**Title 1:** Estimating nosocomial transmission of micro-organisms in hospital settings using patient records and culture data

**Authors:** Jaime Cascante Vega1, Rami Yaari1, Tal Robin1, Lingsheng Wen2, Jason Zucker2, Anne-Catrin Uhlemann2, Sen Pei1,\*, Jeffrey Shaman1,3,\*

**Background and Aims:**

Pathogenic bacteria are a major threat to patient health in hospitals, increasing the length of stay, contributing to mortality, and increasing overall hospital health-care-associated costs. Understanding the natural history and transmission dynamics of bacterial pathogens, including those resistant to antibiotics, are critical for designing better control measures. Similarly, measuring colonization rates in hospitals would help development of a quantitative framework that ultimately could support improved surveillance strategies and culture test allocation. Here we leverage electronic health records from a major New York City hospital system collected during 2020-2021 to support inference of nosocomial transmission and detection for eight micro-organisms.

**Methods and Results:**

We develop an agent-based model informed by patient hospitalization records, including admission, transfer and discharge, to simulate importation from the community, nosocomial transmission, and patient spontaneous decolonization of bacteria. The model is coupled with a Bayesian inference algorithm to estimate the likelihood of detection upon testing and nosocomial transmission rates. We evaluate parameter identifiability for this model-inference system and find that it is able to discriminate nosocomial transmission and effective sensitivity upon clinical culture testing. We apply the framework to estimate both quantities for eight prevalent bacterial pathogens: *Escherichia coli, Klebsiella pneumoniae, Pseudomonas aeruginosa, Staphylococcus aureus*(both sensitive, MSSA, and resistant, MRSA, phenotypes), *Staphylococcus epidermidis*, *Enterococcus faecium* and *Enterococcus faecalis*. We find that nosocomial transmission for *E. coli* is negligible and is lower for MSSA than MRSA. While bacterial pathogens are found to have different importation rates, nosocomial transmission rates were similar among organisms, except *E. coli*. We also find a similar estimated likelihood of detection  for all pathogens.

**Implications:**

This work highlights how fine-scale patient data can support inference of the epidemiological properties of micro-organisms and how hospital traffic, patient contact and surveillance determine epidemiological characteristics. Evaluation of the surveillance and transmission potential for different pathogens could ultimately support the development of hospital control measures, as well as the design of surveillance strategies.