Trends in Reported Foodborne Illness in the United States: 1996-2013

Mark Powell
U.S. Department of Agriculture, Office of Risk Assessment and Cost-Benefit Analysis
Washington, DC

ORACBA Science, Policy, and Risk Forum
17 September, 2015
Washington, DC

INTRODUCTION

- Current statistical methods for analyzing FoodNet data make pair-wise comparisons between most recent surveillance year and one or more baseline periods.
- Advantage: avoids specifying model form for trend
- Disadvantage: can’t distinguish trends from year to year variability
- Objective: Analyze trends in reported U.S. foodborne illness with/without specifying model form for trend

DATA

- FoodNet (Foodborne Diseases Active Surveillance Network) data: 1996-2013
  - Reported illness counts by site (State_EntYr) and year
  - Population size by site and year (increased over time)
  - FoodNet composition stable since 2004

- *Campylobacter*
- *Listeria*
- *Salmonella*
- *Shigella*
- STEC O157
- *Vibrio*
- *Yersinia*

METHODS

- Poisson Log-Linear Model with Site (1996-2013)
  - $\text{Log}(E[y_{ij}]) = \log(\text{population}_{ij}) + \beta_0 + \beta_1(\text{year}_i) + \beta_j(\text{site}_j) + \varepsilon_{ij}$
  - $y_{ij} = \text{count}_{ij}$; $E[y_i] = \mu_i$ (point on predicted curve)
  - Poisson (count) model is heteroscedastic
    - Generalized Poisson dispersion: $\text{Var}[y_i] = \phi\mu_i$
    - Negative Binomial dispersion: $\text{Var}[y_i] = \mu_i + \delta\mu_i^2$
  - Differs from CDC method in that time is treated as a continuous covariate, considers generalized Poisson as well as negative binomial dispersion
- A conventional approach, but assumes log-linear trend

METHODS

- Penalized B-spline Regression
  - Semi-parametric method – no assumed trend model form
  - B-spline basis functions provide local control, local fit is insensitive to points far removed
  - Penalized form of B-spline regression is insensitive to number, placement of join-points (“knots”)
- Wide range of applications

METHODS

- Generalized Additive Model for Poisson Regression
- $\text{Log}(E[y_i]) = \log(\text{population}_i)+\beta_0+f(\text{year}_i)+\varepsilon_i$
- Smooth $f(\text{year}_i) = \sum B_k(\text{year}_i)\beta_k$
  - $B_k(x)$ = B-spline basis function
- Year (nx1 vector) $\rightarrow$ **X** (nxk matrix)
  - Fit the model with basis functions as covariates

METHODS

- At any given point, q+1 B-splines are non-zero (local control)
  - q = B-spline degree (e.g., q=3 for cubic)
  - B-splines sum to 1
- Basis dimension (k) = q + n'
  - n' = no. intervals along domain
  - e.g., 2 internal knots divides domain into n' = 3 intervals
- Eilers and Marx (1996) provides recursive algorithm for B-spline basis functions for uniformly spaced knots
- In practice, need to impose identifiability constraint → k-1 orthogonal columns (QR decomposition)
- Smoothness controlled by penalty term, fit insensitive to basis dimension

METHODS

- P-IRLS to obtain GLM likelihood maximization, s.t. smooth
- Given λ, min: $\left\|\sqrt{W}(z - X\beta)\right\|^2 + \lambda\beta^T S\beta$
  - λ = curvature penalty parameter
  - $w_i \propto [V(\mu_i)g'(\mu_i)^2]^{-1}$
  - $z_i = g'(\mu_i)(y_i - \mu_i) + X_i\beta$
  - X = design matrix (constrained)
  - g = link function (log)
  - S (penalty matrix) = D^TD
    - penalize differences among neighboring β coefficients
  - For D = second order difference matrix ~ $\int[f''(x)]^2 dx$
    - measure of total curvature

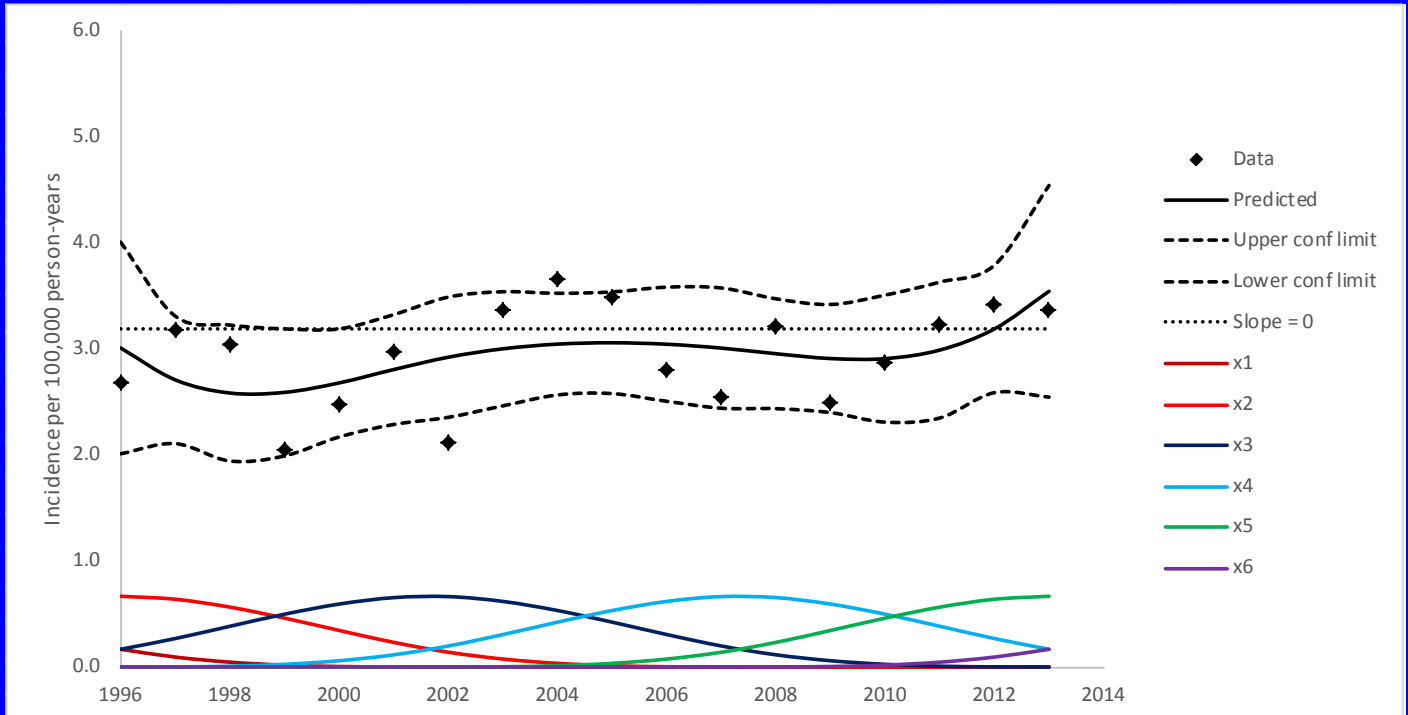
METHODS

- Select degree of smoothness (λ) based on model selection criterion (e.g., GCV, AIC)
- Effective degrees of freedom (edf) = tr(A)
  - where $\hat{\mu} = Ay$
- With λ = 0, tr(A) = k
- As λ→∞, GAM → Poisson Log-Linear Model (X →1 edf (Year))

METHODS

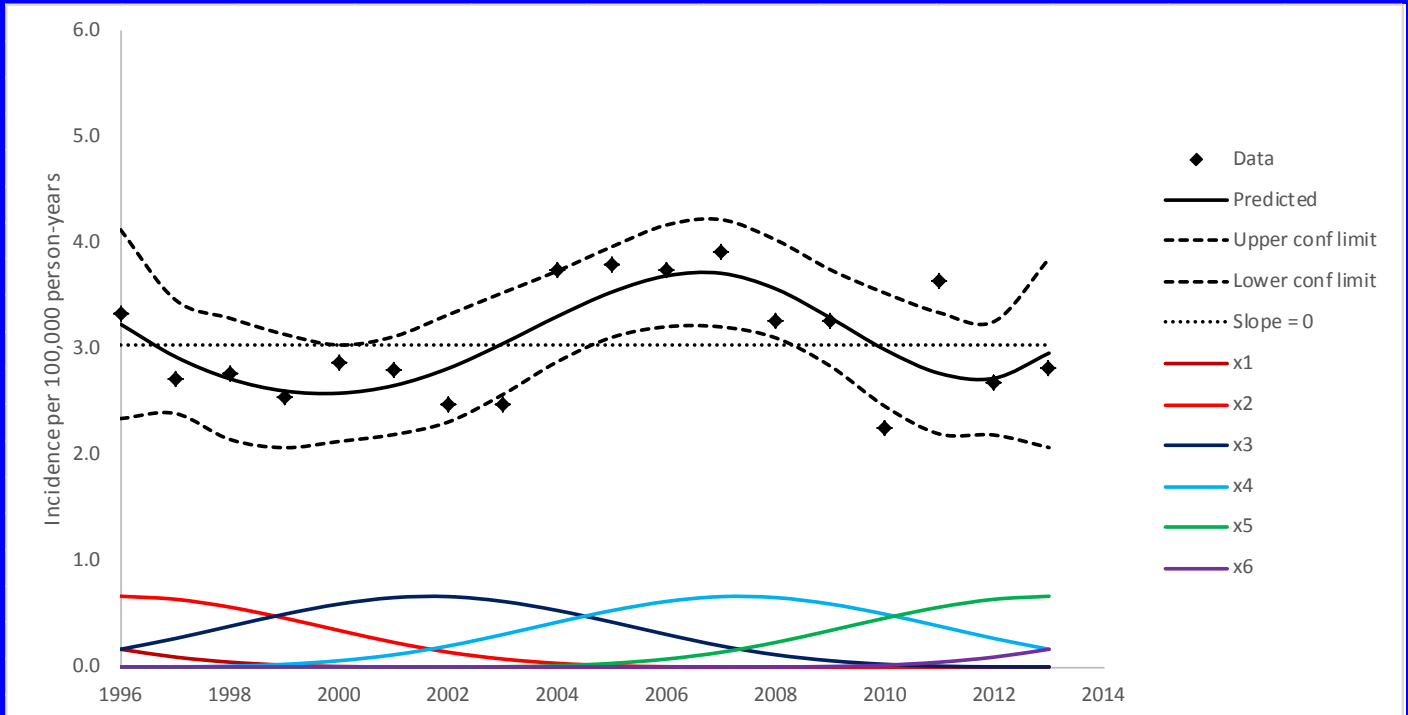
- X: cubic B-spline basis with 2 internal knots
  - k = q(3)+ n’(3) = 6 unconstrained basis functions
- S (penalty matrix): 2nd order difference matrix
- All Sites (1996-2013)
  - Composition of FoodNet sites stable since 2004
- Original 5 Sites (1996-2013)
  - Attempt to control for changes in FoodNet composition over time
- Ljung-Box test for serial correlation (AR(1))

METHODS
B-Spline Provides Flexibility



	PENALTY	0.00										
	AICC	0.995143										

METHODS
Penalized B-Spline Avoids Overfitting



	PENALTY	0.00										
	AICC	0.766499										

PRELIMINARY RESULTS

- Poisson Log-Linear Model

Poisson Log Linear Model

CAMPYLOBACTER	Generalized Poisson				Negative Binomial			
Parameter	Est	SE	ChiSq	p	Est	SE	ChiSq	p
Intercept	5.2702	3.7713	1.9528	0.1623	1.3240	4.1002	0.1043	0.7468
Year	-0.0073	0.0019	15.2107	0.0001	-0.0054	0.0020	6.8889	0.0087
CA_1996	1.4694	0.0579	644.8468	0.0000	1.4770	0.0627	555.1890	0.0000
CA_2000	1.1268	0.0701	258.5311	0.0000	1.1290	0.0668	285.8118	0.0000
CO_2001	0.5648	0.0674	70.2924	0.0000	0.5703	0.0674	71.6515	0.0000
CO_2002	1.0996	0.1021	116.0965	0.0000	1.1015	0.0739	221.9737	0.0000
CT_1996	0.5866	0.0670	76.6550	0.0000	0.5945	0.0638	86.8874	0.0000
CT_1997	0.9476	0.0719	173.7944	0.0000	0.9543	0.0650	215.5177	0.0000
CT_1998	0.6183	0.0798	60.0296	0.0000	0.6237	0.0668	87.1792	0.0000
GA_1996	-0.0885	0.0670	1.7447	0.1866	-0.0648	0.0641	1.0219	0.3121
GA_1997	-0.2269	0.0919	6.0899	0.0136	-0.2020	0.0684	8.7062	0.0032
GA_1999	-0.0152	0.0651	0.0546	0.8152	-0.0125	0.0653	0.0365	0.8485
MD_1998	0.0524	0.0701	0.5584	0.4549	0.0572	0.0653	0.7685	0.3807
MD_2001	-0.1294	0.0857	2.2819	0.1309	-0.1295	0.0700	3.4192	0.0644
MD_2002	0.0414	0.0939	0.1942	0.6595	0.0401	0.0725	0.3058	0.5803
MN_1996	0.8194	0.0571	205.5809	0.0000	0.8272	0.0626	174.8439	0.0000
NM_2004	0.7641	0.0705	117.4649	0.0000	0.7639	0.0716	113.7315	0.0000
NY_1998	0.6488	0.0740	76.8452	0.0000	0.6537	0.0659	98.3130	0.0000
NY_1999	0.4326	0.0822	27.7239	0.0000	0.4377	0.0678	41.6936	0.0000
NY_2002	0.1121	0.0932	1.4470	0.2290	0.1155	0.0724	2.5486	0.1104
NY_2003	0.7309	0.0963	57.5967	0.0000	0.7327	0.0741	97.7662	0.0000
NY_2004	0.6802	0.1294	27.6526	0.0000	0.6812	0.0818	69.2886	0.0000
OR_1996	0.8676	0.0585	220.2090	0.0000	0.8727	0.0626	194.3326	0.0000
TN_2000	-0.2411	0.0746	10.4626	0.0012	-0.2372	0.0675	12.3719	0.0004
TN_2003	0.0000	0.0000			0.0000	0.0000		
Scale/Dispersion	2.7423	0.0000			0.0234	0.0023		
SSPR	324.4895				367.1452			

Poisson Log Linear Model

LISTERIA	Generalized Poisson				Negative Binomial			
Parameter	Est	SE	ChiSq	p	Est	SE	ChiSq	p
Intercept	31.8912	11.7758	7.3344	0.0068	30.6845	11.5738	7.0289	0.0080
Year	-0.0224	0.0059	14.5477	0.0001	-0.0218	0.0058	14.2617	0.0002
CA_1996	0.7329	0.1663	19.4344	0.0000	0.7411	0.1647	20.2577	0.0000
CA_2000	0.2308	0.2367	0.9508	0.3295	0.2349	0.2171	1.1700	0.2794
CO_2001	0.2513	0.1915	1.7225	0.1894	0.2559	0.1854	1.9048	0.1675
CO_2002	-0.1350	0.4529	0.0888	0.7657	-0.1325	0.3852	0.1182	0.7309
CT_1996	1.0383	0.1655	39.3409	0.0000	1.0441	0.1643	40.3740	0.0000
CT_1997	0.6363	0.2048	9.6535	0.0019	0.6420	0.1923	11.1473	0.0008
CT_1998	0.5177	0.2168	5.7050	0.0169	0.5222	0.2014	6.7204	0.0095
GA_1996	0.2290	0.1686	1.8456	0.1743	0.2333	0.1666	1.9594	0.1616
GA_1997	-0.1576	0.2370	0.4424	0.5060	-0.1537	0.2158	0.5072	0.4763
GA_1999	-0.2658	0.1808	2.1595	0.1417	-0.2639	0.1765	2.2340	0.1350
MD_1998	0.3572	0.1752	4.1573	0.0415	0.3583	0.1721	4.3349	0.0373
MD_2001	0.4628	0.1936	5.7173	0.0168	0.4642	0.1869	6.1660	0.0130
MD_2002	0.2546	0.2330	1.1941	0.2745	0.2565	0.2160	1.4107	0.2349
MN_1996	-0.4035	0.1744	5.3538	0.0207	-0.4024	0.1706	5.5605	0.0184
NM_2004	0.1277	0.2224	0.3299	0.5657	0.1307	0.2106	0.3854	0.5347
NY_1998	0.4262	0.2070	4.2384	0.0395	0.4263	0.1945	4.8038	0.0284
NY_1999	0.8784	0.1934	20.6396	0.0000	0.8841	0.1852	22.7803	0.0000
NY_2002	0.2756	0.2347	1.3786	0.2403	0.2770	0.2173	1.6249	0.2024
NY_2003	0.5925	0.2697	4.8268	0.0280	0.5983	0.2440	6.0108	0.0142
NY_2004	-0.0717	0.4813	0.0222	0.8816	-0.0699	0.4089	0.0292	0.8642
OR_1996	0.1809	0.1681	1.1592	0.2816	0.1860	0.1660	1.2549	0.2626
TN_2000	-0.3213	0.2018	2.5339	0.1114	-0.3185	0.1919	2.7537	0.0970
TN_2003	0.0000	0.0000			0.0000	0.0000		
Scale/Dispersion	1.2178	0.0000			0.0576	0.0166		
SSPR	312.7524				343.3450			

Poisson Log Linear Model

SHIGELLA	Generalized Poisson				Negative Binomial			
Parameter	Est	SE	ChiSq	p	Est	SE	ChiSq	p
Intercept	70.2796	14.8840	22.2957	0.0000	92.9580	14.4239	41.5342	0.0000
Year	-0.0400	0.0074	29.1268	0.0000	-0.0513	0.0072	51.0144	0.0000
CA_1996	0.7497	0.2071	13.1059	0.0003	0.6995	0.2123	10.8562	0.0010
CA_2000	-0.2412	0.3472	0.4828	0.4872	-0.2859	0.2261	1.5987	0.2061
CO_2001	-0.1156	0.2629	0.1933	0.6602	-0.1501	0.2270	0.4376	0.5083
CO_2002	-0.1075	0.5604	0.0368	0.8479	-0.1055	0.2423	0.1896	0.6633
CT_1996	-0.9577	0.3310	8.3697	0.0038	-1.0313	0.2165	22.6990	0.0000
CT_1997	-0.5555	0.3728	2.2202	0.1362	-0.5737	0.2190	6.8636	0.0088
CT_1998	-1.5342	0.5896	6.7705	0.0093	-1.5275	0.2313	43.5984	0.0000
GA_1996	0.7893	0.1975	15.9691	0.0001	0.7718	0.2115	13.3164	0.0003
GA_1997	0.5018	0.2448	4.2020	0.0404	0.4750	0.2147	4.8963	0.0269
GA_1999	0.8001	0.1956	16.7389	0.0000	0.8156	0.2184	13.9476	0.0002
MD_1998	0.1473	0.2270	0.4208	0.5166	0.0908	0.2172	0.1747	0.6760
MD_2001	-0.2599	0.2962	0.7695	0.3804	-0.2795	0.2277	1.5060	0.2198
MD_2002	-1.1408	0.4994	5.2174	0.0224	-1.1658	0.2394	23.7056	0.0000
MN_1996	-0.0538	0.2067	0.0677	0.7947	-0.0945	0.2121	0.1986	0.6558
NM_2004	0.3747	0.2613	2.0560	0.1516	0.3853	0.2413	2.5500	0.1103
NY_1998	-0.0669	0.2965	0.0509	0.8215	-0.0206	0.2181	0.0090	0.9246
NY_1999	-1.1297	0.4786	5.5724	0.0182	-1.1165	0.2281	23.9606	0.0000
NY_2002	-1.5188	0.6027	6.3502	0.0117	-1.5461	0.2448	39.8839	0.0000
NY_2003	-2.0499	1.0981	3.4849	0.0619	-2.0442	0.2841	51.7593	0.0000
NY_2004	-1.7235	1.3391	1.6566	0.1981	-1.6948	0.3102	29.8559	0.0000
OR_1996	-0.6654	0.2445	7.4080	0.0065	-0.7284	0.2136	11.6238	0.0007
TN_2000	0.6466	0.2089	9.5844	0.0020	0.6808	0.2217	9.4262	0.0021
TN_2003	0.0000	0.0000			0.0000	0.0000		
Scale/Dispersion	6.7682	0.0000			0.2969	0.0239		
SSPR	388.0268				442.6833			

Poisson Log Linear Model

STECO157	Generalized Poisson				Negative Binomial			
Parameter	Est	SE	ChiSq	p	Est	SE	ChiSq	p
Intercept	60.4425	8.1942	54.4097	0.0000	61.8459	9.7252	40.4410	0.0000
Year	-0.0358	0.0041	77.0799	0.0000	-0.0365	0.0048	56.8915	0.0000
CA_1996	0.0216	0.1306	0.0274	0.8686	0.0348	0.1388	0.0629	0.8020
CA_2000	-0.0439	0.1810	0.0589	0.8082	-0.0487	0.1593	0.0934	0.7600
CO_2001	0.2873	0.1330	4.6705	0.0307	0.2775	0.1461	3.6056	0.0576
CO_2002	0.7558	0.2168	12.1475	0.0005	0.7463	0.1753	18.1318	0.0000
CT_1996	-0.0116	0.1398	0.0068	0.9341	-0.0473	0.1426	0.1102	0.7399
CT_1997	-0.0037	0.1716	0.0005	0.9828	-0.0216	0.1531	0.0199	0.8879
CT_1998	0.3837	0.1566	6.0022	0.0143	0.3542	0.1494	5.6238	0.0177
GA_1996	-1.0131	0.1529	43.8881	0.0000	-1.0206	0.1457	49.0583	0.0000
GA_1997	-0.8448	0.2100	16.1853	0.0001	-0.8557	0.1667	26.3378	0.0000
GA_1999	-1.0657	0.1532	48.4024	0.0000	-1.0672	0.1489	51.3870	0.0000
MD_1998	-0.6242	0.1530	16.6468	0.0000	-0.6288	0.1478	18.1078	0.0000
MD_2001	-1.3010	0.2446	28.2914	0.0000	-1.3084	0.1843	50.3875	0.0000
MD_2002	-0.5041	0.2141	5.5460	0.0185	-0.5041	0.1742	8.3709	0.0038
MN_1996	1.0093	0.1057	91.2273	0.0000	0.9923	0.1328	55.8516	0.0000
NM_2004	-0.5208	0.1948	7.1502	0.0075	-0.5263	0.1720	9.3640	0.0022
NY_1998	0.3425	0.1472	5.4164	0.0199	0.3321	0.1461	5.1669	0.0230
NY_1999	0.5312	0.1471	13.0363	0.0003	0.4783	0.1479	10.4520	0.0012
NY_2002	-0.1659	0.1913	0.7521	0.3858	-0.1778	0.1663	1.1424	0.2851
NY_2003	0.2577	0.2150	1.4367	0.2307	0.2572	0.1767	2.1183	0.1455
NY_2004	0.2940	0.2875	1.0458	0.3065	0.2759	0.2073	1.7723	0.1831
OR_1996	0.7089	0.1105	41.1495	0.0000	0.7054	0.1336	27.8627	0.0000
TN_2000	-0.2787	0.1395	3.9923	0.0457	-0.2877	0.1463	3.8664	0.0493
TN_2003	0.0000	0.0000			0.0000	0.0000		
Scale/Dispersion	1.8314	0.0000			0.0955	0.0123		
SSPR	339.9408				383.9270			

Poisson Log Linear Model

YERSINIA	Generalized Poisson				Negative Binomial			
Parameter	Est	SE	ChiSq	p	Est	SE	ChiSq	p
Intercept	71.1822	10.8817	42.7905	0.0000	54.3429	11.5051	22.3104	0.0000
Year	-0.0419	0.0054	59.6345	0.0000	-0.0335	0.0057	34.1159	0.0000
CA_1996	0.7576	0.1693	20.0160	0.0000	0.7601	0.1675	20.5870	0.0000
CA_2000	0.3920	0.2301	2.9012	0.0885	0.3845	0.2049	3.5222	0.0606
CO_2001	-0.1261	0.2163	0.3396	0.5600	-0.1259	0.1975	0.4064	0.5238
CO_2002	0.1699	0.4067	0.1745	0.6761	0.1641	0.3205	0.2620	0.6087
CT_1996	0.4827	0.1839	6.8846	0.0087	0.5019	0.1751	8.2133	0.0042
CT_1997	0.5620	0.2119	7.0362	0.0080	0.5841	0.1916	9.2929	0.0023
CT_1998	0.1967	0.2426	0.6576	0.4174	0.2141	0.2107	1.0329	0.3095
GA_1996	1.0606	0.1583	44.8977	0.0000	1.0648	0.1620	43.2090	0.0000
GA_1997	-0.3102	0.2529	1.5053	0.2199	-0.2984	0.2163	1.9025	0.1678
GA_1999	0.1568	0.1726	0.8257	0.3635	0.1757	0.1706	1.0617	0.3028
MD_1998	-0.0580	0.1940	0.0895	0.7648	-0.0496	0.1820	0.0743	0.7852
MD_2001	-0.3103	0.2470	1.5780	0.2090	-0.3116	0.2159	2.0817	0.1491
MD_2002	-1.1229	0.4067	7.6250	0.0058	-1.1276	0.3206	12.3728	0.0004
MN_1996	0.3298	0.1624	4.1228	0.0423	0.3507	0.1631	4.6218	0.0316
NM_2004	-1.0001	0.3461	8.3476	0.0039	-1.0183	0.2825	12.9912	0.0003
NY_1998	0.3008	0.2177	1.9091	0.1671	0.3180	0.1958	2.6387	0.1043
NY_1999	0.2455	0.2340	1.1015	0.2939	0.2596	0.2062	1.5849	0.2081
NY_2002	0.6712	0.2147	9.7711	0.0018	0.6726	0.1980	11.5405	0.0007
NY_2003	0.3831	0.3009	1.6210	0.2030	0.3797	0.2520	2.2698	0.1319
NY_2004	0.1228	0.4568	0.0723	0.7880	0.1108	0.3567	0.0966	0.7560
OR_1996	0.3923	0.1672	5.5057	0.0190	0.4315	0.1651	6.8312	0.0090
TN_2000	0.4560	0.1763	6.6902	0.0097	0.4559	0.1739	6.8775	0.0087
TN_2003	0.0000	0.0000			0.0000	0.0000		
Scale/Dispersion	1.3696	0.0000			0.0832	0.0159		
SSPR	303.0346				318.5979			

Poisson Log Linear Model

VIBRIO	Generalized Poisson				Negative Binomial			
Parameter	Est	SE	ChiSq	p	Est	SE	ChiSq	p
Intercept	-135.0769	12.7410	112.3980	0.0000	-139.7475	12.6327	122.3765	0.0000
Year	0.0605	0.0063	91.0847	0.0000	0.0629	0.0063	99.8885	0.0000
CA_1996	1.9104	0.2077	84.6356	0.0000	1.9634	0.1930	103.5117	0.0000
CA_2000	1.5084	0.2421	38.8298	0.0000	1.5194	0.2169	49.0542	0.0000
CO_2001	0.4548	0.2551	3.1789	0.0746	0.4729	0.2265	4.3604	0.0368
CO_2002	0.8712	0.4117	4.4778	0.0343	0.8729	0.3355	6.7691	0.0093
CT_1996	1.1023	0.2323	22.5219	0.0000	1.0794	0.2079	26.9539	0.0000
CT_1997	1.8181	0.2308	62.0297	0.0000	1.7949	0.2075	74.8268	0.0000
CT_1998	1.3594	0.2528	28.9156	0.0000	1.3530	0.2225	36.9818	0.0000
GA_1996	0.7561	0.2215	11.6529	0.0006	0.8212	0.2021	16.5049	0.0000
GA_1997	0.6558	0.2706	5.8724	0.0154	0.6857	0.2349	8.5210	0.0035
GA_1999	0.8528	0.2157	15.6372	0.0001	0.8656	0.1991	18.8939	0.0000
MD_1998	1.6179	0.2106	59.0278	0.0000	1.6027	0.1950	67.5274	0.0000
MD_2001	0.7839	0.2518	9.6941	0.0018	0.7967	0.2242	12.6333	0.0004
MD_2002	1.8375	0.2263	65.9417	0.0000	1.8328	0.2086	77.1963	0.0000
MN_1996	0.1651	0.2270	0.5288	0.4671	0.1406	0.2046	0.4721	0.4920
NM_2004	-0.4443	0.3644	1.4869	0.2227	-0.4444	0.3039	2.1382	0.1437
NY_1998	0.1527	0.3259	0.2195	0.6394	0.1473	0.2723	0.2928	0.5884
NY_1999	0.7408	0.2855	6.7352	0.0095	0.7378	0.2450	9.0669	0.0026
NY_2002	0.2930	0.3257	0.8092	0.3684	0.2988	0.2747	1.1834	0.2767
NY_2003	-0.4513	0.5689	0.6292	0.4277	-0.4518	0.4502	1.0067	0.3157
NY_2004	0.3860	0.5318	0.5270	0.4679	0.3895	0.4238	0.8450	0.3580
OR_1996	0.9713	0.2152	20.3679	0.0000	0.9838	0.1973	24.8559	0.0000
TN_2000	0.2283	0.2503	0.8316	0.3618	0.2372	0.2224	1.1378	0.2861
TN_2003	0.0000	0.0000			0.0000	0.0000		
Scale/Dispersion	1.3107	0.0000			0.0760	0.0188		
SSPR	299.1456				350.8983			

Poisson Log Linear Model

SALMONELLA	Generalized Poisson				Negative Binomial			
Parameter	Est	SE	ChiSq	p	Est	SE	ChiSq	p
Intercept	-20.6363	3.4426	35.9323	0.0000	-9.1893	3.5286	6.7819	0.0092
Year	0.0059	0.0017	11.7123	0.0006	0.0002	0.0018	0.0088	0.9252
CA_1996	0.1727	0.0477	13.1290	0.0003	0.1559	0.0516	9.1135	0.0025
CA_2000	-0.1217	0.0695	3.0656	0.0800	-0.1291	0.0573	5.0686	0.0244
CO_2001	-0.2012	0.0556	13.1079	0.0003	-0.1990	0.0556	12.7995	0.0003
CO_2002	-0.1040	0.1152	0.8145	0.3668	-0.1007	0.0693	2.1076	0.1466
CT_1996	-0.0770	0.0540	2.0348	0.1537	-0.0909	0.0527	2.9762	0.0845
CT_1997	0.1208	0.0628	3.7040	0.0543	0.1077	0.0545	3.8996	0.0483
CT_1998	-0.1339	0.0705	3.6037	0.0577	-0.1433	0.0565	6.4393	0.0112
GA_1996	-0.0248	0.0457	0.2931	0.5882	-0.0392	0.0514	0.5822	0.4455
GA_1997	0.1068	0.0571	3.4941	0.0616	0.0867	0.0535	2.6220	0.1054
GA_1999	0.8355	0.0401	434.8496	0.0000	0.8208	0.0520	249.4500	0.0000
MD_1998	0.1623	0.0473	11.7939	0.0006	0.1513	0.0524	8.3269	0.0039
MD_2001	-0.2150	0.0605	12.6415	0.0004	-0.2174	0.0564	14.8438	0.0001
MD_2002	0.2238	0.0606	13.6362	0.0002	0.2249	0.0573	15.4093	0.0001
MN_1996	-0.0702	0.0431	2.6497	0.1036	-0.0870	0.0510	2.9154	0.0877
NM_2004	0.1318	0.0567	5.4058	0.0201	0.1358	0.0589	5.3201	0.0211
NY_1998	-0.0548	0.0627	0.7645	0.3819	-0.0659	0.0550	1.4354	0.2309
NY_1999	-0.2100	0.0702	8.9564	0.0028	-0.2181	0.0569	14.7039	0.0001
NY_2002	-0.4300	0.0780	30.3702	0.0000	-0.4308	0.0605	50.6285	0.0000
NY_2003	-0.1997	0.0949	4.4255	0.0354	-0.1975	0.0652	9.1849	0.0024
NY_2004	-0.1953	0.1299	2.2592	0.1328	-0.1903	0.0751	6.4226	0.0113
OR_1996	-0.3257	0.0479	46.2627	0.0000	-0.3404	0.0517	43.3810	0.0000
TN_2000	0.0406	0.0480	0.7173	0.3970	0.0377	0.0537	0.4908	0.4836
TN_2003	0.0000	0.0000			0.0000	0.0000		
Scale/Dispersion	2.5088	0.0000			0.0155	0.0016		
SSPR	316.7114				356.1833			

Poisson Log Linear Model

- Significant Negative Trend Term
  - *Campylobacter*, *Listeria*, *Shigella*, STEC O157, and *Yersinia*
- Significant Positive Trend Term
  - *Vibrio*
- Significant Site-Level Effects
  - All pathogens

Poisson Log Linear Model

- *Salmonella*
  - Significant Positive Trend Term for Generalized Poisson
  - No Significant Trend Term for Negative Binomial
- Generalized Poisson better fit than Negative Binomial
  - All pathogens (incl. *Salmonella*)
- Log linear model not a good overall fit

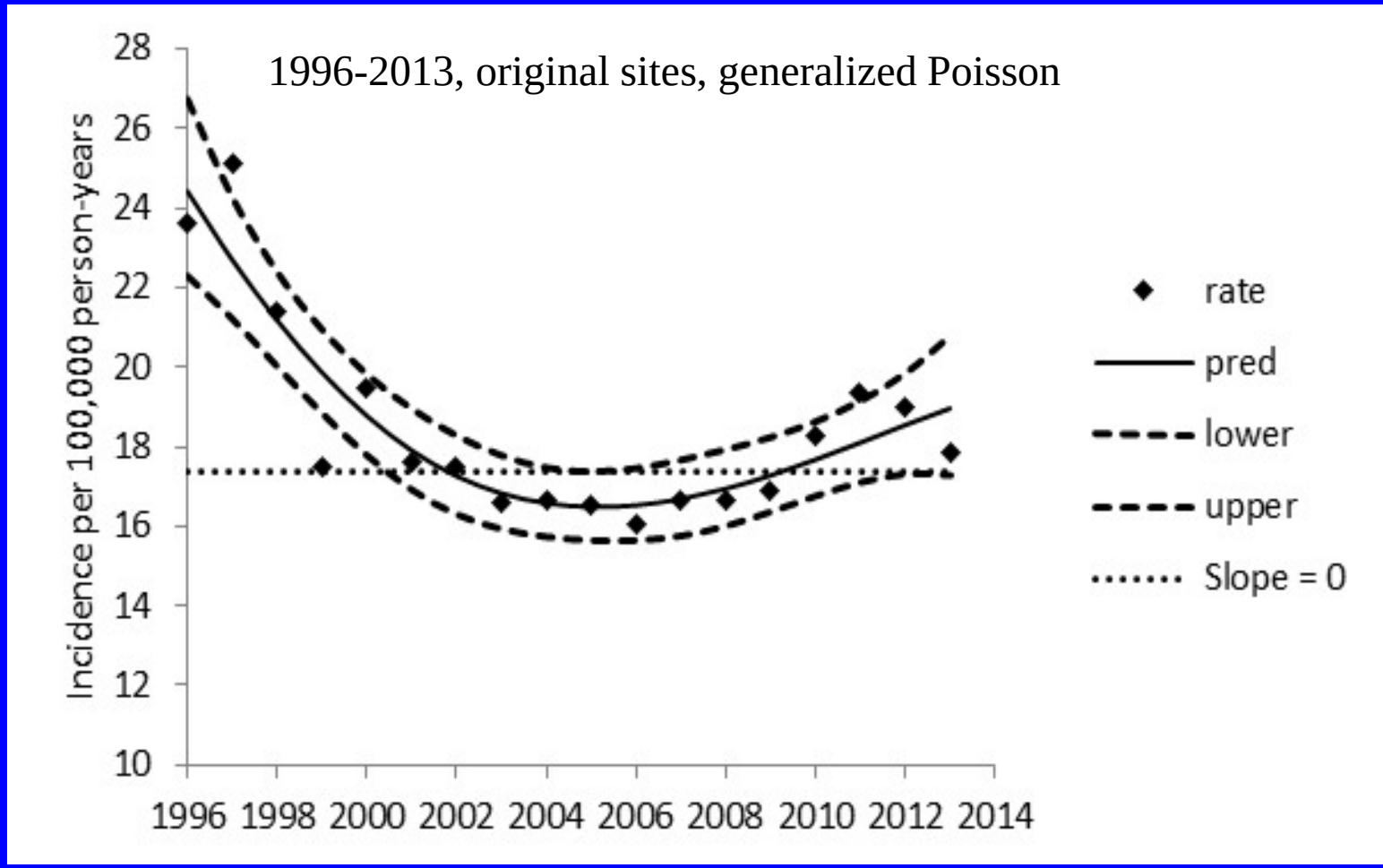
PRELIMINARY RESULTS

- Penalized B-Spline Regression

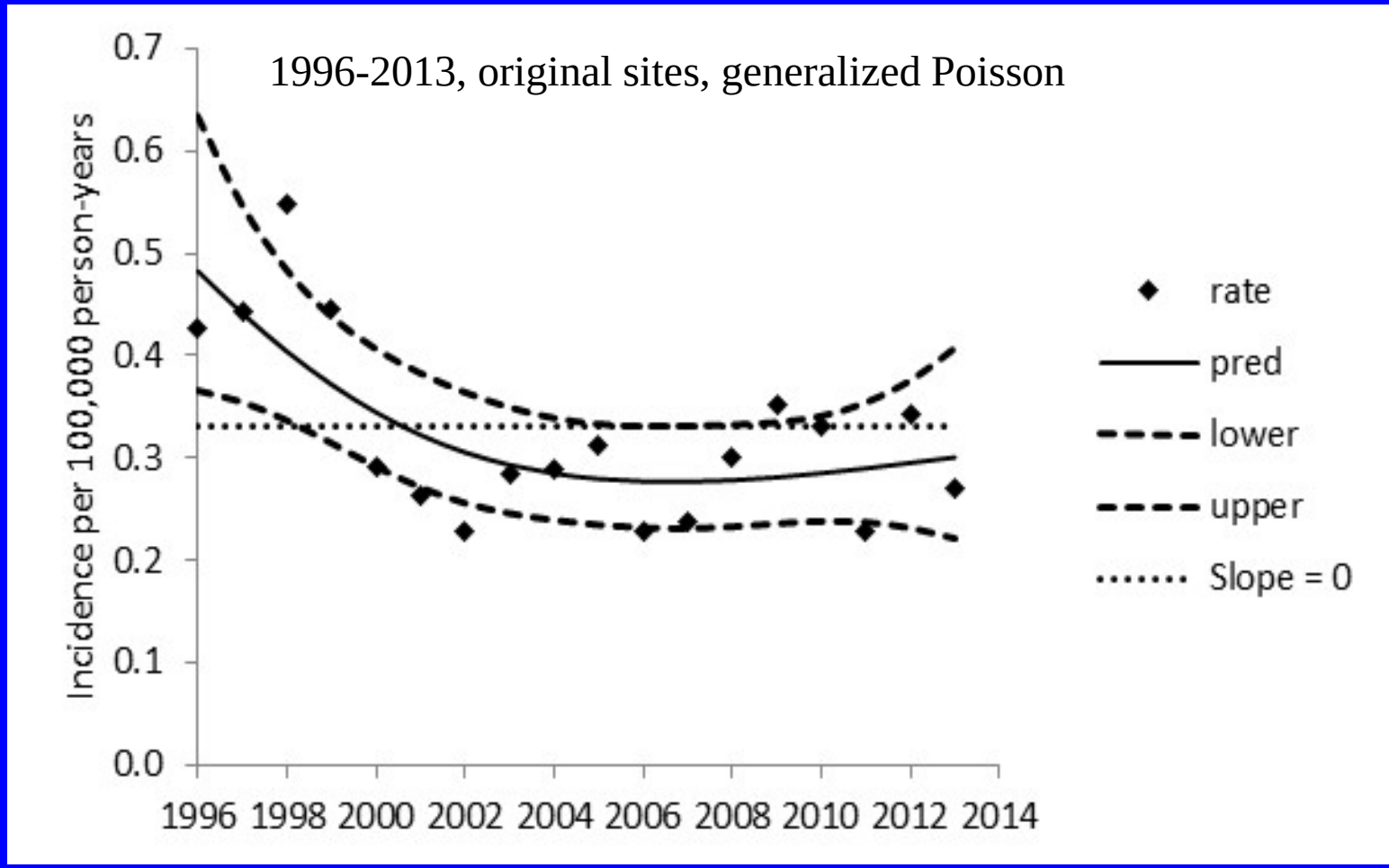
Penalized B-Spline Regression

- *Campylobacter*, *Listeria*, STEC O157, *Yersinia*
- Early declines followed by a period of no significant trend
- Results consistent for All Sites, Original Sites, Generalized Poisson, Negative Binomial

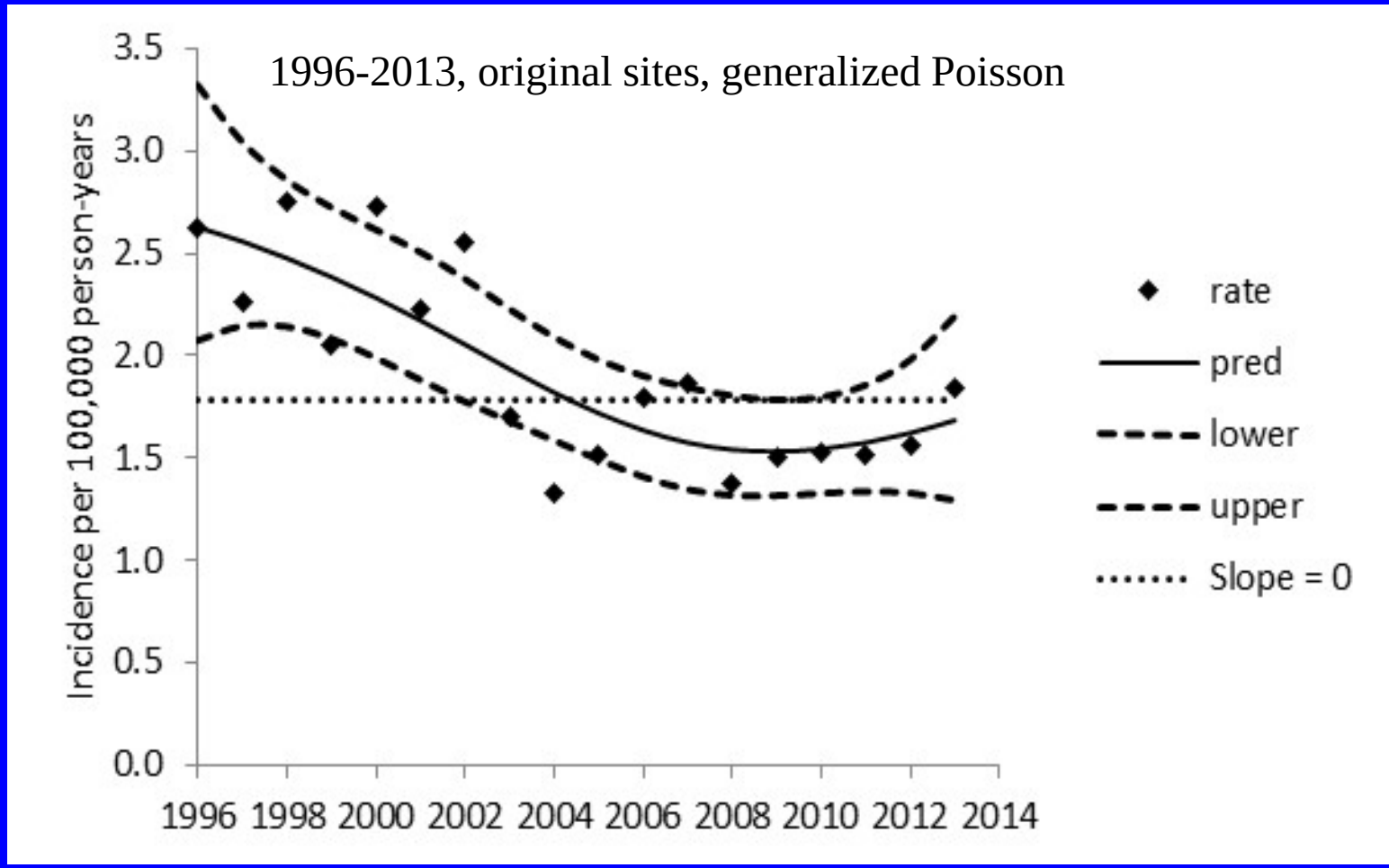
Campylobacter



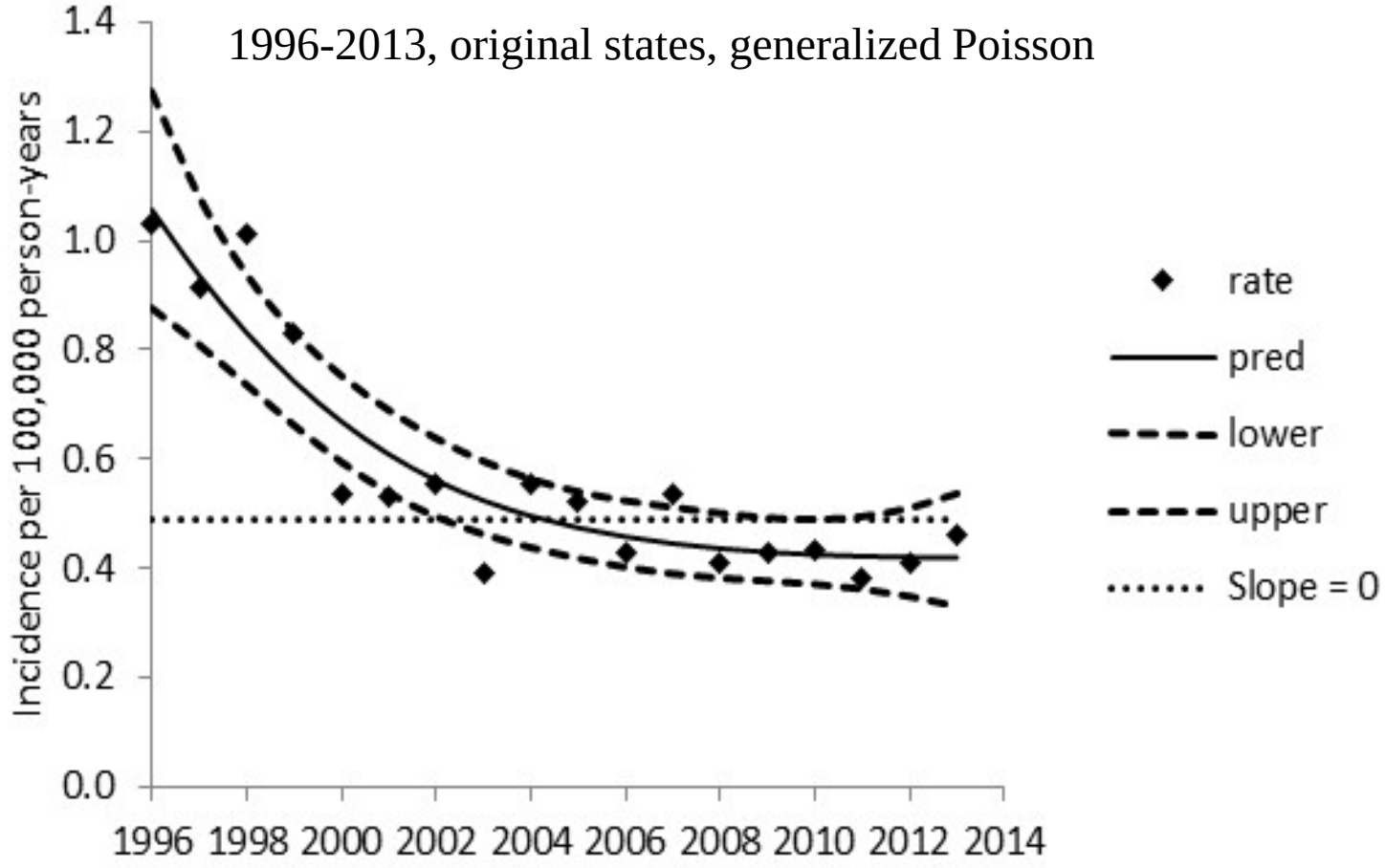
Listeria


1996-2013, original sites, generalized Poisson
Incidence per 100,000 person-years
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
1996 1998 2000 2002 2004 2006 2008 2010 2012 2014
rate
pred
lower
upper
Slope = 0

STEC O157


1996-2013, original sites, generalized Poisson
Incidence per 100,000 person-years
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
1996 1998 2000 2002 2004 2006 2008 2010 2012 2014
rate
pred
lower
upper
Slope = 0

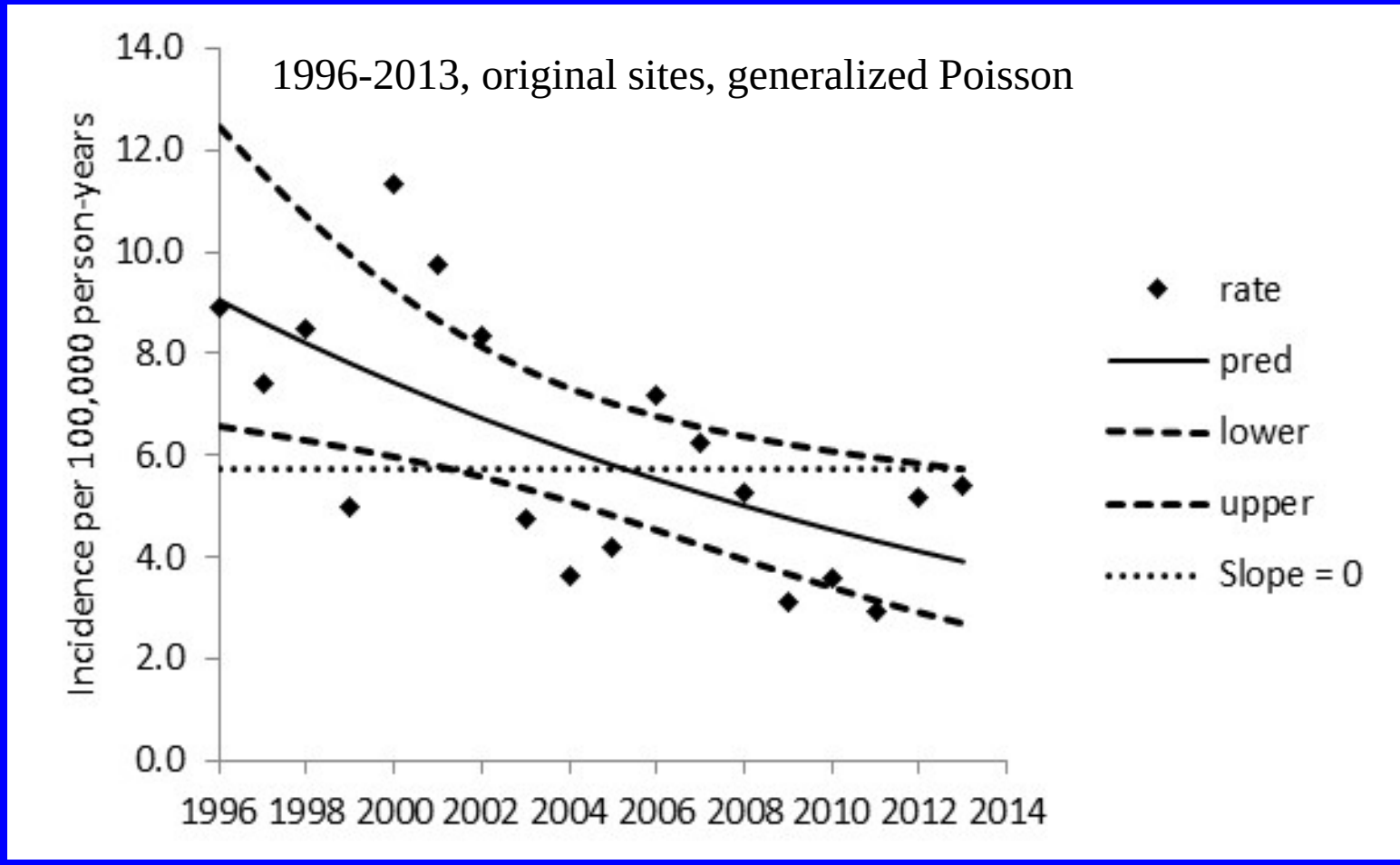
Yersinia


1996-2013, original states, generalized Poisson
Incidence per 100,000 person-years
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0.0
1996 1998 2000 2002 2004 2006 2008 2010 2012 2014
rate
pred
lower
upper
Slope = 0

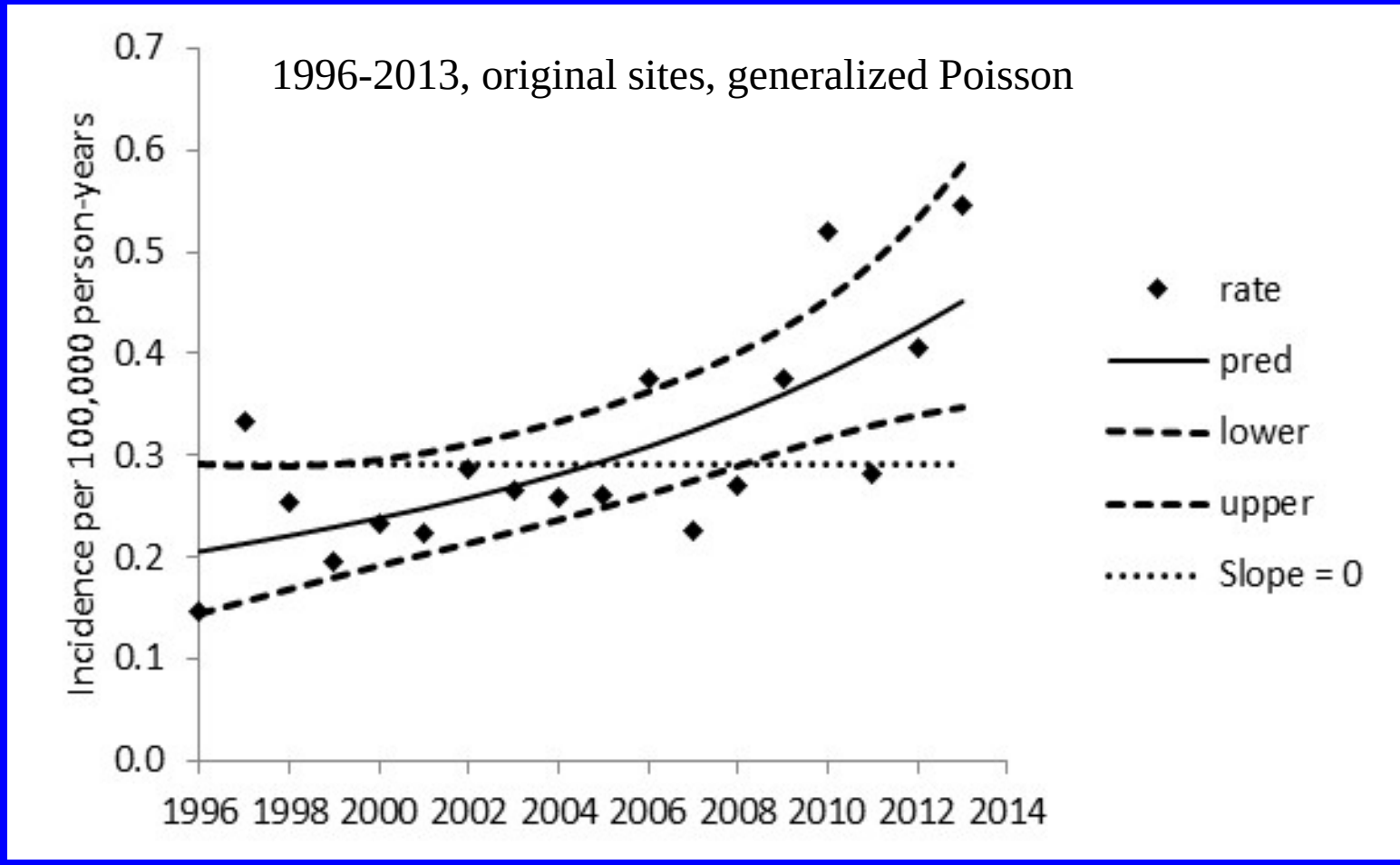
Penalized B-Spline Regression

- *Shigella* and *Vibrio*
- Continuous trends without an apparent plateau
- *Shigella* decreasing
- *Vibrio* increasing
- Results consistent for All Sites, Original Sites, Generalized Poisson, Negative Binomial

Shigella


1996-2013, original sites, generalized Poisson
Incidence per 100,000 person-years
14.0
12.0
10.0
8.0
6.0
4.0
2.0
0.0
1996 1998 2000 2002 2004 2006 2008 2010 2012 2014
rate
pred
lower
upper
Slope = 0

Vibrio

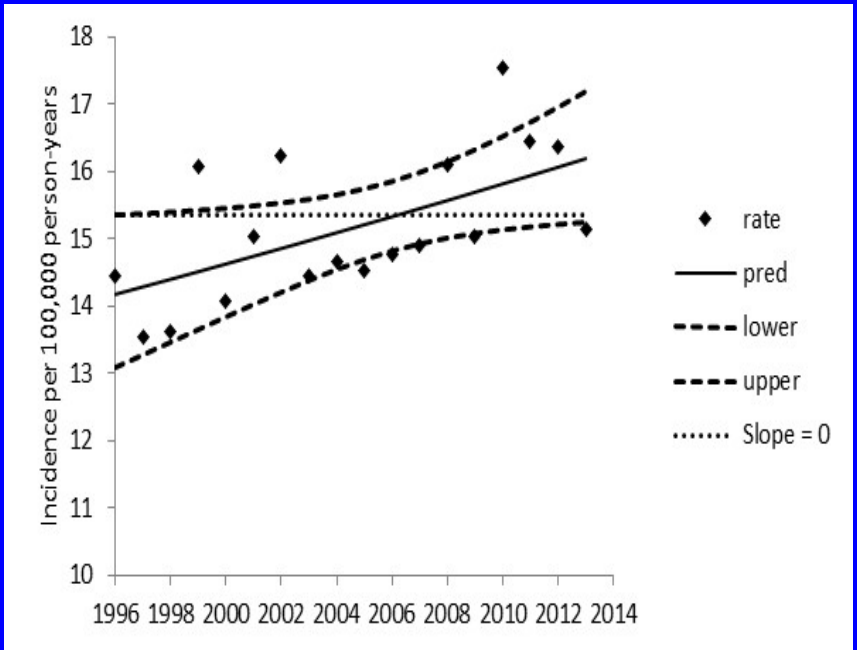

1996-2013, original sites, generalized Poisson
Incidence per 100,000 person-years
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
1996 1998 2000 2002 2004 2006 2008 2010 2012 2014
rate
pred
lower
upper
Slope = 0

Penalized B-Spline Regression

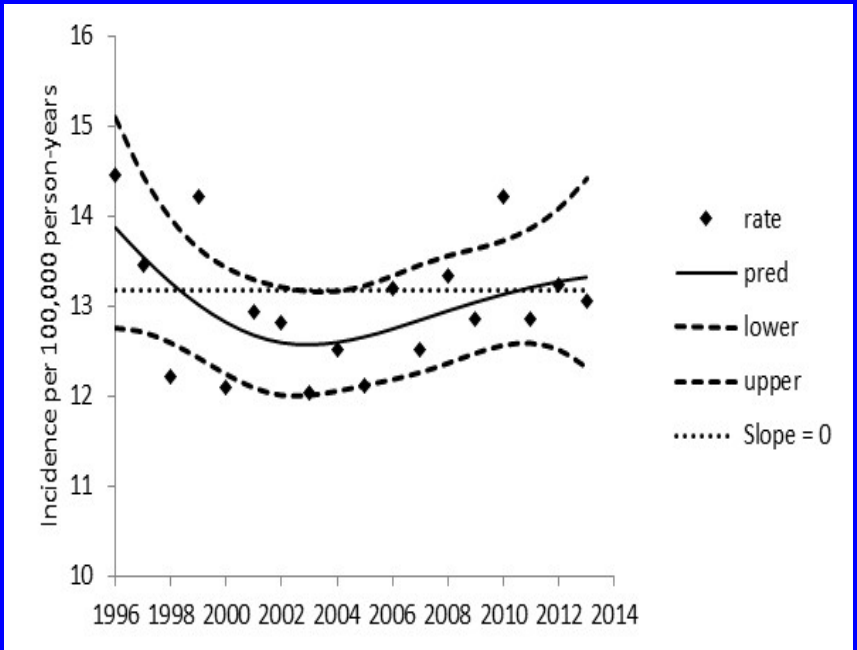
- *Salmonella*
- Inconclusive whether there is no trend or an increasing trend
  - Significant Trend in All Sites, 1996-2013
    - Inf. Smooth → Log-linear model
    - Negative auto-correlation (SE(LP) over stated) - GAMM
  - No Trend in Original 5 Sites, 1996-2013

Salmonella

All Sites, 1996-2013



Original Sites, 1996-2013



Salmonella

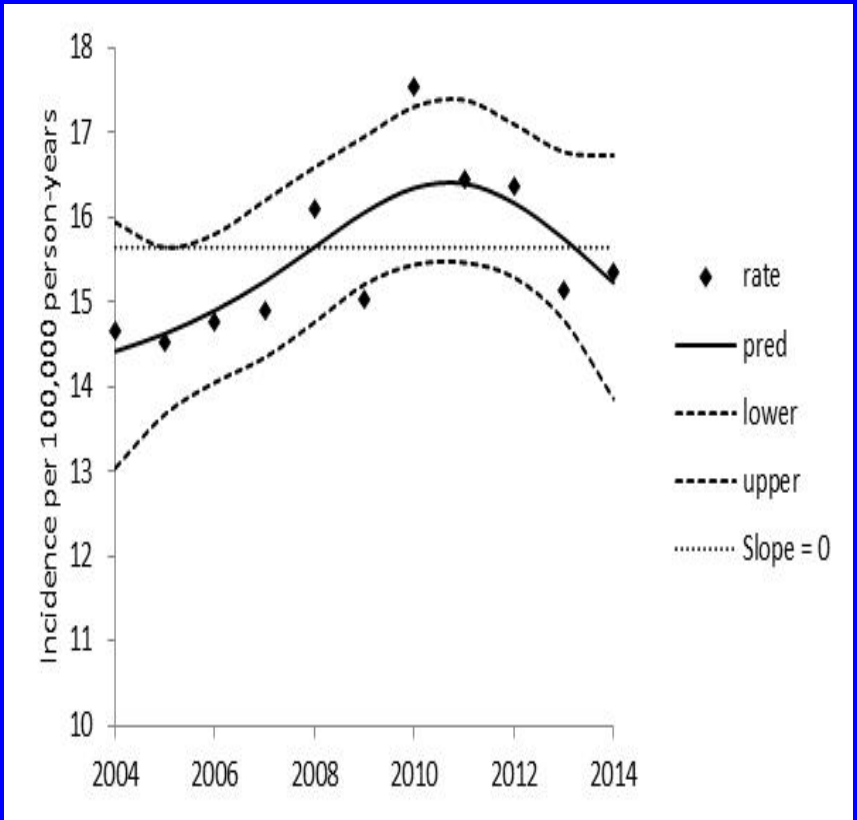
- Restrict the analysis to data since 2004 (FoodNet composition is stable)

Poisson Log Linear Model Salmonella 2004-13

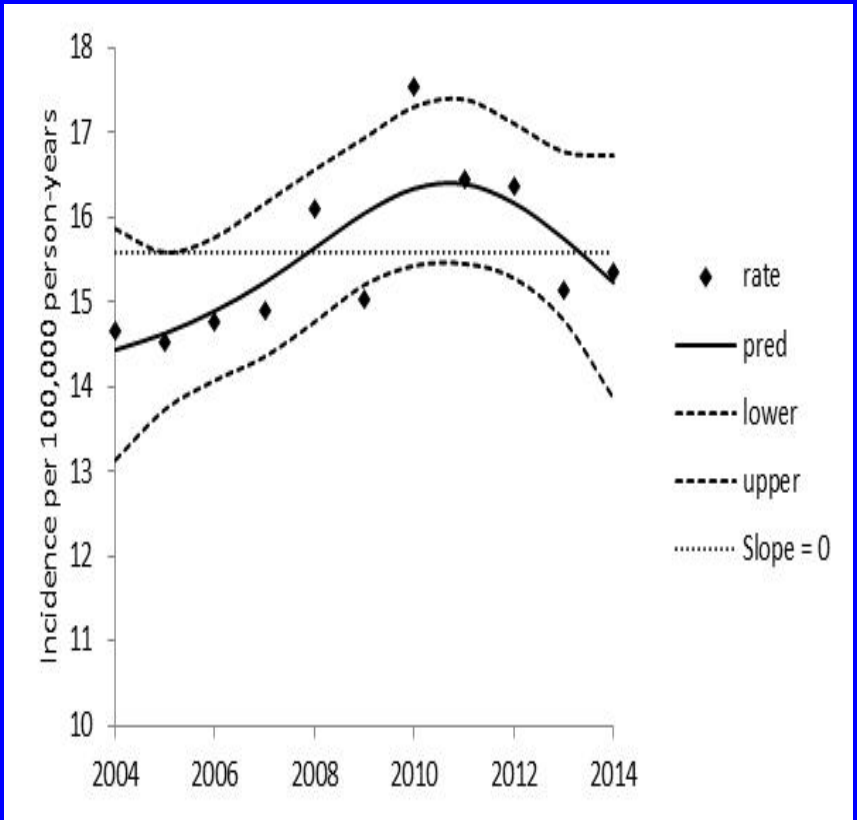
SALMONELLA 2004-13	Generalized Poisson				Negative Binomial			
Parameter	Est	SE	ChiSq	p	Est	SE	ChiSq	p
Intercept	-33.7745	5.8629	33.1853	0.0000	-18.8233	6.0539	9.6677	0.0019
Year	0.0124	0.0029	18.0847	0.0000	0.0050	0.0030	2.7168	0.0993
CA_1996	0.1392	0.0496	7.8812	0.0050	0.1419	0.0532	7.1116	0.0077
CA_2000	-0.0916	0.0695	1.7378	0.1874	-0.0883	0.0575	2.3628	0.1243
CO_2001	-0.2408	0.0547	19.3724	0.0000	-0.2319	0.0542	18.2878	0.0000
CO_2002	-0.1050	0.1111	0.8930	0.3447	-0.0968	0.0693	1.9552	0.1620
CT_1996	-0.1820	0.0594	9.3717	0.0022	-0.1791	0.0552	10.5310	0.0012
CT_1997	0.1075	0.0676	2.5275	0.1119	0.1120	0.0570	3.8546	0.0496
CT_1998	-0.1694	0.0759	4.9751	0.0257	-0.1659	0.0591	7.8848	0.0050
GA_1996	-0.0512	0.0463	1.2218	0.2690	-0.0478	0.0527	0.8246	0.3638
GA_1997	0.1288	0.0581	4.9172	0.0266	0.1302	0.0549	5.6158	0.0178
GA_1999	0.8754	0.0381	527.7696	0.0000	0.8751	0.0513	290.7412	0.0000
MD_1998	0.1236	0.0479	6.6546	0.0099	0.1256	0.0529	5.6364	0.0176
MD_2001	-0.2490	0.0599	17.2543	0.0000	-0.2461	0.0553	19.8120	0.0000
MD_2002	0.1953	0.0584	11.1779	0.0008	0.1992	0.0550	13.1283	0.0003
MN_1996	-0.0841	0.0430	3.8175	0.0507	-0.0819	0.0521	2.4736	0.1158
NM_2004	0.1219	0.0516	5.5746	0.0182	0.1236	0.0536	5.3109	0.0212
NY_1998	-0.1743	0.0692	6.3411	0.0118	-0.1710	0.0574	8.8691	0.0029
NY_1999	-0.2673	0.0747	12.8201	0.0003	-0.2649	0.0587	20.3443	0.0000
NY_2002	-0.4437	0.0758	34.2684	0.0000	-0.4413	0.0590	55.8896	0.0000
NY_2003	-0.1778	0.0878	4.0961	0.0430	-0.1750	0.0622	7.9029	0.0049
NY_2004	-0.2046	0.1168	3.0710	0.0797	-0.2023	0.0710	8.1124	0.0044
OR_1996	-0.3510	0.0495	50.3779	0.0000	-0.3465	0.0532	42.3902	0.0000
TN_2000	-0.0054	0.0474	0.0130	0.9091	-0.0004	0.0528	0.0001	0.9934
TN_2003	0.0000	0.0000			0.0000	0.0000		
Scale/Dispersion	2.2475	0.0000			0.0117	0.0016		
SSPR	208.2813				241.9529			

Penalized B-Spline Regression Salmonella 2004-14*

Generalized Poisson



Negative Binomial



*MMWR 2015 - preliminary 2014

SUMMARY
Trends for Bacterial Pathogens

- Early decline followed by year-to-year variability about lower level
  - *Campylobacter*, *Listeria*, STEC O157, *Yersinia*
- Continuous decline
  - *Shigella*
- Continuous increase
  - *Vibrio*
- Inconclusive whether increase or no trend
  - *Salmonella*

Further Investigation: Trends in Salmonella Serotypes

- Long term trends (1996-2013)
- Seasonal variation (month-to-month)
- $\text{Log}(E[y_{ij}]) = \log(\text{population}_{ij})+\beta_0+f_1(\text{year}_i)+f_2(\text{month}_j)+\varepsilon_{ij}$

Limitations

- Results are preliminary, work in progress
- Reported illness is a proxy, not true incidence
- Not all FoodNet reported illness is foodborne
- Assumes data missing at random
- Descriptive model, not infer causes
- Uncertainty about generalizing from FoodNet population to national level not quantified

Acknowledgements

- Data Provided by Foodborne Diseases Active Surveillance Network, CDC
  - Olga Henao, Elaine Scallan et al.
  - Stacy Crim, CDC
  - Mike Hoekstra, CDC
  - Mike Williams, FSIS

Disclaimers

- The opinions expressed herein are the views of the author and do not necessarily reflect the official policy or position of the United States Department of Agriculture. Reference herein to any specific commercial products, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government.
- The findings and conclusions in this report are those of the author(s) and do not necessarily represent the official position of the Centers for Disease Control and Prevention.