article is part of the Microbe Profiles collection. *Microbiology* **166**, 30–33 (2020).

24. Reuter, S., Sigge, A., Wiedeck, H. & Trautmann, M. Analysis of transmission pathways of Pseudomonas aeruginosa between patients and tap water outlets\*: *Critical Care Medicine* **30**, 2222–2228 (2002).

25. Harris, A. D. *et al.* How important is patient-to-patient transmission in extended-spectrum β-lactamase Escherichia coli acquisition. *American Journal of Infection Control* **35**, 97–101 (2007).

26. Lu, D., Aleta, A. & Moreno, Y. Assessing the Risk of Spatial Spreading of Diseases in Hospitals. *Front. Phys.* **10**, 882314 (2022).

27. Miller, M. *et al.* Staphylococcus aureus in the Community: Colonization Versus Infection. *PLoS ONE* **4**, e6708 (2009).

28. Sakr, A., Brégeon, F., Mège, J.-L., Rolain, J.-M. & Blin, O. Staphylococcus aureus Nasal Colonization: An Update on Mechanisms, Epidemiology, Risk Factors, and Subsequent Infections. *Front. Microbiol.* **9**, 2419 (2018).

29. Anderson, J. L. An Ensemble Adjustment Kalman Filter for Data Assimilation. *Mon. Wea. Rev.* **129**, 2884–2903 (2001).

30. Arulampalam, M. S., Maskell, S., Gordon, N. & Clapp, T. A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking. *IEEE Trans. Signal Process.* **50**, 174–188 (2002).

31. Subramanian, R., He, Q. & Pascual, M. Quantifying asymptomatic infection and transmission of COVID-19 in New York City using observed cases, serology, and testing capacity. *PNAS* **118**, (2021).

32. Romeo-Aznar, V., Picinini Freitas, L., Gonçalves Cruz, O., King, A. A. & Pascual, M. Fine-scale heterogeneity in population density predicts wave dynamics in dengue epidemics. *Nat Commun* **13**, 996 (2022).

33. Santos-Vega, M. *et al.* The neglected role of relative humidity in the interannual variability of urban malaria in Indian cities. *Nat Commun* **13**, 533 (2022).

34. Li, R. *et al.* Substantial undocumented infection facilitates the rapid dissemination of novel coronavirus (SARS-CoV-2). *Science* **368**, 489–493 (2020).

35. Pei, S. & Shaman, J. Aggregating forecasts of multiple respiratory pathogens supports more accurate forecasting of influenza-like illness. *PLoS Comput Biol* **16**, e1008301 (2020).

36. Pei, S., Teng, X., Lewis, P. & Shaman, J. Optimizing respiratory virus surveillance networks using uncertainty propagation. *Nat Commun* **12**, 222 (2021).

37. Pei, S., Cane, M. A. & Shaman, J. Predictability in process-based ensemble forecast of influenza. *PLoS Comput Biol* **15**, e1006783 (2019).

38. Park, S. W. *et al.* Epidemiological dynamics of enterovirus D68 in the United States and implications for acute flaccid myelitis. *Sci. Transl. Med.* **13**, eabd2400 (2021).

39. Yang, W. *et al.* Estimating the infection-fatality risk of SARS-CoV-2 in New York City during the spring 2020 pandemic wave: a model-based analysis. *The Lancet Infectious Diseases* **21**, 203–212 (2021).

40. Ukawuba, I. & Shaman, J. Inference and dynamic simulation of malaria using a simple climate-driven entomological model of malaria transmission. *PLoS Comput Biol* **18**, e1010161 (2022).

41. Ionides, E. L., Breto, C. & King, A. A. Inference for nonlinear dynamical systems. *Proceedings of the National Academy of Sciences* **103**, 18438–18443 (2006).

42. Ionides, E. L., Nguyen, D., Atchadé, Y., Stoev, S. & King, A. A. Inference for dynamic and latent variable models via iterated, perturbed Bayes maps. *Proc Natl Acad Sci USA* **112**, 719–724 (2015).

43. Harris, C. R. *et al.* Array programming with NumPy. *Nature* **585**, 357–362 (2020).

44. Virtanen, P. *et al.* SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nat Methods* **17**, 261–272 (2020).

**Figures and Tables**Calendar

Description automatically generated with low confidence

**Figure 1. Empirical colonization of microbial organisms, hospital admissions and testing.** **A)** Incident observations for microorganisms of interest, from left to right and top to bottom: *E. coli*, *K. pneumoniae*, *P. aeruginosa*, *MSSA*, *S. epidermis*, *Candida albicans*, *MRSA*, *E. faecalis*. **B)** Numbers of in-hospital patients (red) and admitted patients (blue) during the study period at daily resolution (weekly variability is evident). **C)** Numbers of in-hospital patients normalized by ward size (average occupancy per day during the study period); blue lines show the 10 most populated wards, green lines the 10 least populated, and the remaining wards are shown in gray in the background **D)** Heatmap showing the number of weekly cultures in each ward during the study period. **E)** Heatmap plot showing the number of patients admitted weekly to each ward during the study period.

Graphical user interface

Description automatically generated

**Figure 2. A)** Adjacency matrix of transfers between wards during the study period. **B)** Adjacency matrix of transfers between clusters (aggregation of wards) during the study period. Color bar is in Log10 scale; darker colors indicate greater movement of individuals between each pair of wards or ward clusters. **C)** Incident colonization for the 8 study species; line color designates the cluster. **D)** Heatmap plot showing the number of weekly cultures identified in each ward cluster during the study period. **E)** Heatmap showing the number of patients admitted weekly to each ward cluster during the study period.

Text

Description automatically generated with low confidence

**Figure 3. Posterior parameter estimates.** Estimates for the **A)** importation rate and **B)** nosocomial transmission rate. Violin plots show the posterior distribution an estimates for the last iteration of the IF-EAKF (see Methods section) for 3 different runs of the model-inference system. Red dots show the value used in the synthetic simulation (i.e. the truth).Hospital-level simulations ofthe number of **C)** imported, **D)** nosocomial, and **E)** detected colonizations. Each column represents one scenario used for studying the identifiability of the system. Light and dark ribbons show the 95% and 50% uncertainty, respectively, constructed from 300 simulations of the posterior, and dots show the hospital-level observation of the synthetic simulation. Note, hospital level data were not used to optimize the model.

A picture containing diagram

Description automatically generated

**Figure 4. Posterior parameter estimates (A-B):** Violin plots show the posterior distribution and point estimates from the last iteration of the IF-EAKF (Methods section) for each of the microbial species (Data section). **C)** Hospital level fit: Light and dark ribbons show the 95% and 50% quantiles, respectively, constructed from 300 simulations with the posterior estimates of parameters. Observed carriage is plotted with dots at weekly time scale. D) Calibration plot. Hospital-level fit with 4 different confidence intervals (25%, 50, 75%, 95%). The black dotted line is the reference perfect calibration.

Calendar

Description automatically generated

**Figure 5. Importation and nosocomial contribution.** Weekly incident colonization of microbial species. Light and dark ribbons show the 95% and 50% quantiles, respectively, constructed from 300 simulations with the posterior estimates of parameters. Salmon/red color shows importation; light blue shows nosocomial transmission (left and right axis, respectively).

Diagram

Description automatically generated

**Figure 6. A)** Schematic of agent states, S: susceptible to colonization and C: colonized. **B)** Model diagram showing the colonization process for a single ward facility; is the force of Infection for ward i, . **C)** Schematic of a hospital with 2 wards of different size; arrows show the movement process within the hospital at the ward level: admission to the hospital network, transfer between wards, and discharge from the hospital network.

**Table 1.** Parameters used for synthetic scenarios and investigation of ABM identifiability.

|  |  |  |
| --- | --- | --- |
|  | **Importation probability** | **Nosocomial contact transmission rate** |
| Scenario 1 | 0.10 | 0.04 |
| Scenario 2 | 0.05 | 0.04 |
| Scenario 3 | 0.07 | 0.04 |
| Scenario 4 | 0.10 | 0.01 |
| Scenario 5 | 0.03 | 0.02 |
| Scenario 6 | 0.03 | 0.025 |
| Scenario 7 | 0.015 | 0.02 |
| Scenario 8 | 0.15 | 0.04 |
| Scenario 9 | 0.20 | 0.03 |
| Scenario 10 | 0.30 | 0.035 |

**Table 2.** **Posterior estimates**. Mean posterior estimates for the eight species studied; both

importation and nosocomial transmission rates, γ and β, are presented as rates per day.

|  |  |  |
| --- | --- | --- |
| **Organism** | **Importation probability** | **Nosocomial transmission rate** |
| *E. coli* | 0.64 | 0.0022 |
| *K. pneumoniae* | 0.26 | 0.00875 |
| *MSSA* | 0.145 | 0.020 |
| *MRSA* | 0.11 | 0.0148 |
| *C. albicans* | 0.11 | 0.0244 |
| *S. epidermis* | 0.10 | 0.011 |
| *P. aeruginosa* | 0.09 | 0.055 |
| *E.faecalis* | 0.09 | 0.0052 |