Food Source Prediction of Shiga Toxin-Producing *Escherichia coli* (STEC) Outbreaks using Demographic and Outbreak Characteristics, United States, 1998-2014

Alice White

InFORM, November 20, 2015

Introduction

- Outbreak investigation identify food safety interventions
  - Pathogen, food, contributing factors
- Identifying food vehicle is critical
  - 42% of outbreaks do not confirm food vehicle
- Data from prior outbreaks can help identify the food vehicle

Purpose

- Use existing outbreak data to develop methods and tools to more efficiently determine food vehicles
- Used Shiga toxin-producing *Escherichia coli* (STEC) outbreaks to describe differences in demographic and outbreak characteristics and predict food vehicles using differential characteristics

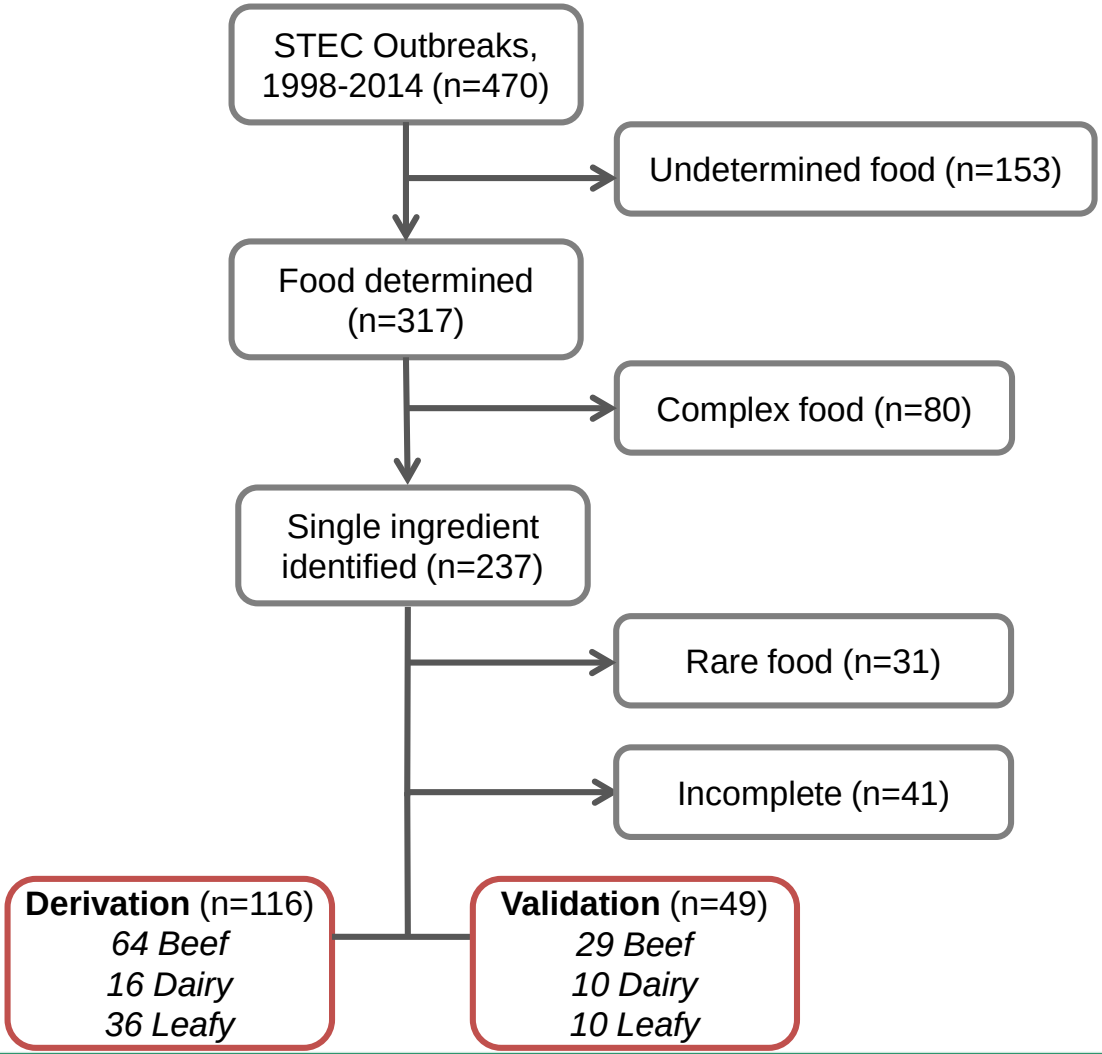
Methods

- Data source
  - Electronic Foodborne Outbreak Reporting System (eFORS), 1998-2008
  - National Outbreak Reporting System (NORS), 2009-2014
- Food source categories
  - Interagency Food Safety Analytics Collaboration (IFSAC) Food Categorization Scheme
- Predictors
  - Percent female, percent hospitalized, percent age categories (< 5 years, 5-19, 20-49, >50), multistate exposure, setting (private or non-private), season, duration, serotype

Methods

- Statistical analysis
  - Compare demographic and outbreak factors across food sources univariately
  - Build prediction model (70% derivation set)
  - Validate model (30% validation set)

Results


STEC Outbreaks, 1998-2014 (n=470)
Undetermined food (n=153)
Food determined (n=317)
Complex food (n=80)
Single ingredient identified (n=237)
Rare food (n=31)
Incomplete (n=41)
Derivation (n=116)
64 Beef
16 Dairy
36 Leafy
Validation (n=49)
29 Beef
10 Dairy
10 Leafy

Univariate Results

Predictor	Beef	Dairy	Leafy	P value
Gender *				
% Female	50 (34-70)	50 (33-67)	64 (56-74)	0.009
Age (years) *				
% <5	0 (0-13)	22 (0-33)	0 (0-4)	0.014
% 5-19	35 (12-61)	50 (50-67)	20 (15-33)	0.017
% 20-49	21 (0-33)	17 (0-33)	44 (30-54)	<0.001
% >50	14 (0-33)	0 (0-0)	19 (9-40)	0.002
Cases *	9 (3-17)	5 (3-9)	18 (10-33)	<0.001

* Data presented as median (IQR), p value from Kruskal-Wallis

Univariate Results

Predictor	Beef	Dairy	Leafy	P value
Setting†				<0.001
Private Home	45 (52%)	14 (70%)	5 (13%)	
Non-private	33 (38%)	4 (20%)	32 (82%)	
Multistate Exposure†	31 (36%)	2 (10%)	22 (56%)	0.002
Season‡				0.108
Fall	17 (20%)	7 (35%)	17 (44%)	
Spring	28 (33%)	6 (30%)	8 (21%)	
Summer	34 (40%)	6 (30%)	9 (23%)	
Winter	7 (8%)	1 (5%)	5 (13%)	
Non-O157:H7 Serotype†	3 (4%)	5 (25%)	7 (18%)	0.003

† Presented as proportion (%), p value from Chi-square
‡ Presented as proportion (%), Fisher's Exact

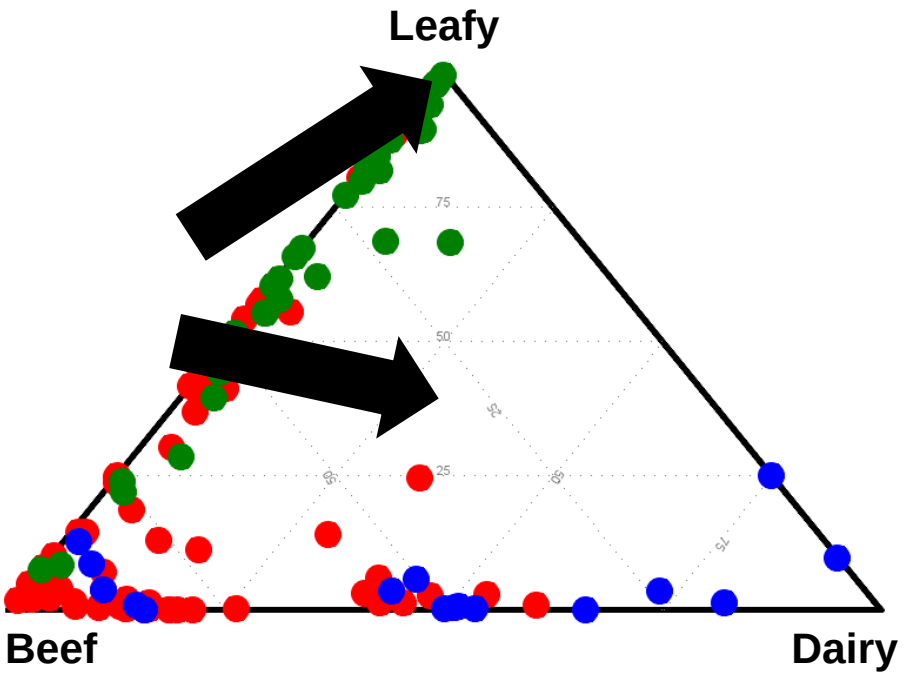
Prediction Model

PREDICTED OUTCOME	OBSERVED OUTCOME: Beef	Dairy	Leafy	Total
Beef	**56**	7	8	71
Dairy	2	**9**	0	11
Leafy	6	0	**28**	34
Total	64	16	36	116
Sensitivity	*0.88*	*0.56*	*0.78*	
Specificity	*0.71*	*0.98*	*0.93*	

***Final model: Percent female, cases, multistate exposure, setting, season, serotype**

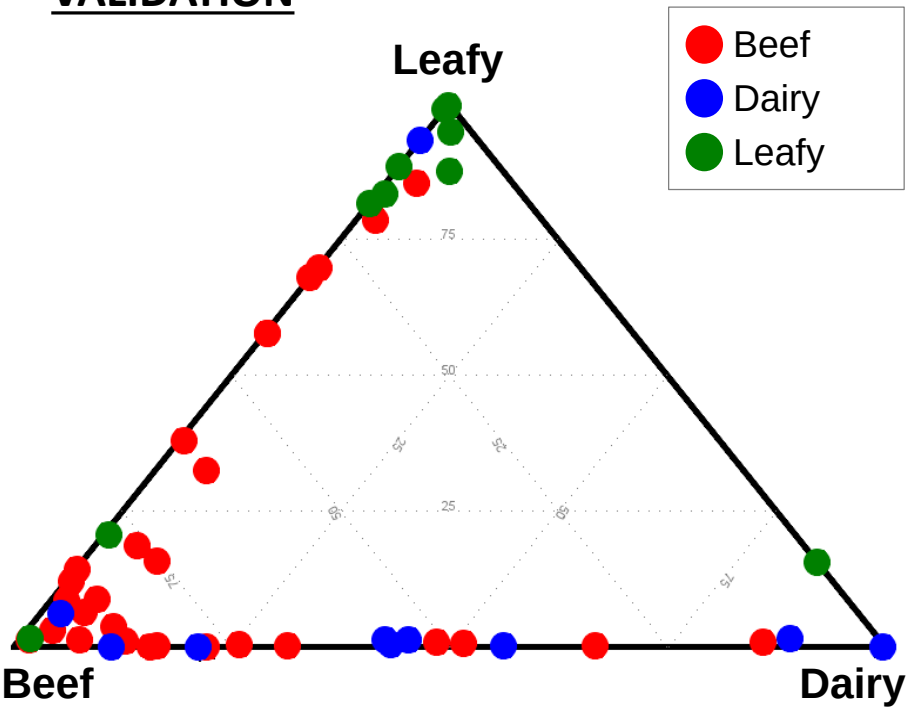
Prediction Model

DERIVATION



80% predicted by model's 1st choice
98% predicted by model's 1st or 2nd choice

VALIDATION



61% predicted by model's 1st choice
96% predicted by model's 1st or 2nd choice

Discussion

- Prior outbreak data can be used to describe and model outbreaks by food source
- Prediction model accurately differentiated food sources
  - Performed better for beef and leafy outbreaks
  - Performed well as a rule-out method
- Could be used for hypothesis generation and attribution

Discussion

- Limitations
  - Known sources of STEC illness
  - Outbreaks with complete data
- Future directions
  - Prospective validation
  - Alternative methods
    - Food categorization schemes
    - Other prediction rules
  - Other major pathogens

Acknowledgements

- Co-authors
  - Elaine Scallan, CSPH
  - Ed Bedrick, CSPH
  - Alicia Cronquist, CDPHE
- NORS team at CDC

Questions?