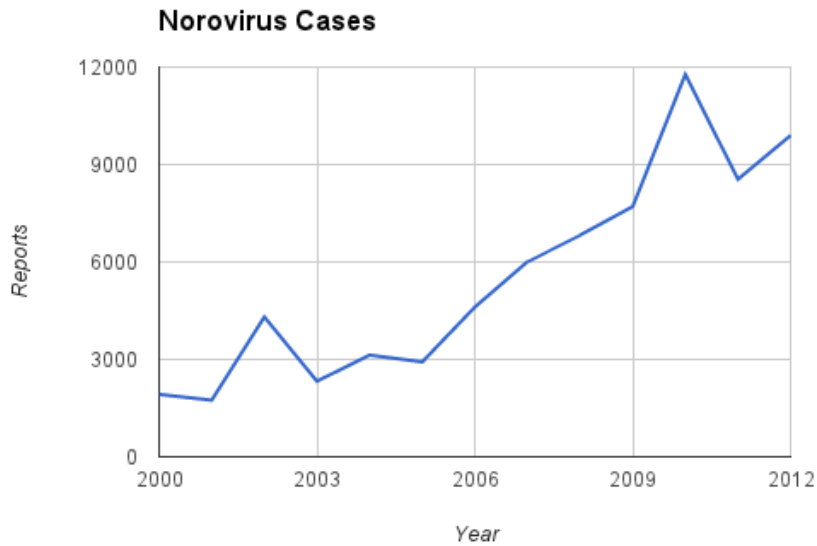


Lies, Damned Lies and Food Safety

Dan Strongin



The previous posts on this topic, edited and integrated into one.

A handful of epidemiologists applied probability theory to a hodgepodge of data to come up with a number of cases of foodborne illness annually in the United States. The number they come up with is over 9 million illnesses and 2,000 dead a year. Scary Stuff!

This posting on the Center for Disease Control website* became the basis for a report from *UNITED STATES PUBLIC INTEREST RESEARCH GROUP* that was widely distributed in the Media, named **“Total Food Recall: Unsafe Foods Putting American Lives at Risk,”** which paints a somewhat doomsday scenario, concluding;
“More needs to be done to protect Americans from the risk of unsafe food.”

Of course it does, but the point should be, what would be effective. Just doing things because you feel something needs to be done accomplishes nothing, and can make things worse. The article asserts:

“Important rules, standards, and inspections that could significantly improve food safety have been blocked, under-funded, or delayed, allowing the drumbeat of recalls to continue.”

Is the answer to the world’s problems more inspection, more regulation, more targets and, oddly, more studies, or better process? Inspection comes after the fact. They go on...

“In other words, instead of things getting better, they appear to be getting worse,” “Our food safety practices are falling short.”

Note the word “appear” before you panic!

- 🕒 Are they getting worse or is reporting getting better?
- 🕒 What exactly is the data that their conclusion is based on?

I am not an industry defender, or a special interest, but someone who would like to see something done that makes a real difference: not just a knee jerk reaction to apparently alarming news.

When people reach unreasonable conclusions, it is not to say they are unreasonable people. They are not alone, a number of articles over the years on food product safety draw unreasoned conclusions, based on incomplete, and imperfectly interpreted data, IMHO.

The problem is a reliance on the tools of enumerative statistics, rather than analytic or pragmatic. Think big data and weather prediction. Lot of good that has done!

A typical example of enumerative statistics is a political opinion poll; all you do is count and report. Drawing conclusions from enumerative data can be risky because it does not provide much in the way of context.

Analytic statistics is what Nate Silver used to so accurately predict the outcome of the last election, and is all about understanding data in context.

Pragmatic statistics is what is used in process improvement, taking the theories generated by analytical statistics, and testing them, where opinion rubs up against reality.

Statistical data, studied in the abstract, is less science than abstract math. Logical connections can be found in the abstract that simply do not exist in the real world.

When looking at data presented in tables, in muddled sequence, arranged without real affinity, mixing areas of opportunity (a fancy way of saying comparing apples to oranges), it is easy to find things that don't exist. Finding sense in nonsense is all too

human. I watch people do it with their financials all the time, until they learn a new way of seeing data.

Let me say, before we begin, that no sickness is the right amount. But before we jump whole hog into a bunch of new regulations and standards, doesn't it make sense that we should take a look at what is really going on first? Find the real cause, the root cause, the one that will really make things better?

The article continues, with a seemingly reasonable statement,

“The prominence of dairy in the study model reflects a relatively high number of reported outbreaks associated with raw milk compared with the quantity of raw milk consumed and issues related to Campylobacter spp. infection; these factors likely resulted in an overestimation of illnesses attributed to dairy.”

While true, it is hardly the only, nor the most important influence on the “overestimation” of illnesses attributed to dairy. Far more than pasteurized or unpasteurized; whether the food was commercially produced or not is far more importance.**

Is it fair to lump the results from non-commercial and commercial sources together?

The other thing that kills understanding is relying on measures of comparison, rather than real numbers.

“An estimated 629 (43 percent) deaths each year were attributed to land animal, 363 to plant, and 94 to aquatic commodities... followed by dairy (10 percent) etc.”

Are they saying that every year 62.9 people die from dairy, or are they saying 10 percent of all illnesses? A cow is a land animal after all, as are goats, sheep, yaks and buffalo. Or am I supposed to figure out what 639 is 42% of and then multiply by 10% for myself?

Percentages are meaningless unless you know the base number, and worse than meaningless if they are based on different numbers. As an example, a dollar compared to 10 dollars is 10 percent, to 100 dollars, only 1 percent, but it is still a dollar!

How many deaths from dairy were there? The actual database is not that clear.

Averages presented as if they are real things,

is a serious problem, especially when there are wide swings in variation point to point. If you average your pay and Bill Gates', does that give you any idea about his salary or yours? Hardly! Averages are comparisons, not real things. When articles throw a lot of numbers at you, without the hard data to back it up, or the context it sits in, how can you judge? Quantity is not quality.

And then comes this nonsense, parading as science:

“One surprising fact consumers should take away from the CDC study of food borne illnesses between 1998 and 2008 is that dairy products, including milk, cheese, and ice cream, are big contributors to food borne illness,”

Caroline Smith DeWaal, food safety director at the Center for Science in the Public Interest (CSPI).

“Dairy products ranked as the leading cause of hospitalizations linked to food borne illness; second to leafy greens in the numbers of illnesses; and second to poultry in the numbers of deaths,”

OH MY GOD!!! They are Scientists! We have to DO something!!!! or do we? Data in tables and single datapoints give little to no context. Without an understanding, from the ground up, in the industry itself, someone from an interest group can end up making what sounds like sense, but isn't. The desire for hot button issues for publicity, and the lack of training in how to interpret data correctly, leads to conclusions that will make things worse, not better, and take the focus away from what could help.

“Therefore, the incidence of reported outbreaks involving non pasteurized dairy products was $\approx 150\times$ greater, per unit of dairy product consumed, than the incidence involving pasteurized products. If, as is probably more likely, $<1\%$ of dairy products are consumed non pasteurized, then the relative risk per unit of non pasteurized dairy product consumed would be even higher.”

Using other people's interpretations of data, **you get 150x more nonsense**. How do they define incidence: by outbreak? Number of people sickened? What is less than one percent? Is it .9 or .5 or .1 percent? With the population of the US hovering around 314 million people, that represents a spread of more than 2,800,000 people, out of which, how many consume dairy products?

It may seem like it, but I am not nitpicking here: Major changes are being called for, up to and including rules that could put an end to what real people depend on for their livelihood: aged raw milk cheese; and based on nothing more than a song and a throw of the dice

Shooting from the hip misses the target, and kills innocent bystanders.

Is a knee jerk reaction to less than stellar analysis based on imperfectly collected, organized, and presented data enough to put an end to one of the last bastions for small family farmers; while sending raw fluid milk production underground, where how it is processed will never be improved.

The facts, seen in context, do not bear out these conclusions. I know, because I went back to the source, the CDC food borne outbreak database. I will share what I found with you.

A report that was published In the CDC's Journal of Emerging Infectious Diseases, Volume 17, Number One, from January 2011, is quoted widely. Now published on the CDC website as official estimates, it has ignited a debate from which major policy decisions will be made.

The data does not support the conclusions:

- ⌚ what data they use compares apples to oranges,
- ⌚ and they make no attempt to separate the signals from the noise, basing their calculations on the assumption of a normal distribution (footnote ***), something that rarely exists in complex messy systems like the food chain,
- ⌚ and they make no attempt to verify.

Are these predictions based on nothing more than "Stuff?" I downloaded the actual data, the same source they used, and applied the skills I have been lucky enough to learn, and the results point to very different conclusions than those posted on the CDC website, with some promising potential solutions, if we keep in mind that they must be tested first. I was looking for pragmatic ways to have real impact on lowering the incidence of foodborne illness in dairy products. To my surprise I uncovered more, some insight into commonly shared beliefs surrounding the relative safety of dairy products, both pasteurized and raw.

To review, I downloaded all reported food outbreaks from 1998 to 2010 from <http://wwwn.cdc.gov/foodborneoutbreaks/Default.aspx>. I removed all with unconfirmed

causes, or where the food involved was unclear, (which whacked off about 60% of the data.)

I kept any that had “suspected” causes, as having been involved in a few incidents myself, I know how hard it is to fix the exact cause, and what suspected means: a pretty good idea.

I removed all that came from bacteria that are not associated with Dairy Foods, and that can only contaminate after the product leaves the processing plant, like staph. The poor quality of the data would make it difficult or impossible to determine if any single incident is signal or noise.

1998-2010 Confirmed or Suspected Dairy Related Food Outbreaks

Non-Commercial	Outbreaks	Ill	Hospital	Dead
Raw	129	2262	224	3
Unspecified	41	598	116	0
Homemade	20	154	44	0
Pasteurized	10	153	35	4
TOTALS	200	3167	423	7

Commercial	Outbreaks	Ill	Hospital	Dead
Unspecified	30	1184	72	0
Raw	2	244	12	0
Pasteurized	5	243	12	1
TOTALS	37	1671	96	1

Tables of data can be misleading; they don't provide enough context. It would be easy to assume from this that there is a huge difference in risk between pasteurized and raw milk, one of the beliefs commonly shared, but when seen in the context of commercial production, looks very different.

What you are looking at includes *Non-Commercial* product from private homes, farmhouses, church events, and picnics as well as commercially produced product. The “homemade” in fact is homemade ice cream. The available data is not clear if raw milk outbreaks came from farmers drinking their own milk, or from those who buy raw milk locally from the farmer. More research would need to be done.

The numbers for commercial products tell a different story, don't they? Is it possible that the raw milk controversy consuming the dairy industry and some legislatures may be a

classic red herring, taking attention and resources away from what could really make a difference?

Rather than trying to distance itself from raw milk products, the industry would be better served trying to ensure the industry as a whole is not compared to non-commercial products.

It is fundamentally unfair to include commercially produced with non-commercial product in the same analysis, and then draw conclusions affecting both.

The single death reported from confirmed commercial product over those 12 years occurred in a pasteurized cheese bought in a supermarket in Oregon in 2006, and was caused by listeria. Even among non-commercial deaths, most came from “bath-tub” producers of fresh Mexican style cheese, both raw and pasteurized.

I am not trying to minimize the tragedy of anyone dying, nor the inherent risks in raw milk, but trying to understand how to deal with what really happened and is likely to happen again. To prevent it, we need to find the real causes and stop blaming the easy targets.

Commercial Incidents

The largest number of outbreaks took place in Restaurants, and the largest number of illnesses per outbreak in Schools and Camps. This makes sense since the number of people served the same food in camp or school is greater, and kids immune systems may not yet be fully developed.

Where the Commercial Incidents Took Place

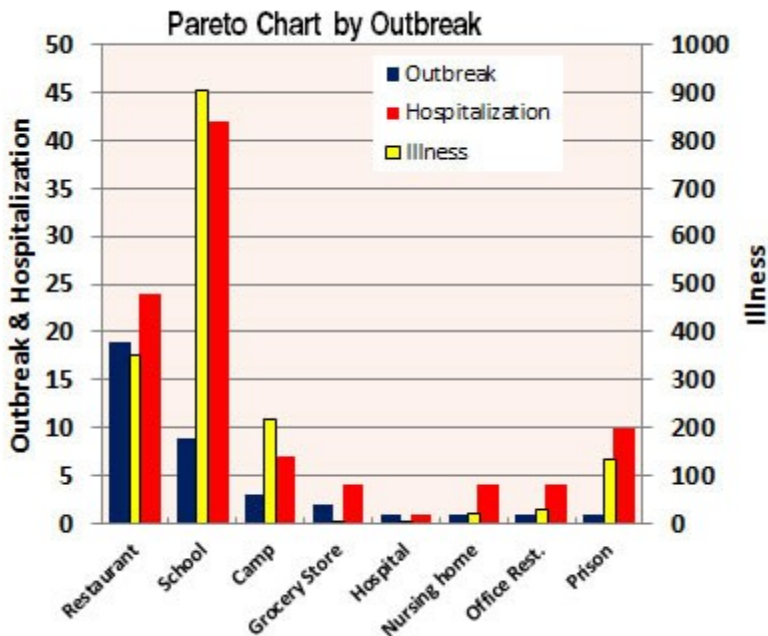
LOCATION	Outbreaks	Ill	Hospital	Dead
Restaurant	19	351	24	0
School	9	904	42	0
Camp	3	217	7	0
Grocery Store	2	6	4	1
Hospital	1	6	1	0
Nursing Home	1	21	4	0
Office Setting	1	31	4	0
Prison	1	135	10	0
Raw	2	244	12	0
Pasteurized	5	243	12	1

TOTALS 37 1671 96 1

Two-thirds of the 904 illnesses in the schools took place in only five incidents from 2001 to 2005. Three of those were caused by salmonella, with one from pasteurized 2 percent milk, and two others from foods made with cheese, which, based on experience, raises the question of *food handling*, rather than the innate integrity coming from the factory. In fact, with the exception of the two incidents from grocery stores, most of the other incidents involved post process food handling.

What Matters

A Pareto chart is a useful tool that helps us discern the relative few that really matter.



In the chart above we can see the data organized by most outbreaks to least. This tool helps us figure out where to begin to look to improve the right processes, the ones that will make the most difference. *Focusing on Restaurants and Schools would have a huge impact*, based on the data we have.

Looking at it from the point of view of illnesses caused, the first place to find a solution would be in schools, the second restaurants.

Schools would be easier as the number of illnesses per outbreak tend to be greater.

If solutions can be found, these two would eliminate over half the illnesses reported for Dairy Products, if the trends from the missing and incomplete data hold. (TESTING REQUIRED)

If the arrow clearly points to schools and restaurants as the first point of entry to analyze and find root causes to improve the system, what will more regulations and inspections for the manufacturer do to help?

By Germ

The following table includes both commercial and non-commercial sources in the more reliable data. While the most outbreaks were caused by Campylobacter, the most illnesses were caused by Salmonella, by a large margin.

GERM	Outbreaks	Ill	Hospital	Dead
Salmonella	84	2497	285	1
Campylobacter	119	1592	80	2
E-Coli	26	826	132	0
Listeria	7	68	43	6
Brucella	5	18	7	0
Other	3	14	1	0
TOTALS	244	5015	548	9

Putting Salmonella aside for a moment, one part of the story in the data becomes clear when you consider the link between Campylobacter and Raw Milk: only a tiny number of outbreaks from this germ are linked to anything else.

This is good news. If a way to mitigate campylobacter in raw milk can be found, and processing improved, a huge number of cases could be eliminated: gone, solved, no one sick.

With most illnesses coming from Enteric Salmonella a close study would have to be done where the incidents happen to find out how the contamination takes place. It will most certainly involve food handling, as most of the incidents involve secondary processing, meaning after the product leaves the manufacturer, of pasteurized dairy products.

But some reasonable assumptions can be made, based on an understanding of statistical thinking, and the intensity of some things, as you will see are clearly signals, based on knowledge of food production, handling and distribution.

A More Reasonable Approach

I will summarize now what I found in my research, and make what I think are reasonable and effective suggestions on which to focus resources to find real solutions. If more of us can learn about pragmatic statistics, we could accomplish so much more, with less.

A handful of epidemiologists applied probability theory to a hodgepodge of data to come up with a number of cases of foodborne illness annually in the United States. The number they come up with is over 9 million illnesses and 2,000 dead a year.

They provide no empirical evidence to back this rather drastic assertion. By empirical I mean verifiable by observation or experience rather than just theory or pure logic. Are we supposed to accept it at face value, and make major changes in public policy?

Sadly, that is what happens in society, and in businesses all the time. The boss says I want 10 percent more without ever understanding what would really be required to make that 10 percent; what Lee Iacocca once called the mushroom school of management: make absurd demands, keep them in the dark and scare them into growing!

Mental Gymnastics

Mental gymnastics done to try to help decide where to apply limited resources in order to proactively confront the problems the US “may” face in the future seem laudable, but the road to perdition is paved with good intentions.

The solutions people come up with, based on untested, unproven theories include, surprise, surprise: more inspection, placing a greater financial burden on industry to maintain arbitrary standards that most likely will not make a difference, based on the real evidence. These ill wrought solutions will take eyes off what really makes a difference, process improvement, leading to calls, within some government agencies, for banning whole classes of products, and sectors of the food and dairy industries.

But without understanding the real causes of the unwanted results, we risk a huge expenditure of already limited resources to accomplish little more than the destruction of one of the last great hopes for the survival of American Family Farming: aged raw milk cheese, among other products. It doesn't have to be this way. Valuable things can be gained from looking at data analytically, even when incomplete, then testing pragmatically.

Rather than make grand guesses, find the meaning buried within what reliable data there is. The probability calculations the study authors used were based on counts of what has happened, and guesses, with little to no context provided. To understand, to find concrete actions that can be taken to improve a situation, context needs to be provided. With context, the meaning buried in the data can be found. Some of what has happened in the past will be useful for prediction, and some just plain wacky. You first have to separate the signals from the noise. Otherwise all we end up doing is making logical connections that have nothing to do with reality: stuff and nonsense.

Pragmatism

Understanding generated from analysis can build a theory, but that theory must predict changes in the real world to be Science. Otherwise, at best, all you have is philosophy, at worst, fiction.

We know Nate Silver nailed it because he nailed it! Reality is still the only place to get a good steak.**

Though solutions being called for are logical: ban raw fluid milk, and some raw milk products and increase the time of aging for aged raw milk cheese from 60 to 90 days before it can be sold, but will that solve the problem? **Not based on the data.**

The data points to the need to separate non-commercial from commercial, take a closer look at post manufacturer food handling, particularly in restaurants, schools, and camps; and the dominance of campylobacter and salmonella as where to look to solve the vast majority of the things *that have actually happened*.

Neither of these have been a problem with aged raw milk cheese except where the evidence points to post manufacturer handling. Do we start a war on raw milk, or do we dig deeper, and solve the real problem, through understanding and better process?

The same or similar problems happen with post manufacturer food handling with pasteurized milk, and in fact, the only fully confirmed death from a commercially

approved dairy product was from listeria in a pasteurized cheese. Vague threats of potential under-reporting miss the point. What is, is what matters, not what “could” be, at least if you really want to solve a problem.

The National Institute of Health of England has taken a better path, a pragmatic one. They have called for the investment of resources in finding a rapid test for Campylobacter in animal and in the milk***

Resources squandered on more inspection and policing bans would be better spent on working with the English to develop this test. Any smart supplier of Raw Fluid Milk would want to put a label on their milk ensuring its safety through rigorous testing and continual process improvement. It would be a huge commercial advantage.

Another short term solution could be to limit sales of raw fluid milk to private homes, **not** institutions, so if there is a problem, the number of illnesses remains small, a matter of personal choice, personal liberty.

Real Safety

I feel confident in saying that Campylobacter, and E. coli could be minimized or eliminated from the milk and the herds through better process on the dairy farm.

In fact, it is the only way. Inspecting does nothing to eliminate the cause, as it comes after the fact, when it is too late, and standards forced on people without their involvement are not effective, and are not followed.

E. coli and Campylobacter get into milk from contact with feces from infected animals, who most likely get it from eating feed contaminated with the feces from other infected animals. Ensuring that feed is not cut too close to the ground, along with careful monitoring of the animals, good manufacturing and agricultural processes, and fore-stripping teats before milking has the potential to greatly diminish their impact, if not eliminate them. It is this kind of insight that comes from knowing the context of the data, that people outside of the food industry have no way of knowing, so they throw darts randomly.

Increasing the economic burden on producers based not on what really happens, but on imperfect probability calculations would devastate family farms for which raw milk cheese production has been a godsend.

Banning the sale of raw milk to those who have already chosen an alternative lifestyle, might simply force the industry underground, where illness would occur, but go undetected. Unhygienic conditions that allow pathogens to develop, may one day produce resistant strains like O157. The only solution is continually improving processes.

If the Food Industry, in these examples, the Dairy Industry wants to do something positive, something with vision, rather than merely point the finger at raw milk, or loopy tree huggers, it should invest in and lobby for real resources to develop a rapid test for Campylobacter, and an industry wide effort to continually improve milking parlors, holding tanks, and feed cutting practices.

While some of the milk used in commercial operations is listed as “unspecified,” it is reasonable to assume that the product was pasteurized during manufacture, given the type of products listed (see the database link in the previous article), and therefore, was contaminated after leaving the processing plant.

The industry should invest in, and support educating consumers about better food handling practices, and work with their foodservice customers to ensure safer food through better handling after manufacture, as many of the outbreaks reported involve post-plant secondary processing.

If I could, I would require the study of analytic, pragmatic statistics in Government, Private Business, and Business Schools, so we could start to make a real difference in how our world is really run. We may not get definitive answers, but we get a good hint of where to really look.

*<http://wwwnc.cdc.gov/eid/article/17/1/pdfs/p1-1101.pdf>

** Woody Allen

***www.food.gov.uk-multimedia-pdfs-campylobacterstrategy.pdf