

A Multi-level Modeling Approach to Prioritizing Interventions to Reduce Food Risk

J. Glenn Morris^{a*}, Michael B. Batz^b, Michael R. Taylor^b

^a Dept. of Epidemiology and Preventive Medicine, Univ. of Maryland School of Medicine, Baltimore, MD

^b Resources for the Future, Washington, DC

* Correspondence to: Jmorris@epi.umaryland.edu

Motivation¹

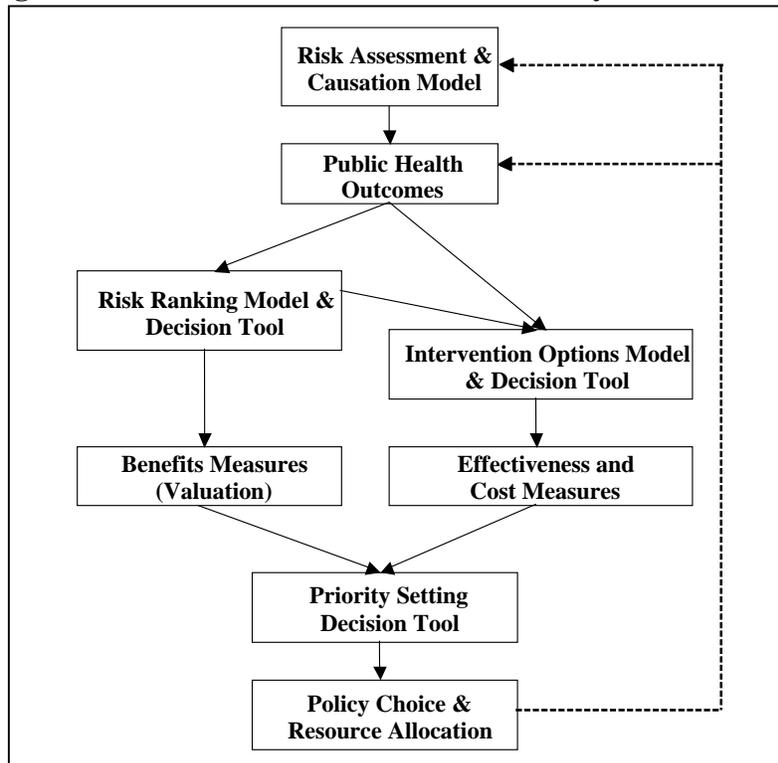
In a science- and risk-based food safety system, decision makers must utilize the best available data to determine the greatest risks to public health and to prioritize interventions to reduce these risks. This requires understanding the many factors that can contribute to the causation and prevention of foodborne illness from the point of production to the point of consumption and being able to systematically target efforts to reduce risk. Unfortunately, the data and analytic tools necessary for such a holistic systems approach to food safety decision-making do not exist.

Decision tools are necessary to help food safety decision makers perform two key tasks: to identify the most significant risks from a public health perspective, and to prioritize opportunities to reduce these risks, taking into account the benefits, effectiveness, and cost of possible interventions. Figure 1, while simplistic, highlights interactions among these various elements.

The initial FSRC modeling project focused on development of a Foodborne Illness Risk Ranking Model (FIRRM). To move further into this diagram, we propose a flexible three-component (or three tiered) modeling framework within which data and models can be designed, built, and ultimately integrated.

¹ This draft was written to encourage discourse at the workshop “Approaches to Predictive Modeling: Building a Framework for Prioritizing Opportunities to Reduce Food Safety Risk,” held at Iowa State University in Ames, Iowa, on June 15 and 16, 2004. It is likely that significant changes will be incorporated into our proposed approach following workshop presentations and discussions.

Figure 1: General Framework of Food Safety Decision Tools



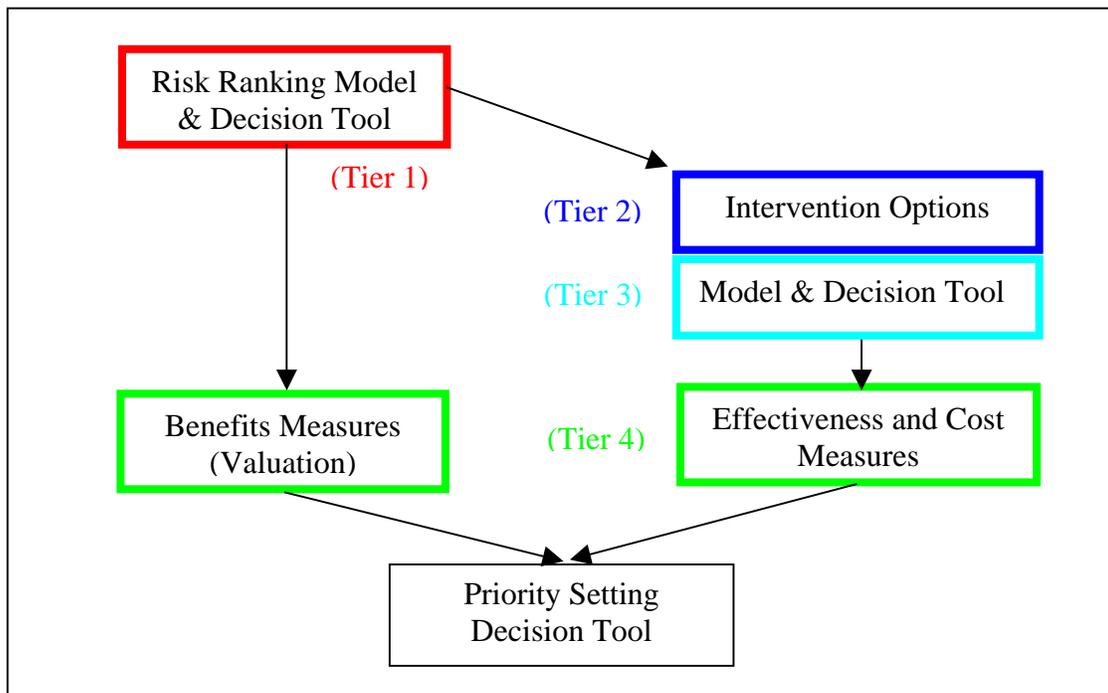
Approach

How do we prioritize opportunities to reduce risk? We need to do so globally; that is, we need to start with a big picture view based on wide, aggregated data and then narrow the target further and further, moving towards greater precision and accuracy as well as greater complexity. No single model will be sufficient to be able to capture the breadth and depth of the food safety system. Thus, we propose a three-component approach that starts with broad, national, public health surveillance data and moves, through each component, toward the selective microbiological modeling of the farm to fork continuum.

The first component of this framework includes the identification and quantification of the food/pathogen pathways responsible for illnesses. This builds on concepts in FIRRM, within which we estimate illnesses due to specific pathogens based on public health surveillance data and then attribute these illnesses to food categories using additional data. In a second level (Tier 2), we propose to model aggregate risk by food commodity to further narrow down possible control points. Specifically, we propose the statistical analysis of temporal contamination data at the farm and retail levels, quantifying the extent to which changes in farm contamination are

responsible for variations in commodity contamination. In a third level (Tier 3), we propose to model the amplification and reduction of pathogen contamination risk along the food supply chain to identify the most important control points and to model the effectiveness of possible interventions. Such a model would provide the necessary basis for cost-effectiveness studies (Tier 4), leading ultimately to a priority-setting decision tool. These elements can be conceptualized as part of the previous diagram, as shown in Figure 2.

Figure 2: Four-Tier Modeling Framework of Food Safety Decision Tools



Overview of Modeling Approaches

Implicit in the three-tiered framework is the belief that different types of data and analytical questions require different modeling approaches to utilize the data most effectively and to provide the most reliable answers. There are a myriad of different types of mathematical models, and while we need not discuss them comprehensively, it may be of service to discuss a few broad categories.

Epidemiological models rely primarily on data on human illnesses, such as public health surveillance data, to identify causes of disease. Relevant sources of data includes aggregate records of outbreaks, active surveillance of cases (eg. FoodNet), case control studies, and

microbial source tracking of clinical isolates (PulseNet), among others. The epidemiological approach is essentially “top-down” and retrospective; it generally relies on aggregated outcome data to identify root causes.

Predictive microbiological modeling, on the other hand, is essentially “bottom-up” and forward-looking. It relies on detailed microbiological data such as pathogen prevalence studies on foods, along with knowledge of pathogen growth under various environmental settings, to estimate pathogen amplification or reductions. Risk assessment modeling takes this one step further and predicts illnesses resulting from these changes to pathogen contamination.

Such predictive models are generally considered “white-box,” meaning that there is a great amount of a priori knowledge about system behavior. That is, the functions relating different variables are known and can be parameterized according to the data. In reality, few models are entirely white-box, due to limitations of scientific knowledge and prohibitive model complexity.

Alternatively, what might be called “toy models” are generally “black-box” models because there is little a priori knowledge about system behavior. In such cases, the functions relating variables must be assumed or estimated, and parameter values must also be assumed or estimated. Such models may be based on statistical regression modeling, differential equations, pattern matching, or alternative methods, and generally imply model specifications drawn from analysis of the data itself.

Most models are not entirely white-box or black-box, but rather, exist somewhere in the continuum between these extremes, based on the amount of available a priori knowledge and modeling resources. Models of complex systems, such as food risk models, likely include both black-box and white-box components.

The Starting Point (Tier 1): Epidemiological Modeling

The first step in prioritizing interventions to reduce risk of foodborne illness is to identify the most critical pathways responsible for illnesses. We believe the best approach is to begin with knowledge of actual illnesses, and work backwards using epidemiological methods, to identify and quantify causes. In Tier 1 models, foods are grouped into fairly broad, yet still meaningful, categories. It may be sufficient for Tier 1 models to differentiate between beef and

poultry, but not necessarily between ground beef and beef cuts, as these commodity-level distinctions will come into play in Tier 2. Ultimately, Tier 1 models should estimate the burden of specific food/pathogen pathways to be used in Tier 2.

FSRC's FIRRM, currently in development, is a starting point in satisfying Tier 1 needs. This model, as discussed in detail in previous FSRC documents, estimates national incidence of foodborne illnesses, in terms of cases, hospitalizations, and deaths, due to 28 pathogens, based on the data and methodology within Mead et al. (1999). The model then estimates the economic costs of these illnesses, as well as the quantified loss in quality of life, measured in Quality Adjusted Life Years (QALYs). Illnesses due to each pathogen are then attributed to food categories based on data from outbreaks or the results of an expert elicitation. The key model outputs are rankings of food/pathogen pathways according to five measures of annual public health burden: cases, hospitalizations, deaths, dollars, and QALYs. It is important to consider disaggregated pathogens according to multiple measures of burden because of the disparity of incidence and severity of foodborne illnesses; Norovirus is exceedingly common but mild, while *Listeria* is fairly uncommon but very severe.

As has become obvious in development of FIRRM, there continue to be substantive data gaps which severely limit our ability to appropriately rank food pathogen combinations. A critical gap relates to our ability to identify food sources for illness (i.e., food attribution). Without such data, rankings become problematic, impacting our ability to prioritize commodity groups for study – and our ultimate ability to assess cost effectiveness and benefits.

Tier 2: Commodity-Level Modeling

Due to the complexity and scope of the food safety system, it is impossible to model the entirety of it as a “white-box” system. From the farm to fork, there are far too many stages in the food supply chain to model for far too many types of food. There is tremendous variation in each stage, and a priori knowledge is not known about interactions between variables in each stage or between stages.

As such, we propose Tier 2 commodity-level models, which will rely on pathogen contamination data gathered at various stages in the food system. For the food/pathogen pathways identified by Tier 1 models, Tier 2 models will further narrow foods of interest into

commodities to target with interventions. Such models will be quite black-box and will not rely on internal assumptions about pathogen growth or decline in individual stages. Rather, contamination data collected at the farm and retail stages, as well as major stages between, will be analyzed for the mathematical relationships between inputs and outputs. Different mathematical and statistical approaches might be used in the analysis of data to look for patterns. Such analysis should indicate which commodities are of most interest and, given rich sampling, which stages are most likely to warrant interventions.

The resulting “toy models” will indicate, in a broad sense, the impact of interventions at various stages in the food supply chain. For example, a spike in contamination in beef on the farm may be correlated to a spike in contamination of ground beef, while no corresponding spike in beef cuts may be witnessed. Such a pattern would result in a model that pointed towards interventions specific to ground beef processing or further upstream. Such black-box models would be insufficient, however, to predict the effect of specific interventions.

Tier 3: Intervention Modeling

After Tier 1 models have identified food/pathogen pathways and Tier 2 models have narrowed down to commodities, Tier 3 models will focus on interventions at process-specific control points. In Tier 3 models, pathogen contamination will be modeled through the food supply chain of a particular commodity. Depending on the commodity and interventions being investigated, Tier 3 models may go upstream as far back as the farm and forward as far as consumption. Each stage in this chain, including including slaughter or harvest, processing, manufacturing, transportation, storage, retail handling will be modeled individually.

These complex Tier 3 models will involve both black-box and white-box modeling of individual stages. Where a priori knowledge is available, microbiological systems will be modeled to the fullest extent possible; parameters such as temperature, light, pressure, acidity, surface area, and so on will be used to predict pathogen growth or decline. Where this is not possible, simple curves will be created to translate a change in contamination level prior to a particular stage to a change in contamination after that stage. Interventions will be modeled as changes to the functional forms of white-box systems or to the curves of black-box systems.

Using Tier 3 models, we will be able to compare interventions at different stages in processing in terms of effectiveness in lowering contamination at retail or consumption. If cost information is available for interventions, we will be able to compare interventions in terms of cost-effectiveness. Ultimately, Tier 3 models should translate changes in contamination to changes in the number of foodborne illnesses using risk assessment methods. FIRRM or other Tier 1 models could be used to estimate the economic benefits of reductions in illnesses, and traditional cost-benefit analyses (Tier 4) could be performed on interventions across pathogens and foods.

Concerns and Conclusion

Though we feel that such a multi-level approach to prioritizing interventions to reduce the risk of foodborne illness provides a useful framework for future modeling, many questions and concerns remain.

First are the theoretical concerns. Is such a multi-tiered approach legitimate? Or is a single modeling framework a better approach? Do our suggested models have precedent and can they be verified? We are fairly confident that the proposed framework satisfies these concerns. Our framework builds off existing modeling approaches and data and merely combines information in a more comprehensive manner. Furthermore, we believe that if done properly, the various models can be used together to inform and validate each other. Having multiple models instead of one increases the likelihood that we may encounter inconsistencies, but resolving these inconsistencies likely involves improving our understanding of the system as a whole.

The practical concerns are another matter, however. The primary practical concerns are twofold: first, does the data exist and if not, can we collect it, and second, how will the models interact?

The data issue is paramount. In regards to Tier 1 models, the food attribution component has been previously highlighted as a particularly thorny issue without simple resolution. Can food attribution be improved given the current state of affairs such that results are reliable? In Tier 2 models, the data problems are even more severe. Where will we get data on pathogen contamination in the food supply chain? Does this data exist and if so does it satisfy temporal and sampling concerns? If it does not exist, can we collect it, and how? Tier 3 models have a rich

precedent in risk assessment and other microbiological models, but we do not know about precedence in food supply chain modeling specifically. Are there available models and data to build upon? Do we have the data or knowledge about processes to model individual stages in the food supply chain? If we do not have the information to produce white-box models of system components, will black-box models be sufficient?

Another significant concern is figuring out how the tiered models will interact. The transition from Tier 1 to Tier 2 is straightforward and uncontroversial, but the transition between Tier 2 and Tier 3 is not. These models are structurally similar in that they both model the food supply chain in different levels of complexity and detail. To what extent will the Tier 2 model inform the Tier 3 model? Will the Tier 2 model only provide broad suggestions as to where to focus attention, or will it somehow interact with the Tier 3 model on a more substantial level, such as giving contamination reduction targets? Will or should both models rely on the same data on food contamination? Lastly, in this regard, how might Tier 3 model results be folded back into the Tier 1 model to see how changes affect food/pathogen rankings?

There are other concerns, of course, but really it comes down to one question: are we on the right track?