Inferring sources of dangerous products in pharmaceutical supply chains

1 Supply chain inference

Post-market surveillance in itself can provide valuable information with respect to substandard and falsified product (SFP) prevalence and patterns. However, the effectiveness of scarce sampling resources can be greatly improved through knowledge of supply-chain structures and inventory policies since SFPs "...tend to cluster in areas or situations where risk factors that favor the manufacture or sale of such products are high..."(World Health Organization, 2017). Pisani et al. (2019) and Nkansah et al. (2017) emphasize the link between SFPs and supply chain factors. The simulation of relevant supply chains heightens the ability to discern robust policies in the face of different unknown scenarios, while expanded statistical techniques permit refined estimates and analyses of SFP prevalence.

The project will progress under the following steps: (i) develop a modular framework that includes testing and supply chain behavior; (ii) include supply chain behavior such as ordering policies, stocking and distribution; (iii) survey relevant parameters that determine the size and structure of the supply chain of interest; (iv) construct a suite of implementation scenarios that can be used to investigate the effectiveness of end-node testing.

1.1 Constructing a flexible simulation environment for analysis

A realistic simulator of pharmaceutical supply chains permit the evaluation of sampling polices' ability to detect illicit products given particular supply chain structure and agent behavior. We will develop a systematic, flexible approach that integrates available data into a simulation model that best represents the illicit supply chains under study. For example, when faced with a significant shortage, a pharmacy may expand its supplier options to less reliable sources, which can in turn increase the risk of SFPs entering the market.

An initial simulation model has been constructed in `python` as part of the NSF EAGER project. Our initial simulation model utilizes a graph-based structure to represent a pharmaceutical supply chain similar to that observed in our field study in Liberia. Nodes in each echelon, such as manufacturers, importers and pharmacies, are connected by arcs that represent procurement and distribution preferences. This construction allows us to flexibly structure the echelons beyond a simple analytic model and to explicitly describe supply chain decision making between echelons. Based on data from our collaborators, we simulate testing at the pharmacy (sales outlets) echelon and potential results. With this simulation environment, we can also investigate testing throughout the supply chain and evaluate the impact of changes in structure, agent behavior, etc. This experimentation will examine sampling paradigms in complex environments, leveraging limited field testing to gain the maximal amount of data.

Model development will be done in an open-source framework to improve widespread adoption and best practices. The simulation will use file-based i/o structures to enable deployment on computer clusters for large-scale experimentation. This experimentation will examine sampling paradigms in complex environments, leveraging limited field testing to gain the maximal amount of data. Parameters and supply chain information will be documented and tracked to replicate the scenarios faced by our collaborators.

As an example of the potential inferences from our simulation model, consider a setting with three intermediate nodes supplying end-nodes: two "good" intermediate nodes with high quality products and one "bad" intermediate node through which SFPs can enter. This example shows how supply chain conditions (here, stock-out rates due to supply levels) impact detection abilities

with end-node PAD testing. Figure 1 presents the estimated SFP rates at each intermediate node from simulated PAD testing data at the end-nodes as generated by the simulator. Our sampling policy is effective if we can connect SFPs to the correct source (here, the third intermediate node). However, we find that this underlying system is only detectable in the low stock-out scenario. These simulation results show how limited supplies (stock-outs in pharmaceutical markets) impacts not only the presence of illicit products as shown in Pisani et al. (2019), but also the ability to detect the source of these products due to the ways in which agents adapt to stock-outs.

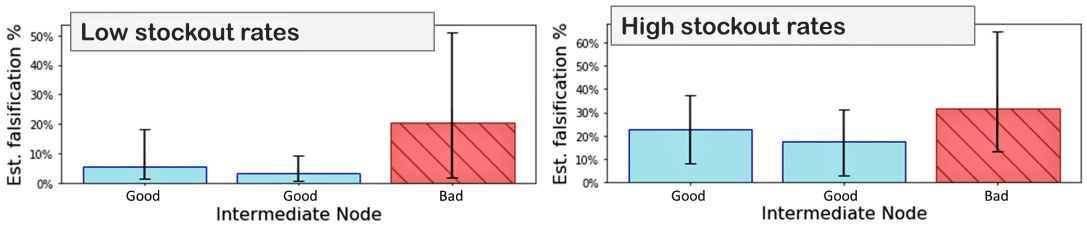

Figure 1: *Estimates and uncertainty bounds of falsification rate under different stock-out rates using an initial simulation model.*

1.2 Statistical modeling, inference and adaptive sampling

We aim to create a principled statistical inference framework that encompasses findings about the supply chain network structure and diagnostic properties. This inference builds on a novel analytic statistical model that explicitly encodes observational randomness in diagnostics (sensitivity and specificity) but also allows for statistical estimation via model inversion beyond the capabilities of the simulation model described. This analytic model quantifies information about the supply chain structure (size, ordering policies) through information obtained over a partially observable network.

Our analytic model is built on hierarchical components where end-node sampling (post-market surveillance) is the primary source of testing. The upper level corresponds to m importers (where no domestic production occurs, as in our Liberia study) and the lower level corresponds to n outlets (pharmacies). For the settings we are considering, it is unlikely that the procurement relationships between outlets and importers will be fully known, making the network partially observable. Each importer or outlet represents an potential entry point for SFPs. Let π_i be the fraction of product that outlet i acquires from an SFP supplier and θ_j be the fraction of product that importer j acquires from an SFP supplier. Our goal is to uncover the SFP entry points using testing at outlets (i.e., estimates of π and θ). This testing yields readings, labeled $p_1, \ldots, p_n$, that are noisy due to imperfect sensitivity and specificity.

Our analytic model will allow researchers and practitioners to move beyond the trivial assessment that the presence of positive tests at an end-node is the direct fault of the end-node. A representation of a supply chain structure is $p = A\theta + \pi + \varepsilon$, which is a matrix-vector expression of the end-node testing observations. The $n \times m$ transition matrix A has rows for each end-node where the ijth value represents the probability of end-node i receiving product from source-node j. The ε corresponds to the error caused by limited PAD testing and sampling levels.

The core novelty of this research is that out-of-the-box statistical calculations directly depend on the supply network matrix A, which we will glean from results of the simulation model. In environments with low information access, it is feasible to estimate with some uncertainty the transition matrix A through surveys at the end-nodes. The proposed Bayesian scheme will be

flexible enough to carry through this uncertainty in A. Additionally, one can derive analytical results from this hierarchical structure. For example, in a completely vertical supply chain, where all end-nodes procure from a single importer, the preliminary findings are that $\hat{\theta}_j$ is the average of all end-nodes that receive product from the jth importer, and $\hat{\theta}_i$ is the difference between that end-node's performance and all other end-nodes that supply from that importer. The left panel of Figure 1 illustrates this situation, with positive performance results. Alternatively, the right panel of Figure 1 shows that these results can change with more diverse distribution networks.

We will expand this modelling framework to consider a broader array of settings, including a time-dependent setting with temporal properties of A, θ and/or π.

References

Nkansah, P., Smine, K., Pribluda, V., Phanouvong, S., Dunn, C., Walfish, S., Umaru, F., Clark, A., Kaddu, G., Hajjou, M., Nwokike, J., and Evans, L. (2017), "Guidance for Implementing Risk-Based Post-Marketing Quality Surveillance in Low- and Middle-Income Countries," Tech. rep., U.S. Pharmacopeial Convention. The Promoting the Quality of Medicines Program. Rockville, Maryland.

Pisani, E., Nistor, A.-L., Hasnida, A., Parmaksiz, K., Xu, J., and Kok, M. O. (2019), "Identifying market risk for substandard and falsified medicines: an analytic framework based on qualitative research in China, Indonesia, Turkey and Romania," *Wellcome Open Research*, 4, 70.

World Health Organization (2017), "A study on the public health and socioeconomic impact of substandard and falsified medical products," Tech. rep., World Health Organization, licence: CC BY-NC-SA 3.0 IGO.