An Economic Evaluation of the Whole Genome Sequencing Program in the U.S.

Brad Brown, Travis Minor, Marc Allard, Michael C. Bazaco, & Joseph Blankenship
United States Food and Drug Administration (FDA), Center for Food Safety and Applied Nutrition, College Park, Maryland, 20740

Abstract

The FDA created the GenomeTrakr Whole Genome Sequencing (WGS) Network in 2013, as a tool to improve food safety. The study presents the impact of WGS implementation on potential food contamination and related illnesses using a social welfare model. We conduct empirical tests using data from FDA regulated food commodity outbreaks garnering FDA response from 1999 through 2019 and examine the effect of WGS program isolates collected in the U.S. on outbreak illnesses. We connect the empirical results with existing literature and conduct a Monte Carlo analysis to estimate benefits and costs. Empirical results are consistent with the theoretical model and suggest that each additional 1,000 WGS isolates added to the public National Institute of Health's (NIH) National Center for Biotechnology Information (NCBI) database is associated with a reduction of approximately 6 illnesses per WGS pathogen, per year. Further, the analysis suggests that this effect holds even though we identify slightly more outbreaks for WGS pathogens each year relative to other pathogens, because these outbreaks tend to be smaller in magnitude. By 2019, we estimate annual health benefits of nearly $500 million, compared to a roughly $22 million investment by public health agencies. Even under conservative assumptions, the program likely paid for itself in its second year of implementation. We expect increasing public health benefits as the WGS GenomeTrakr network matures.

Introduction

Despite significant effort to improve and modernize the food safety system in the United States, foodborne pathogens remain a major public health threat, causing 9.4 million illnesses each year, including 56,000 hospitalizations, and 1,400 deaths. Illnesses caused by foods regulated by the FDA account for roughly 80% of those illnesses. The FDA's Center for Food Safety and Applied Nutrition (CFSAN) created the GenomeTrakr Whole Genome Sequencing (WGS) Network in 2013, as a tool to help improve food safety. WGS provides more precise, high resolution source tracking and predictions for food and environmental genomic data. Results from the analysis of the data enables faster and more precise public health and regulatory actions (such as public messaging and recalls), thus decreasing the number of illnesses associated with outbreaks and decreasing the breadth of products recalled on average. The program makes it easier to evaluate cases across time and geography, helping to solve ongoing contamination events that would otherwise go unidentified. Through empirical and cost benefit analyses, we evaluate the effects of the WGS program on the FDA's ability to detect, investigate, and limit the spread of outbreaks linked to FDA-regulated commodities, and the costs of the program. Even though the program is relatively new, and the theoretical effects of WGS on observed illnesses are varied, measuring the effect of such a program on foodborne illness already shows clear improvements over time.

Theoretical modeling shows the net effect of WGS on social welfare is ambiguous. We cannot know without further analysis, weighing the benefits of illness reduction against the costs of implementation, whether society is better off with WGS. Especially in the early stages of implementation of any program, it is possible the costs could outweigh the benefits. It takes time for the food industry to adopt and invest in contamination controls and therefore, in the short term, it is possible the detection of illnesses will drive the results of WGS implementation on the number of observed illnesses.

Materials and Methods

The data for this analysis is primarily drawn from the FDA's Coordinated Outbreak Response and Evaluation (CORE) database. The CORE database includes detailed information on foodborne outbreaks that were investigated by the FDA. While this data does not represent all outbreaks related to (or likely associated with) FDA-regulated human foods, it does represent the scope of outbreaks with direct FDA involvement. The database includes information on the number of confirmed illnesses, the associated pathogen and food vehicle, and the timing of FDA's investigations for each outbreak. From this data, we construct a pathogen and year panel comprised of the annual number of illnesses and outbreaks for each observed pathogen between 1999-2019.

Additional data on WGS isolates is drawn from the National Institute of Health's (NIH) National Center for Biotechnology Information (NCBI). NCBI collects sequencing data submitted by public health officials, academic researchers, or industry sources as a central repository designed to facilitate analysis and aid in outbreak and traceback investigations. We use the annual number of unique isolates compiled by NCBI as an indicator of the preventive power of this program.

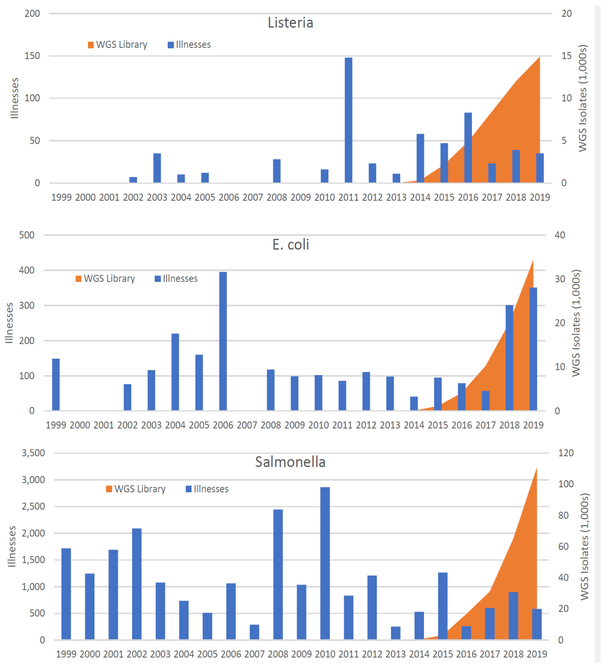

Figure 1. ILLNESSES & WGS ISOLATES BY PATHOGEN

Results and Discussion

Results from Table 1 suggest that the WGS program is associated with fewer illnesses for the collected pathogens over time, and this effect is largely observed because of smaller but slightly more frequent detection of outbreaks for those pathogens. Moving forward with the benefit-cost analysis, we focus on the estimated reduction in illnesses from Model 3.

Estimates for reduction in illnesses due to WGS by 2019 range from 210 illnesses annually, or a 13% reduction for *Listeria*, the most heavily sequenced pathogen at this point, relative to the number of associated illnesses occurring each year, to roughly 19,800 illnesses, or about a 1.5% reduction for *Salmonella*, that has been sequenced relatively less and is more recently ramping up (see Table 2 & Figure 2). The total burden of illness reduction in 2019 is nearly $500 million, or a little over 1.5% of total burden of illness attributed to FDA-regulated foods (see Table 2 & Figure 3).

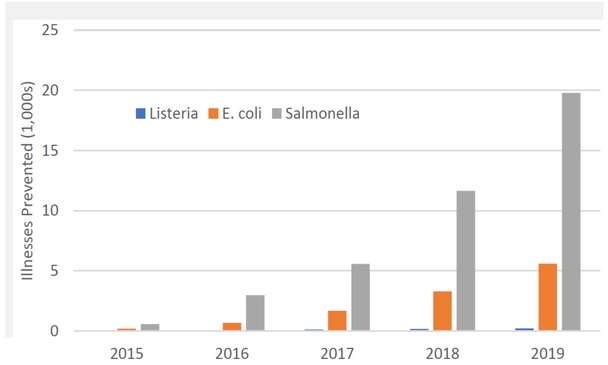

Figure 2. ESTIMATED ILLNESSES PREVENTED PER PATHOGEN

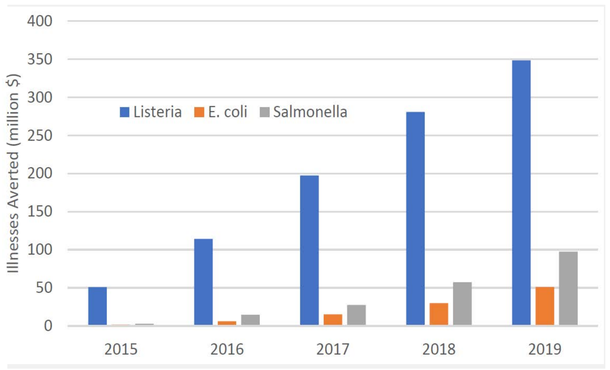

Figure 3. ESTIMATED BURDEN OF ILLNESSES PREVENTED PER PATHOGEN

Table 1. ESTIMATED EFFECT OF WGS LIBRARY ON ILLNESSES AND OUTBREAKS

	Illnesses			Outbreaks			Average Ill		
	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3
	Direct Effects Estimation								
WGS Lib.	7.36***	-6.11**	-6.09**	0.14***	0.01*	0.01*	0.58**	-1.06***	-1.07***
	(0.67)	(2.25)	(2.25)	(0.01)	(0.01)	(0.01)	(0.21)	(0.32)	(0.31)
R^2	0.03	0.71	0.71	0.07	0.99	0.99	0.00	0.26	0.26
FE Controls		X	X		X	X		X	X
FSMA			X			X			X

Notes: N. Obs. = 462. Significance levels are indicated as: *** significant at beyond the 1 percent level; ** significant at the 5 percent level; * significant at the 10 percent level; + significant at the 15 percent level.

Table 2. ESTIMATED BURDEN OF ILLNESS AVERTED

	Observed Effects Only					With Underreporting and Underdiagnosis Multipliers				
	List.	E. coli	Sal.	Yearly Total	95% CI	List.	E. coli	Sal.	Yearly Total	95% CI
	Estimated Illnesses Averted									
2014	2	0	3	5	(2 - 8)	4	13	80	98	(37 - 166)
2015	13	7	20	40	(16 - 64)	31	185	574	789	(297 - 1,339)
2016	30	25	102	157	(62 - 252)	69	671	2,982	3,722	(1,398 - 6,339)
2017	51	63	190	304	(119 - 489)	119	1,670	5,577	7,366	(2,770 - 12,534)
2018	73	123	397	593	(233 - 954)	169	3,281	11,636	15,085	(5,670 - 25,683)
2019	91	210	675	976	(383 - 1,569)	210	5,592	19,792	25,595	(9,619 - 43,589)
	Monetized Illnesses Averted in Millions of $									
2014	$3.22	$0.00	$0.01	$3.24	($1.22 - $5.51)	$7.43	$0.12	$0.39	$7.94	($2.96 - $13.61)
2015	$22.07	$0.06	$0.10	$22.23	($8.36 - $37.87)	$50.95	$1.68	$2.83	$55.46	($20.79 - $94.89)
2016	$49.48	$0.23	$0.50	$50.21	($18.89 - $85.49)	$114.23	$6.13	$14.69	$135.04	($51.03 -$229.39)
2017	$85.51	$0.57	$0.94	$87.01	($37.72 - $148.09)	$197.39	$15.24	$27.46	$240.09	($90.87 - $406.78)
2018	$121.56	$1.12	$1.96	$124.64	($46.92 - $211.99)	$280.62	$29.94	$57.30	$367.86	($139.56 - $620.41)
2019	$150.96	$1.91	$3.33	$156.19	($58.83 - $265.47)	$348.48	$51.03	$97.47	$496.98	($188.62 -$835.92)

Notes: Underreporting/underdiagnosis multiplier as well as illness burden by pathogen reported in Table 3.[1, 21, 22, 23]

Conclusion

Using data collected on outbreaks associated with FDA-regulated foods, we estimate the effect additional WGS isolates have on the burden of foodborne illness for *E. coli*, *Listeria*, and *Salmonella*. Results suggest that WGS has been successful. Illness numbers of heavily sequenced pathogens are falling faster relative to non-sequenced pathogens and observed WGS pathogen outbreaks are getting smaller. Under current funding and growth levels, the net benefits are somewhere between $100 million and $450 million. These estimated benefits of the WGS program easily outweigh the estimated costs of implementation after the second year. Once the program is fully implemented, we may see net benefits measure in billions of dollars.

Other countries and different geographic regions have different baseline regulatory climates and food safety cultures, so the marginal effect of WGS may vary from what we observe in US data, but this study provides strong evidence for a significant improvement in food safety anywhere WGS is implemented. The benefits of the program are applicable to COVID-19 and other infectious disease control applications (hospitals, nursing homes, medical manufacturing, waste management, composting, agricultural water use and reuse). In each of these instances, results from this analysis demonstrate that incorporating WGS may provide real positive public health benefits even in the early stages of implementation.