Foundations of Data Science

*Capstone Project*

## Rob L. Wallace, roblwallace@gmail.com

# Abstract

In 1847 Ignaz Semmelweis proposed the practice of washing hands with chlorinated lime solutions while working in Vienna General Hospital obstetrical clinic. He published a book of his findings in 1861. Turns out the doctors’ clinic had a birth mother mortality rate three times higher than midwives’ clinics. His ideas were roundly rejected by the medical community of that time. Disappointing. However, Pasteur’s germ theory was still a decade or more away. One hundred and sixty-six years later we are still working on solutions and innovations to improve healthcare worker hand hygiene to reduce hospital-associated infections.

A hospital-associated infection is an infection that develops as a result of medical care. This may occur in a hospital, outpatient surgery center, nursing home, rehabilitation facility, or while receiving wound care services. In order to develop an infection while receiving these services, bacteria must enter your body. This can happen in many ways: through a wound, a device such as a catheter, or even by way of the lungs.[[1]](#footnote-1) The CDC estimates that one in 25 hospital patients has at least one healthcare-associated infection.

Hospital-associated infections (HAI) in the US affect more than 1.7 million patients annually. It’s estimated that in the US 99,000 deaths occur annually as a result of an HAI. The medical costs of healthcare-associated infections in US hospitals for inpatient hospital services are estimated to be between $28B and $45B annually. That’s a ‘B’ for billion.

Healthcare workers’ hands are the most common vehicle for the transmission of healthcare-associated pathogens from patient to patient and within the healthcare environment. Hand hygiene is the leading measure for preventing the spread of antimicrobial resistance and reducing healthcare-associated infections, but healthcare worker compliance with optimal practices remains low in most settings.

There are many technologies, services, and care practices that are emerging to address the human, and capital costs of HAIs. In particular, hand hygiene compliance system vendors need better tools to identify care delivery locations (hospitals, rehab facilities, etc.) that could avail themselves of their offerings. Further, once identified, the care delivery locations can be studied to better understand more about infection rates, drivers, and restraints. Identified sites can be interviewed to better understand what innovations are still needed to help drive adoption of technologies.

# Questions

Understanding that a lack of appropriate hand hygiene compliance contributes to infection rates in healthcare settings:

* Can we Identify these underperforming healthcare settings geographically?
* Are these underperforming healthcare settings concentrated geographically, by zip code, or county?
* Are there predictors of HAI performance by examining Medicare spending, per claim or per patient, and HAIs?
* Are there predictors of HAI performance using HCAHPS (quality scores)?
* Are there differences in hospital size (licensed beds) that drive better or worse HAI rates?

Can we, in short, identify and predict performance of a healthcare setting, predominately hospitals for this project, through a collection, preparation, and examination of the meta data available about US hospitals?

# The Data

There are a number publicly available data sources to extrapolate our story with. Two immediate sources of infection rates are Data.gov, and Healthdata.gov. Centers for Medicare and Medicaid Services (CMS) also hosts a number of public use data files. Initially identified data sources are:

* Hospital-Acquired Infections: Beginning 2008
* Healthcare Associated Infections – Hospital
* CMS Claims Data
* General Hospital Information
* Hospital Value Based Purchasing Healthcare Associated Infections
* Medicare Hospital Spending by Claim
* Medicare Hospital Spending by Patient
* Patient Survey HCAHPS
* Healthcare Associated Infections - State

I’ve a number of hardcopy books of hospital general information that may, or may not, lend themselves to assisting in the project. The strategy is to wrangle the data into a cohesive set of tables whereby we can answer the questions. Which in short shall be used to identify lower infection compliance market places to focus product selling and development efforts.

Of import to me as a product executive is to be able to graphically represent many of the results. Envisioned is overlaying result sets onto a US MAP and/or perhaps a county map. Fifteen years of experience in Fortune 15 companies tells me we are going to need cogent pictures to demonstrate findings.

Hospitals come in many sizes, and have different demographics served. How do we measure an infection rate for a 90 bed hospital against a 900 bed hospital? A measure that might be useful to even the analysis “patient days.” A patient day (P.D.) is defined as “a unit in a system of accounted used by healthcare facilities and healthcare planners. Each day represents a unit of time during which the services of the institution or facility are used by a patient; thus 50 patients in a hospital for 1 day would represent 50 patient days.”[[2]](#footnote-2) By using patient days we can compare infection rates of a Mount Sinai Medical Center (1,171 beds) to a Trident Medical Center (90 beds).

# Deliverables

At this stage of project development, a number of deliverables have been identified. The deliverables include;

* R Code
* PowerPoint Presentation
* Recommendation Document

Of course, all materials will be uploaded to GitHub repository and clearly identified.

# In Summary

I wrote, and was granted, a patent in 2015 for a device that helps identify whether or not a healthcare worker sanitized their hands. This is an emerging technology using indoor GPS like devices to detect entry/exit, and sanitizer activation/proximity. My GE Healthcare business sold some of the first systems installed in US hospitals in 2012. We were finding the early adopters, but they are too few to support growth & contribution margins required by Finance. We needed a better way to identify markets & customers faster.

The healthcare system needs our help as inventors and innovators. Self-reporting of hand hygiene compliance by healthcare workers is generally over estimated. Direct observation through “secret shoppers” is also generally over-estimated. One of my customers noted that when self-reporting an 80% hand hygiene compliance rate, his hospital used 1/3 of the soap and sanitizer for a 60% compliance rating my systems were reporting. It just didn’t add up. And one in 25 of us (as patients) with an HAI isn’t ok. We shouldn’t go to a hospital and get sicker because of something largely preventable with better hand hygiene.

We can find a better way to identify & predict markets and customers. And do what your mother told you to do; “wash your hands.”

1. Association for Professional in Infection Control and Epidemiology (APIC) | What is an HAI? [↑](#footnote-ref-1)
2. patient day. (n.d.) *Mosby's Medical Dictionary, 8th edition*. (2009). Retrieved September 8 2016 from <http://medical-dictionary.thefreedictionary.com/patient+day> [↑](#footnote-ref-2)