

$$\sum_{y=0}^n \binom{n}{y} p^y (1-p)^{n-y}$$



Control Groups in the Modern Age

Josh Rounds MPH | Epidemiologist Senior

Binomial Probability Calculations (Bernoulli Trials)

- **Population based case-control studies have become more difficult and resource intensive to conduct**
- **Binomial probability calculations can be used to quickly assess potential exposures during cluster and outbreak investigations**

Binomial Probability Calculations (Bernoulli Trials)

- **Data needed to calculate binomial probability**
 - **Total number of cases in the cluster**
 - **Number of cases in the cluster who ate the food item of interest**
 - **The background consumption rate of that food item**
 - **Rarely do perfect background food consumption data (i.e., specific, current, and geographically representative) exist for a given food item**
 - **Usually, have to use various less than perfect sources and recognize the limitations**

Sources of Background Exposure Estimates

- **FoodNet Population Survey Atlas of Exposures, 2006-2007**
- **Aggregated Case Surveillance Data**
 - **Use your own sporadic case data**
 - **Center of Excellence Project Mercury aggregating data from multiple states**
 - **Minnesota data are posted online**

Food Item	Denominator	Overall %	Overall % (including maybe)	Female	Male	Age: <18	Age: 18 – 65	Age: >65	Spring	Summer	Fall	Winter	Pop Survey MN	Pop Survey National
Eggs	409	65%	69%	68%	61%	62%	69%	65%	67%	67%	61%	60%	77%	75%
Shredded cheese	388	48%	57%	50%	47%	53%	49%	18%	56%	49%	45%	40%	N/A	N/A
Processed cheese slices	417	41%	47%	40%	42%	43%	40%	30%	37%	47%	37%	21%	57%	51%
Block cheese	438	22%	26%	22%	22%	22%	21%	21%	23%	20%	20%	37%	62%	61%
String cheese	436	19%	23%	23%	14%	26%	11%	5%	24%	18%	17%	16%	21%	16%
Cheese curds	457	6%	7%	6%	7%	6%	7%	3%	2%	7%	6%	6%	N/A	N/A
Queso fresco or other Mexican style cheese	457	2%	3%	2%	3%	3%	1%	3%	3%	3%	1%	0%	4%	6%
Gourmet cheese (gouda, blue, other cow, goat or sheep cheeses)*	265	5%	7%	4%	7%	3%	8%	9%	8%	4%	7%	0%	N/A‡	N/A‡
Ice cream	415	47%	52%	46%	48%	54%	36%	41%	33%	52%	46%	32%	63%	59%
Frozen dessert treats	440	17%	20%	18%	15%	24%	7%	8%	8%	22%	12%	7%	N/A	NA
Yogurt	433	47%	50%	56%	37%	52%	40%	44%	48%	48%	43%	52%	45%	43%
Milk	461	89%	89%	87%	90%	95%	80%	82%	87%	90%	89%	79%	85%	79%
Unpasteurized milk	467	2%	3%	2%	3%	2%	2%	2%	2%	2%	3%	3%	2%	3%
Milk alternatives (soy, almond, rice milk)*	266	7%	7%	10%	4%	6%	8%	8%	8%	7%	4%	13%	N/A	N/A
Other dairy (cottage cheese, cream cheese, sour cream)*	249	31%	37%	37%	26%	30%	34%	30%	35%	32%	26%	43%	N/A‡	N/A‡
Buttermilk**	186	1%	2%	1%	0%	0%	0%	8%	0%	0%	0%	6%	5%	6%
Sour cream**	175	15%	22%	16%	14%	16%	15%	15%	10%	14%	18%	24%	37%	32%
Cream cheese**	174	9%	16%	13%	5%	9%	9%	0%	0%	10%	2%	20%	25%	25%
Cottage cheese**	180	3%	8%	4%	2%	4%	2%	8%	0%	4%	2%	6%	25%	22%

‡ Please refer to pages 13-15 of the FoodNet Population Survey Atlas of Exposures, 2006-2007. Question asked by population survey may not have been compatible to MN hypothesis generating form or population survey may include additional details.

Sources of Background Exposure Estimates: Restaurant Outbreaks

- **Restaurant Sales Records (Online Orders or Transaction Records)**
 - Only use a background estimate compiled from restaurant data to assess exposures within a restaurant or restaurant chain
 - The FoodNet Exposure Atlas would not be an appropriate data source in this instance

Considerations for Choosing Background Food Consumption Rate Estimates

- **Recent data are preferred, as food consumption patterns can change over time**
- **Characteristics of the cluster cases**
 - **If cases have distinct gender, age, or racial/ethnic characteristics, background estimates should be derived from similar populations (like matching in traditional case-control studies)**
 - **Regional variability - food consumption patterns can vary across states/regions**
- **Seasonal variation in consumption of particular food items**

Considerations for Choosing Background Food Consumption Rate Estimates

- **Use of sporadic case data for background rates could bias towards the null**
 - **Observed consumption rates for known vehicles (e.g., ground beef for *E. coli* O157) may be higher than in the general population**
 - **This could limit your ability to find a statistically significant association even if that food item is the source of the outbreak**

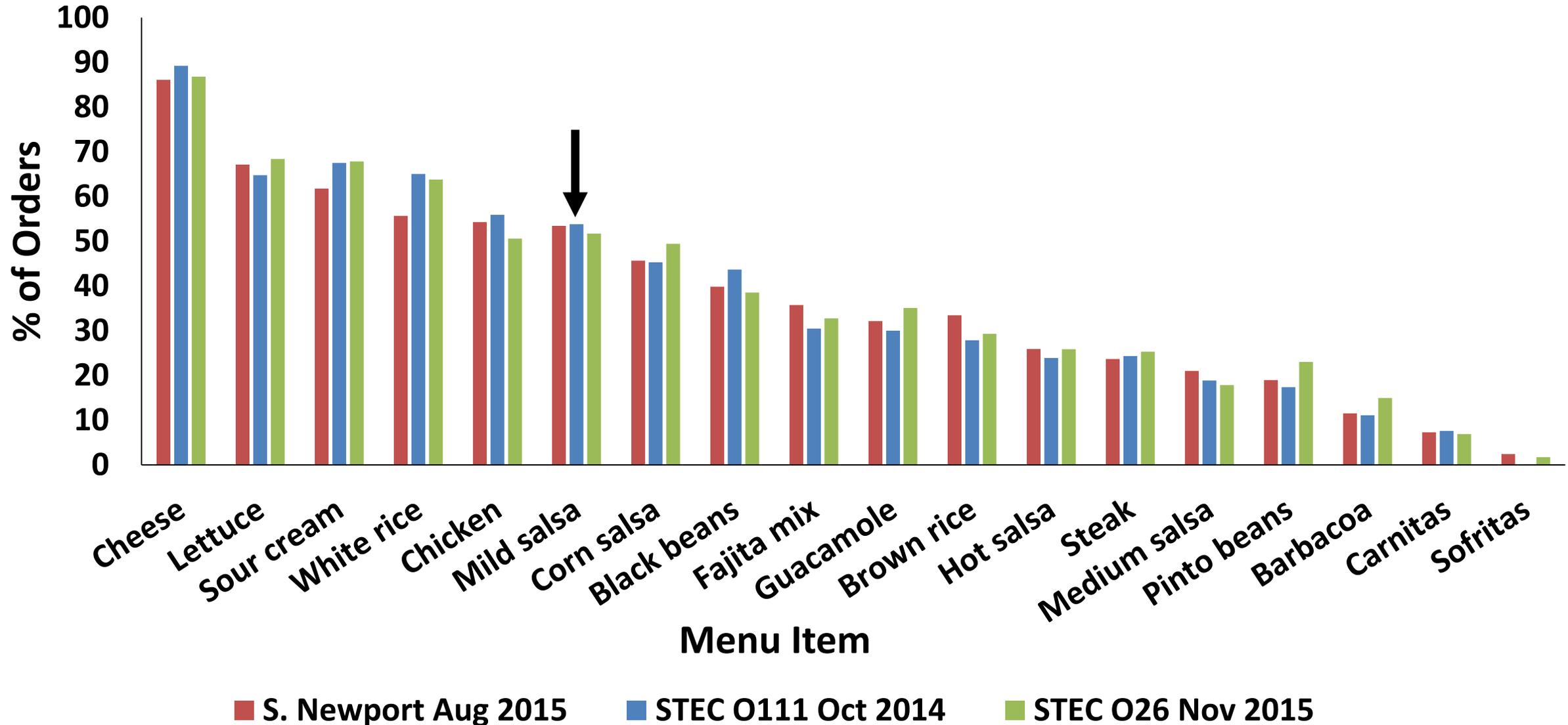
Considerations for Choosing Background Food Consumption Rate Estimates

- **If no data are available for the specific food item of interest, you can use data for a similar food item, or make an educated guess (it is best to be conservative and err on the high side)**
 - **E.g., during an *E. coli* O157 hazelnut investigation we used cashews as a proxy**
- **Can also plug in a variety of background estimates and see how they might change (or not change) your conclusion**
 - **E.g., how high does the background estimate need to be to not find a significant association? Is this plausible?**

Interpretation and Next Steps

- **Interpretation**
 - **By chance alone, how likely are we to find x of n people (or more) eating a given food item**
 - **Is this enough to conclude that the food item is the cause of the outbreak?**
 - **Not usually, but what it does indicate is that efforts to confirm the hypothesis (e.g., informational trace backs, food testing) are reasonable to pursue**
- **Example: *Salmonella* outbreak at a restaurant, 12 of 15 cases report consuming mild salsa**

Frequency of Menu Items Ordered from Burrito Chain C Online Orders During 3 Outbreaks, Minnesota (n=4,554)



Example

The screenshot shows the Epi Info 7 interface with the StatCalc window open. The StatCalc window is titled "StatCalc - Binomial" and displays the following data:

Binomial - Proportion vs. Standard	
Numerator:	12
Total observations:	15
Expected percentage:	53 %
Probability that the number of case	
< 12	0.9697081
<= 12	0.9929149
= 12	0.0232068
>= 12	0.0302919
> 12	0.0070851
Two-tailed p-value	0.0605838
95% confidence interval	8 - 14

- **Next steps**

- Environmental health assessment
- Ingredient specific case-control study
 - Ensure restaurant uses same ingredient in each dish, e.g., the same tomatoes are used in each dish
- Traceback investigation

Online Orders

- **Great source of data for restaurant outbreak investigations**
 - Both directly through restaurant and services (GrubHub)
 - Often include exposure data and contact information
 - Only need illness history
- **Consider potential biases**
 - Are people who order online different from other customers?
 - Less likely to order drinks
- **Keep doing case-control studies!**

Transaction Records

- **Traditional sources of controls may be limited in certain restaurant outbreak investigations**
 - Credit card receipts may not contain a patron name and are often only required for orders over a certain dollar amount
 - Often only a limited number of well meal companions
- **Itemized transaction records can be used as controls or to estimate the background consumption rate of foods in the restaurant**
 - Misclassification of some cases as controls
 - Each record used as a single control even though some transactions represent meals for multiple people, this may vary by restaurant
 - Bias towards the null

Transaction Records

Meal date and time →

Meal items →

```
*****  
* Thu Mar 12 12:36:31 2015 *  
* [redacted], Shannon M Reg#:2 Rcpt#:1915 *  
* *  
* 1 3 Pc Ml Or D $7.99 *  
* 1 Leg $0.00 *  
* 2 Thigh $0.00 *  
* 1 Msh/Grvy Sm $0.00 *  
* 1 Mac & Chez S $0.00 *  
* 1 Biscuit Sm I $0.00 *  
* 1 3 Pc Ml Or D $7.99 *  
* 1 Leg $0.00 *  
* 2 Thigh $0.00 *  
* 2 Coleslaw Sm $0.00 *  
* 1 Biscuit Sm I $0.00 *  
* *  
* Subtotal $15.98 *  
* TAX $1.14 *  
* Gross Receipt $17.12 *  
* Cash $0.00 *  
* Credit Card $17.12 *  
* Change Due $0.00 *
```

Obtaining Grocery Transaction Records from Credit/Debit Cards

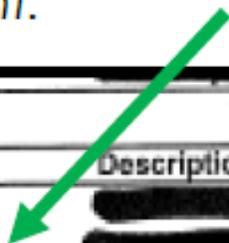
- **Cases can visit the store location(s) where they purchased food eaten before they became ill**
 - **Need to bring the credit/debit card that they used and visit customer service counter (or use in-store kiosk) to request copies of receipts for purchases made during the time period of interest**
- **Public health can also work with the grocery store to find itemized transaction information using details provided by the case**
- **Need to speak with the store to determine what information they need to find the case's receipts in their digital transaction files, which can vary by store**

Obtaining Grocery Transaction Records from Credit/Debit Cards

- Have the case look at paper or online credit card/bank statements to provide the needed information which may include:
 - Location of purchase, date of purchase, transaction number, time of purchase, and total dollar amount of transaction
 - Store may require written consent from the case to access these records

Example credit card/bank statement:

Transaction number



Transactions					
Trans Date	Post Date	Card Reference Number	Description		Amount
[REDACTED]	[REDACTED]	[REDACTED]	[REDACTED]		\$100.00
[REDACTED]	[REDACTED]	[REDACTED]	[REDACTED]		\$75.50
03/05	[REDACTED]	6542950EHRP1N5BY9M	[REDACTED]		\$8.60
03/11	[REDACTED]	7541823EN0618ZAJXM	[REDACTED]		\$24.00
03/12	[REDACTED]	6541734ERTAXGHAVYM	[REDACTED]		\$29.00

Resources

- **Centers of Excellence (coefoodsafetytools.org)**

- **Key Points**

- **Outbreak Investigation Case Studies**

- **Hazelnuts – Binomial**

- ***Salmonella* KFC – Transaction Records**

- **Contact us!**

- **Informal consultations**

- **For Center contact info, visit: www.cdc.gov/foodsafety/centers/sites**



Thank you! Questions?

Joshua.rounds@state.mn.us

651-201-5083