Critical Outcomes Report Analysis

January 10, 2008

Agenda

1:00	Overview of *why* reports are wrong and how to fix them. This will help somewhat in reading them and in contracting for DM but critical outcomes report analysis is about learning how to read these things generally Sample question and answer
2:00	Test
3:00	Return tests and break
3:15	Going over the answers. Email lines will be open
3:45	Adjournment of formal session. I will be available until 5:00 to answer followup questions privately on phone or email

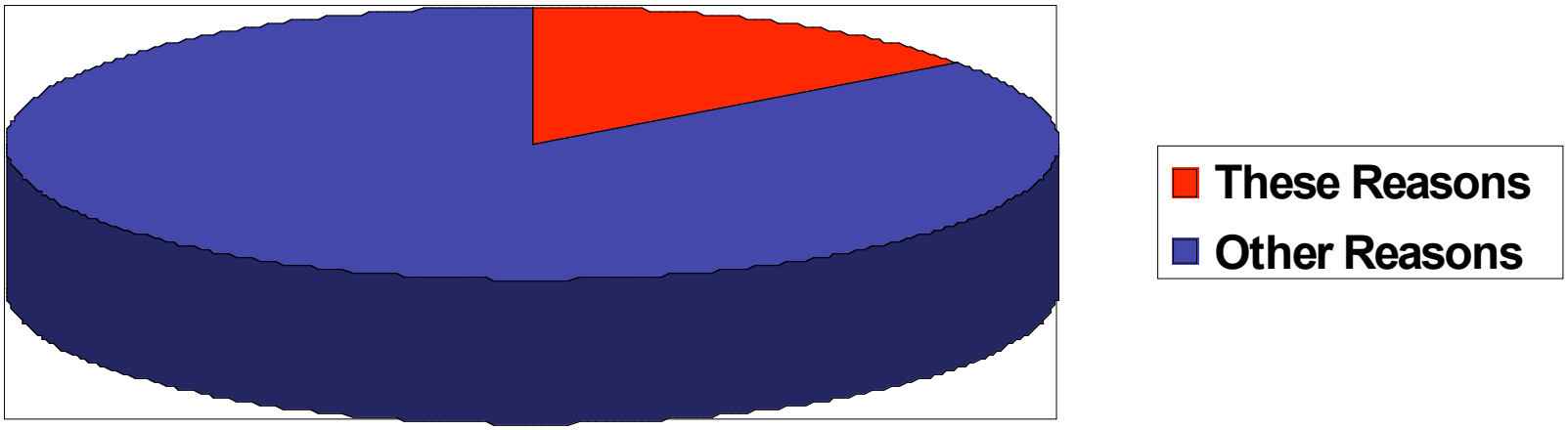
Overview of Why Reports are wrong and how to fix them and be a hero to your organization…

…Rather than rely on others for your measurement

Reasons Why Reporting is often Wrong

- Look at these “checks and balances,” and ask yourself, why aren’t you already doing this in contracts with your vendor?

Plenty of Other Reasons too (Read the DMAA guidelines)



Three reasons reports are wrong

1. No one does a Dummy Year Analysis
   The exact same methodology applied to a year in which you did not have disease management
2. No one checks for plausibility
3. No one says, “wait a second – this doesn’t make sense.” This is Critical Outcomes Report Analysis

Dummy Year Analysis

- Most contracts have a baseline period to which a contract period is compared (adjusted for trend)
- Watch what happens when you have a baseline and then compare a contract period (adjusted for trend)
  - Just the analysis, no program

In this Dummy Year Analysis example

- Assume that “trend” is already taken into account
- Focus on the baseline and contract period comparison

Base Case: Example from Asthma

First asthmatic has a $1000 IP claim in 2005

	2005 (baseline)	**2006 (contract)**
Asthmatic #1	1000	
Asthmatic #2		
Cost/asthmatic		

Example from Asthma

Second asthmatic has an IP claim in 2006 while first asthmatic goes on drugs (common post-event)

	2005 (baseline)	**2006 (contract)**
Asthmatic #1	1000	100
Asthmatic #2	**0**	**1000**
Cost/asthmatic		

What is the Cost/asthmatic In the baseline?

Cost/asthmatic in baseline?

	2005 (baseline)	**2006 (contract)**
Asthmatic #1	1000	100
Asthmatic #2	0	1000
Cost/asthmatic	**$1000**	

Vendors don't count #2 in 2005 bec. he can't be found

Cost/asthmatic in contract period?

	2005 (baseline)	**2006 (contract)**
Asthmatic #1	1000	**100**
Asthmatic #2	0	**1000**
Cost/asthmatic	**$1000**	**$550**

Base Case: How Dummy Year Analysis (DYA) fixes it

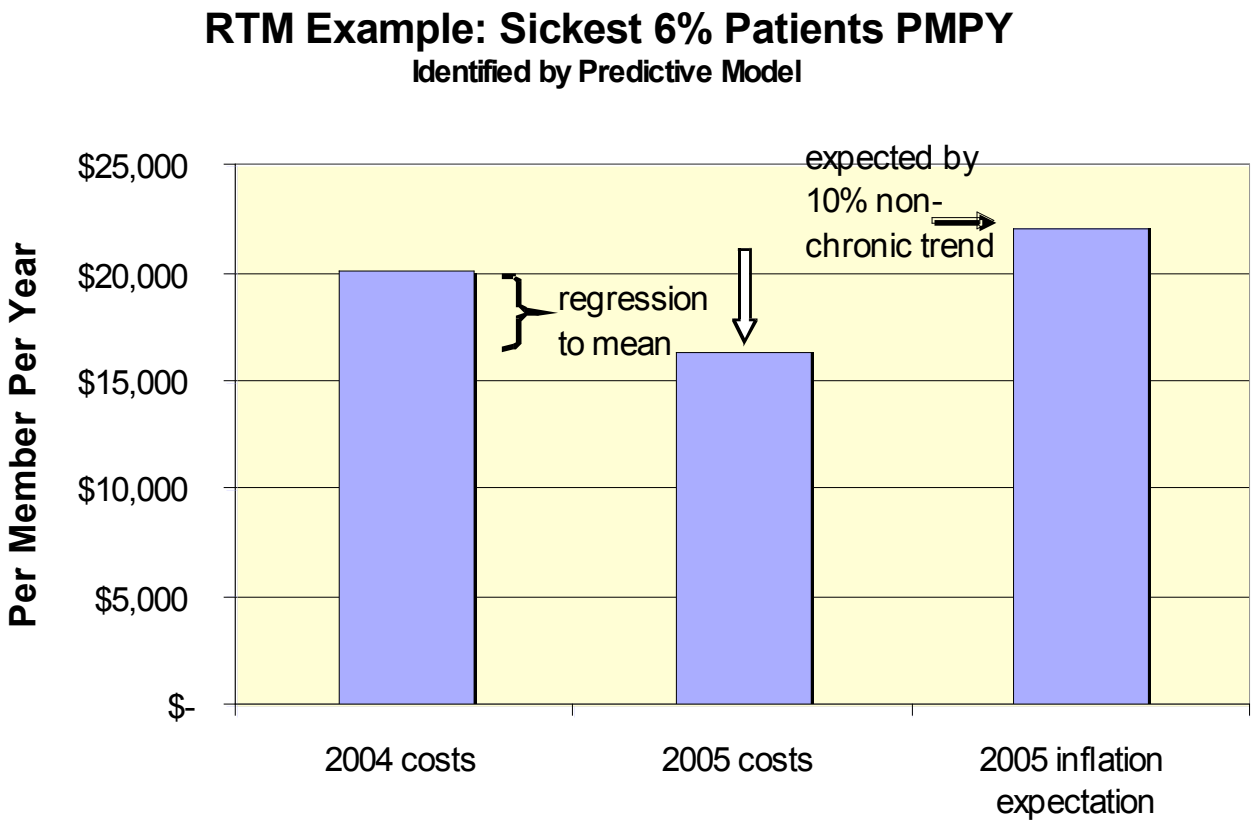
	2005 (baseline)	**2006 (contract)**
Asthmatic #1	1000	**100**
Asthmatic #2	0	**1000**
Cost/asthmatic	**$1000**	**$550**

In this case, a “dummy population” falls 45% on its own without DM

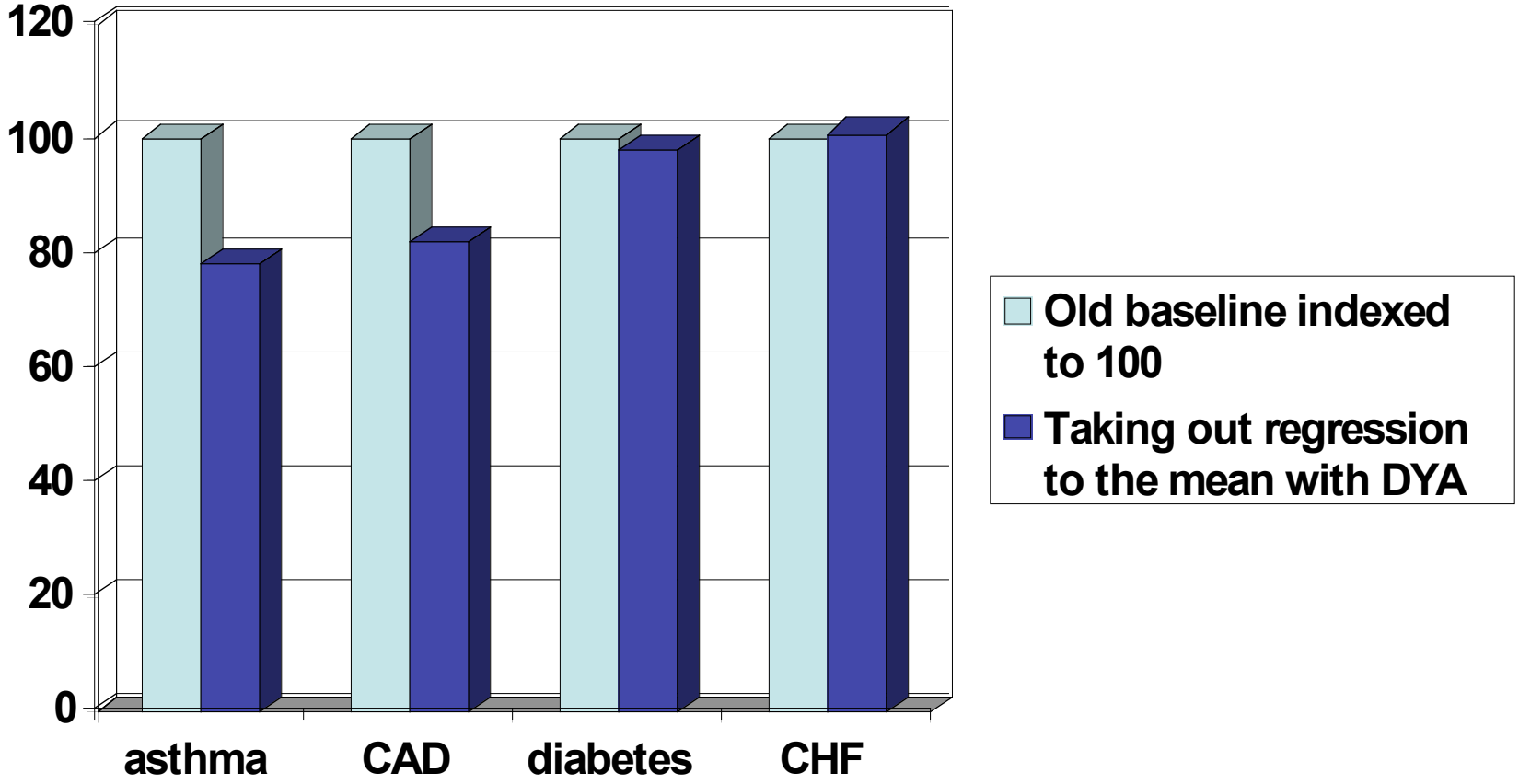
So…

- If you were to do an asthma program the vendor should not get credit for the reduction that happens anyway
  - But they do
  - How do we know that? With a plausibility test, to be discussed later
  - First, some real-world Dummy Year Analyses (DYAs)

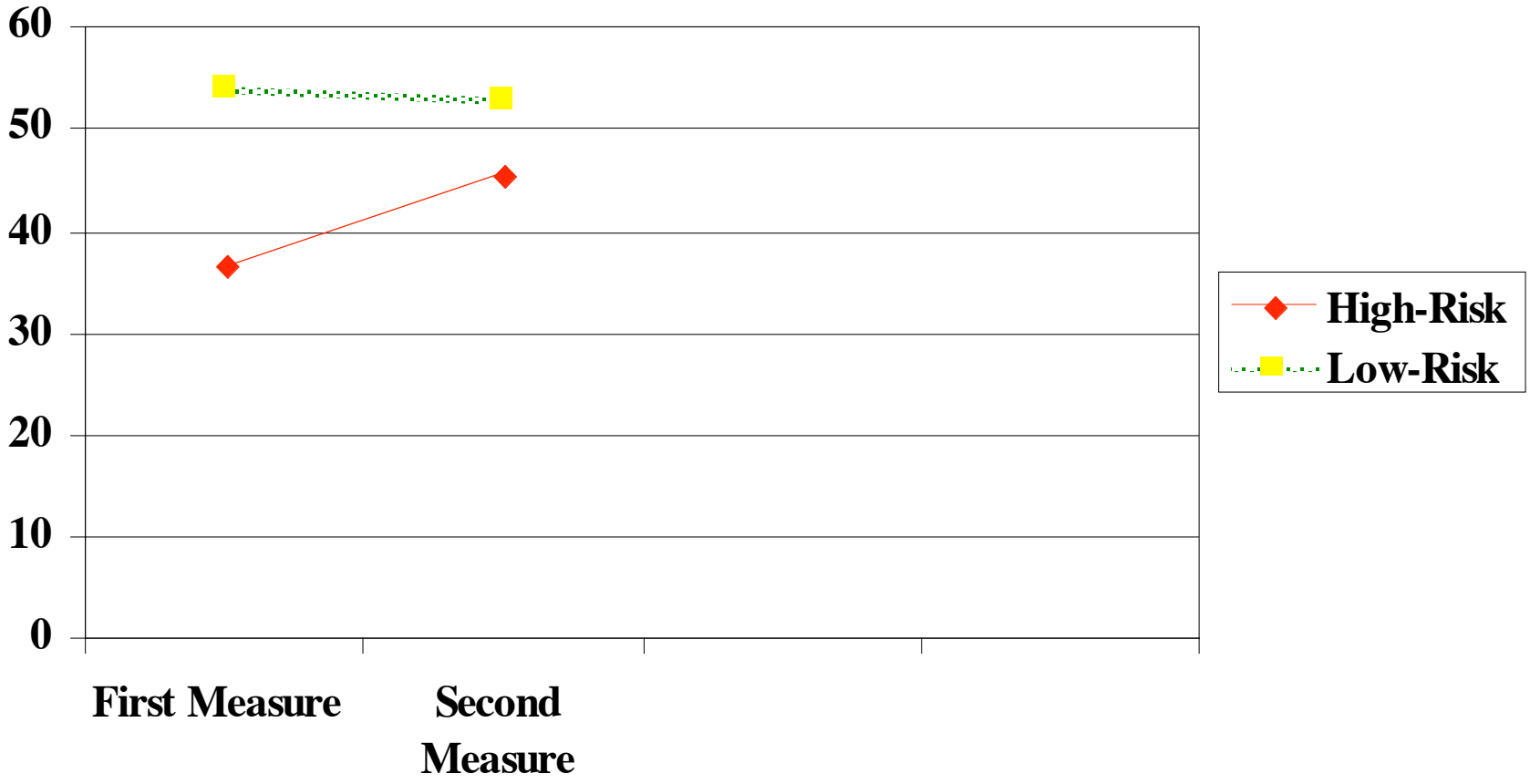
DYA real-world Result: Excerpt from Regence Blue Cross-DMPC study for *Health Affairs* released recently



DYA Result By Disease (using 1-year baseline and standard DMPC algorithms) – what is the difference which is caused automatically by just trending forward?

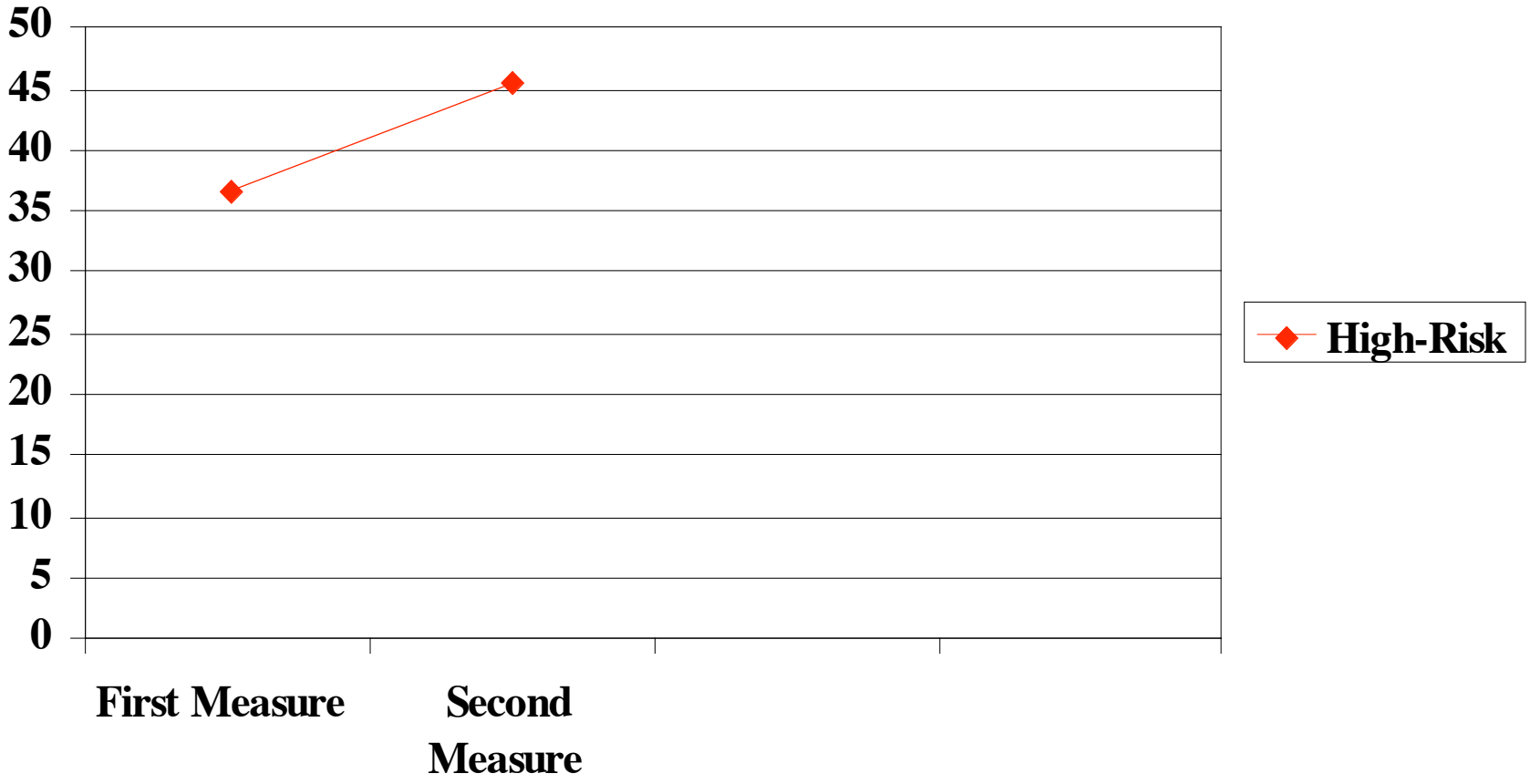


DYA Result in Wellness



Source: Ariel Linden – citation On request

T*here was no program in this case* – just two samplings and the average stayed the same



Source: Ariel Linden – citation on request

Other evidence for Dummy Year Analysis (DYA)

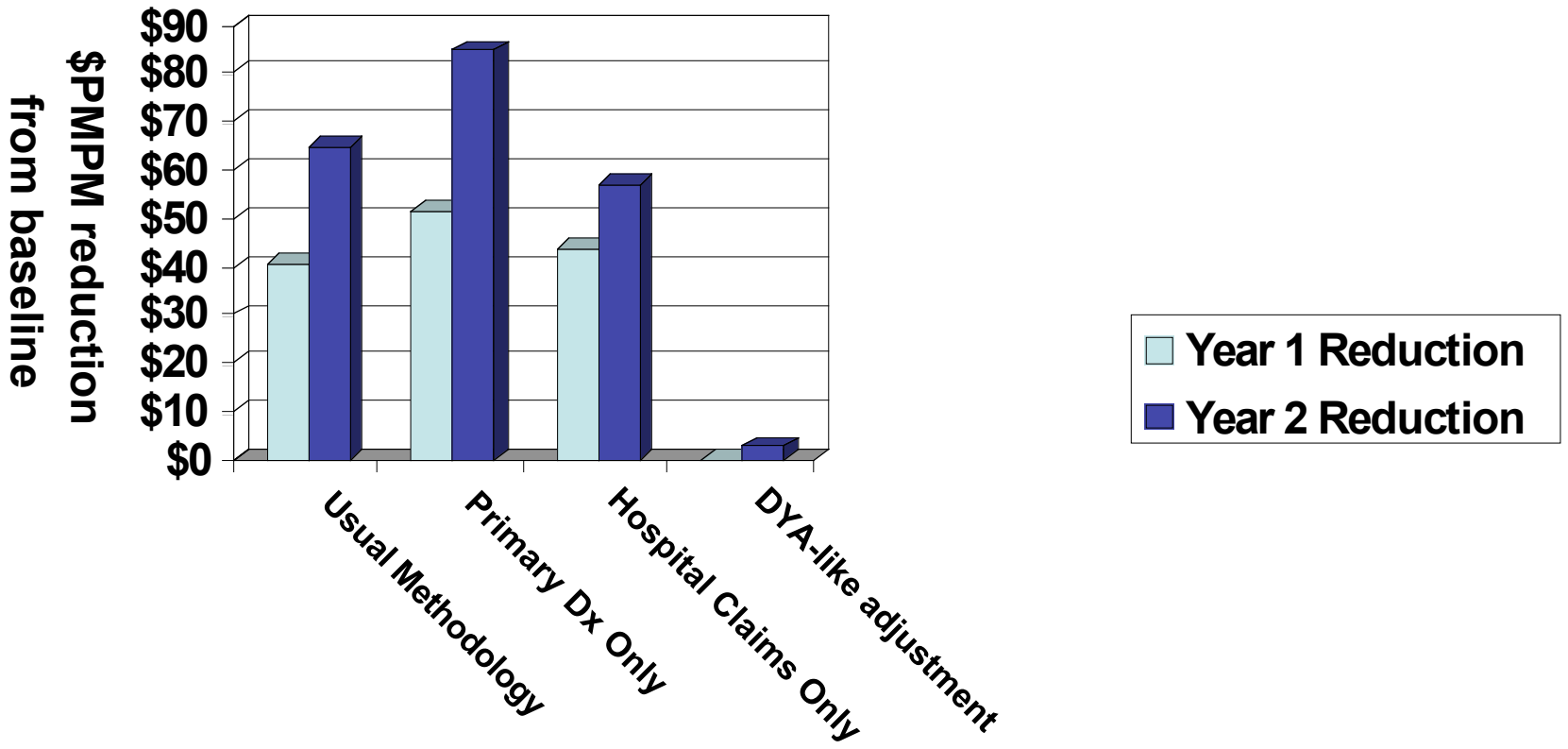
- CMS studies – very carefully designed -- get results opposite those done without DYAs, and consistent with those done with DYAs
- Only one vendor does a DYA-like adjustment
- Watch what happens when you get results "adjusted for trend" --
- ROIs without DYA adjustment flunk plausibility testing…

Actual Report example

Service category	Expected Cost (adjusted for trend)	Actual cost	Savings
Inpatient	$137	$125	$12
ER	$8.00	$7.50	$0.50
Outpatient	$62	$59	$3
Labs	$9.00	$8.80	$0.20
Office Visit	$69	$66	$3
Other	$125	$121	$4

Impact of adjustment similar to DYA on Highmark (Medicare)

Data courtesy of www.soluciaconsulting.com



Other evidence for Dummy Year Analysis (DYA)

- CMS studies – very carefully designed -- get results opposite those done without DYAs, and consistent with those done with DYAs
- Watch what happens when you get results "adjusted for trend" --
- Reports like that just scream out for plausibility testing…

Three reasons reports are wrong

1. No one does a Dummy Year Analysis

   The exact same methodology applied to a year in which you did not have disease management

2. No one checks for plausibility
3. No one says, “wait a second – this doesn’t make sense.” This is Critical Outcomes Report Analysis

What is a plausibility test?

- You do it all the time…outside DM
- An easy way to directionally check results
- Measure total event rates for diseases being managed, like you'd measure a birth rate. Couldn't be easier
- Specific codes on the next page
  - Specific fine-tuning rules available from me
- Example from previous asthma hypothetical

Event rates tracked by disease: the Plausibility Indicators

Disease Program Category	ICD9s (all .xx unless otherwise indicated)
Asthma	**493.xx (including 493.2x[1])**
Chronic Obstructive Pulmonary Disease	**491.1, 491.2, 491.8, 491.9,. 492, 494, 496, 506.4**
Coronary Artery Disease (and related heart-health issues)	**410, 411, 413, 414**
Diabetes	**250**
Heart Failure	**428, 404.01, 404.03, 404.11, 404.13, 404.91, 404.93, 425.0, 425.4**

[1] 493.2x is asthma with COPD. It could fit under either category but for simplicity we are keeping it with asthma

Cost/asthmatic in contract period?

	2005 (baseline)	**2006 (contract)**
Asthmatic #1	1000	**100**
Asthmatic #2	0	**1000**
Cost/asthmatic	**$1000**	**$550**

Asthma events in the payor as a whole – the plausibility check

	2005 (baseline)	**2006 (contract)**
Asthmatic #1	1000	**100**
Asthmatic #2	0	**1000**
Inpatient events/year	**1**	**1**

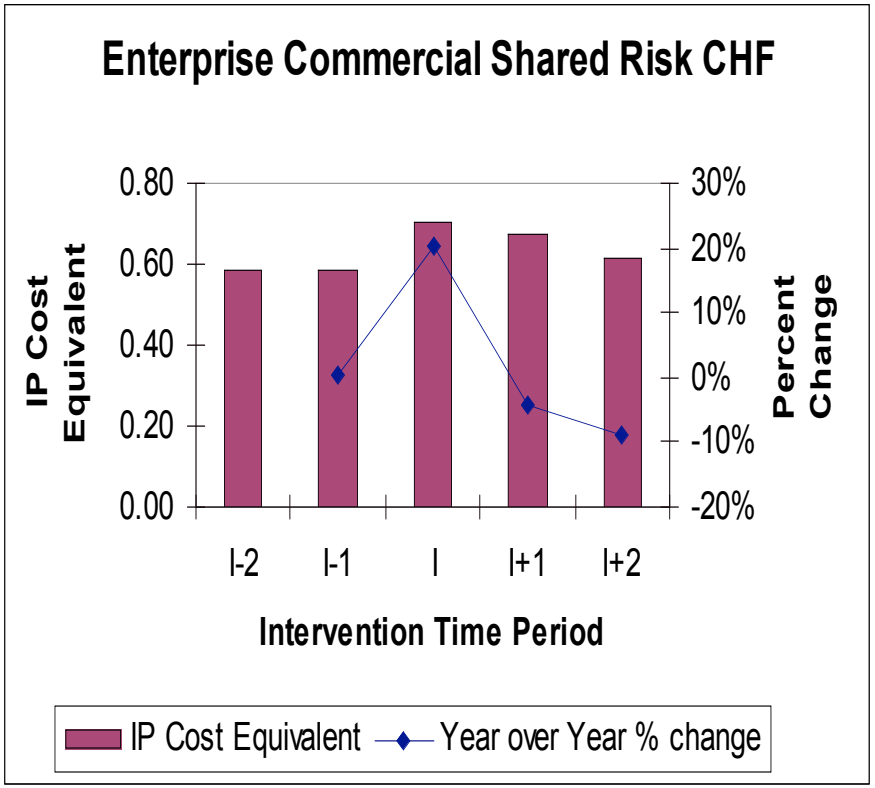
Plausible?

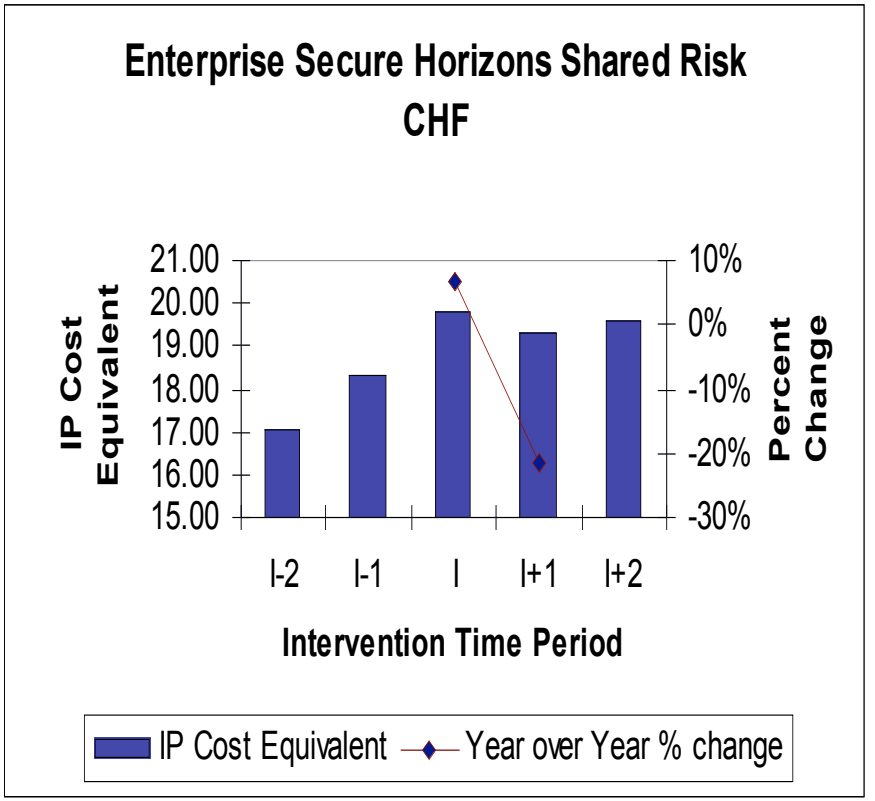
- How can you reduce asthma costs 45% without reducing planwide asthma event rate?
- Answer: You can’t. Not plausible

Several Examples of Plausibility Analysis

- Pacificare
- Some which didn’t turn out so well
- Plausibility-testing generally and benchmarks

PacifiCare HF Results

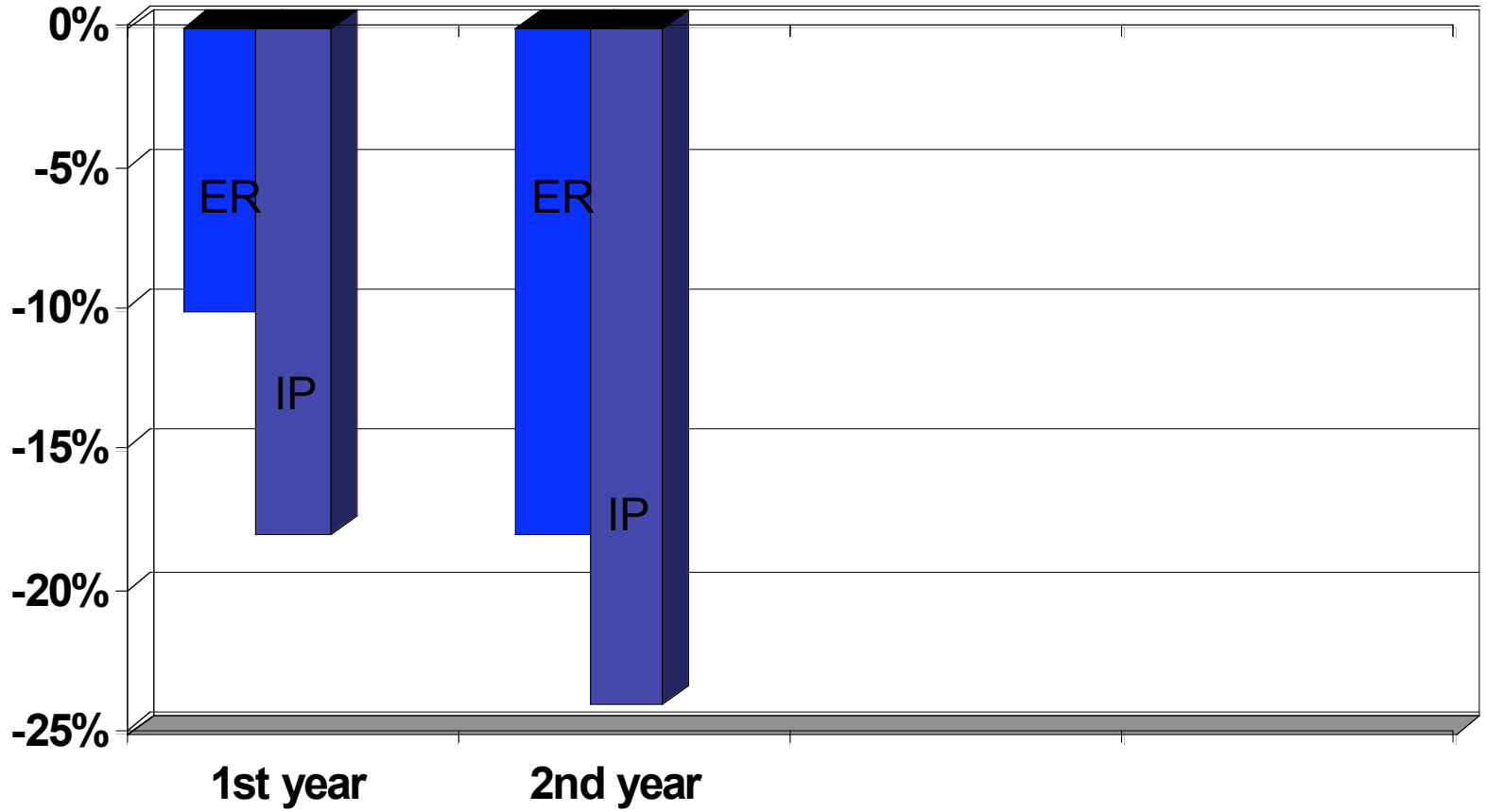




Several Examples of Plausibility Analysis

- Pacificare
- Some which didn’t turn out so well

Example of just looking at Diagnosed people: Vendor Claims for Asthma Cost/patient Reductions

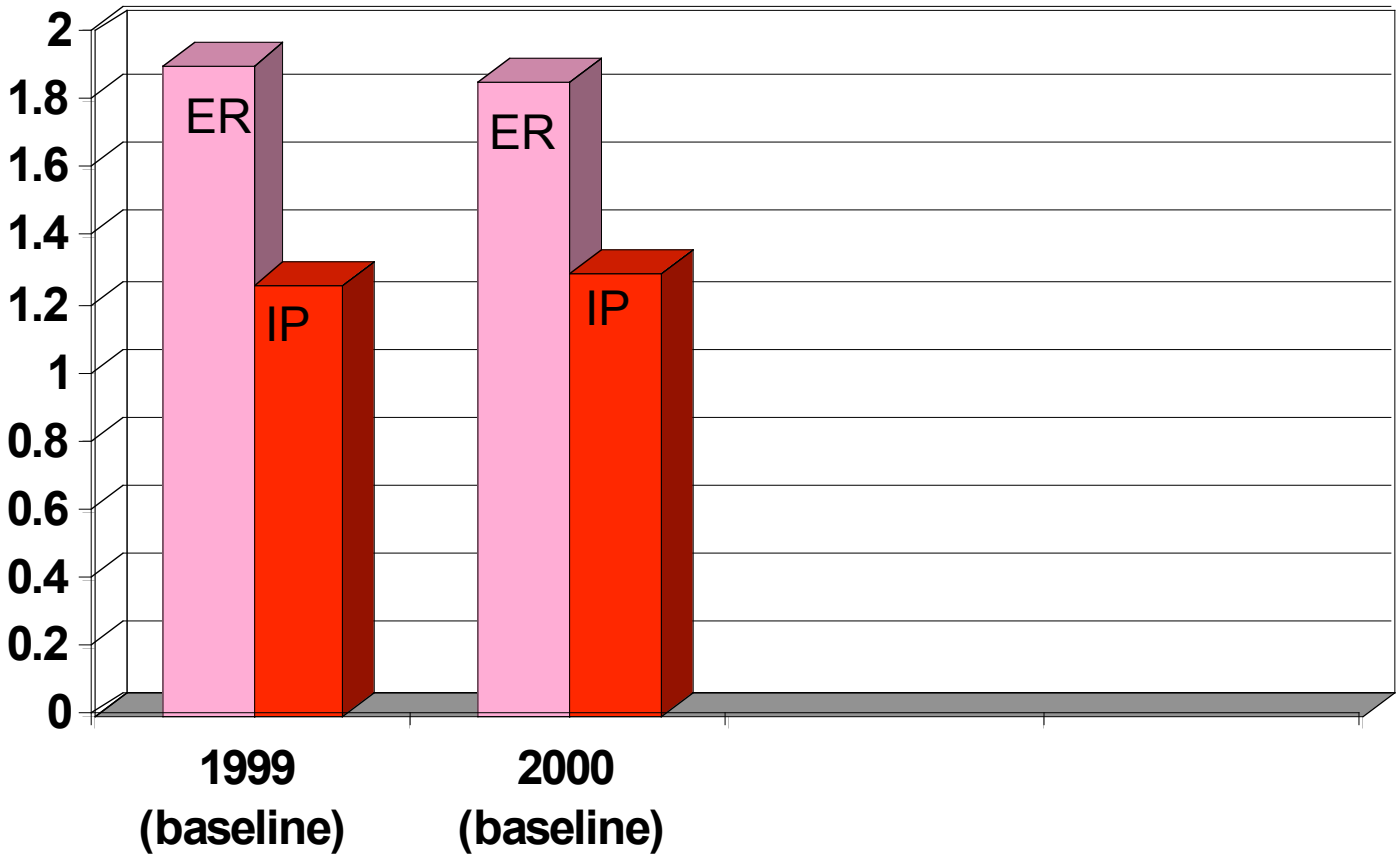


What we did to plausibility-test…

- We looked at the actual codes across the plan
- This includes everyone
- Two years of codes pre-program to establish trend
- Then two program years

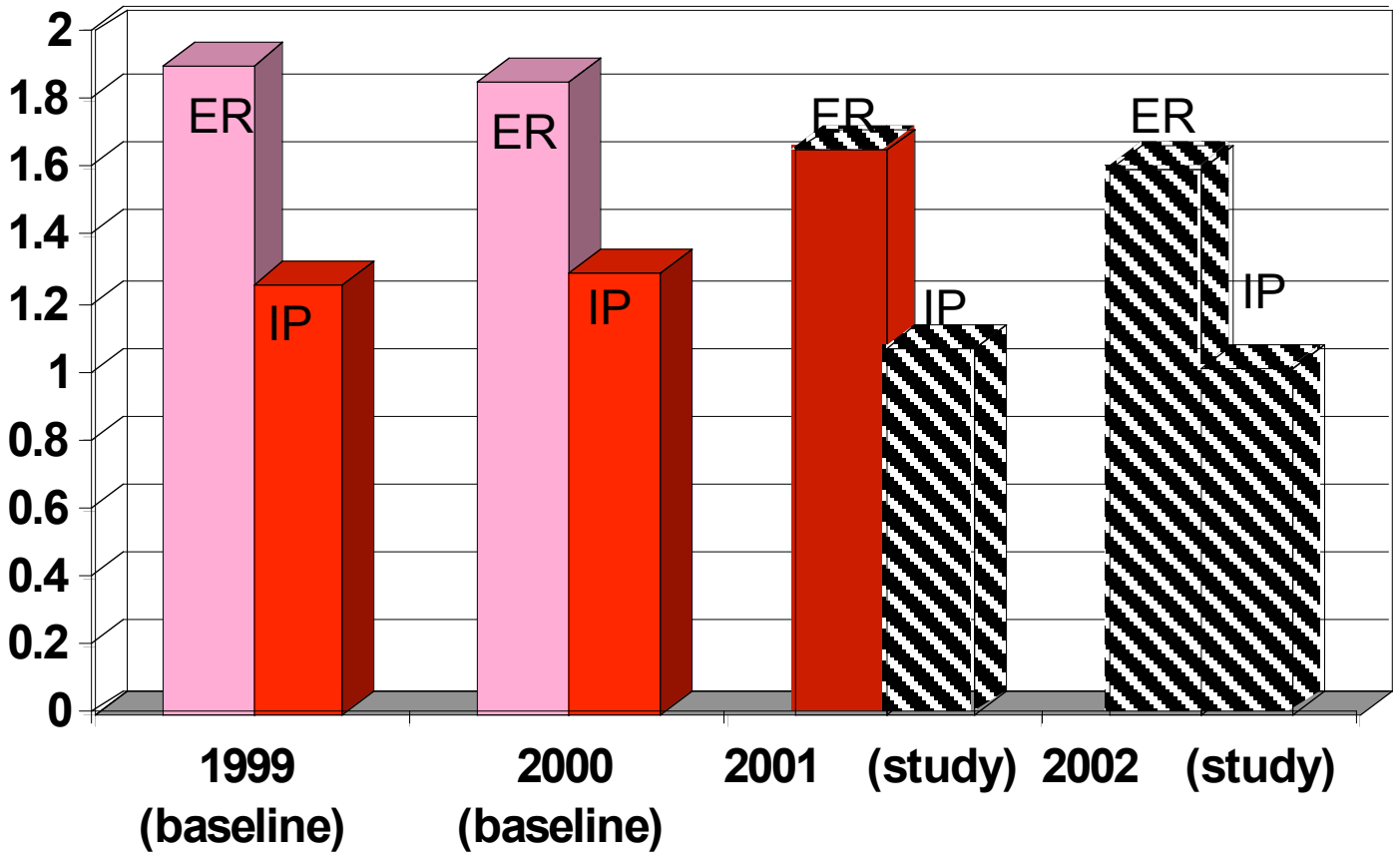
Baseline trend for asthma ER and IP Utilization

493.xx ER visits and IP stays/1000 planwide



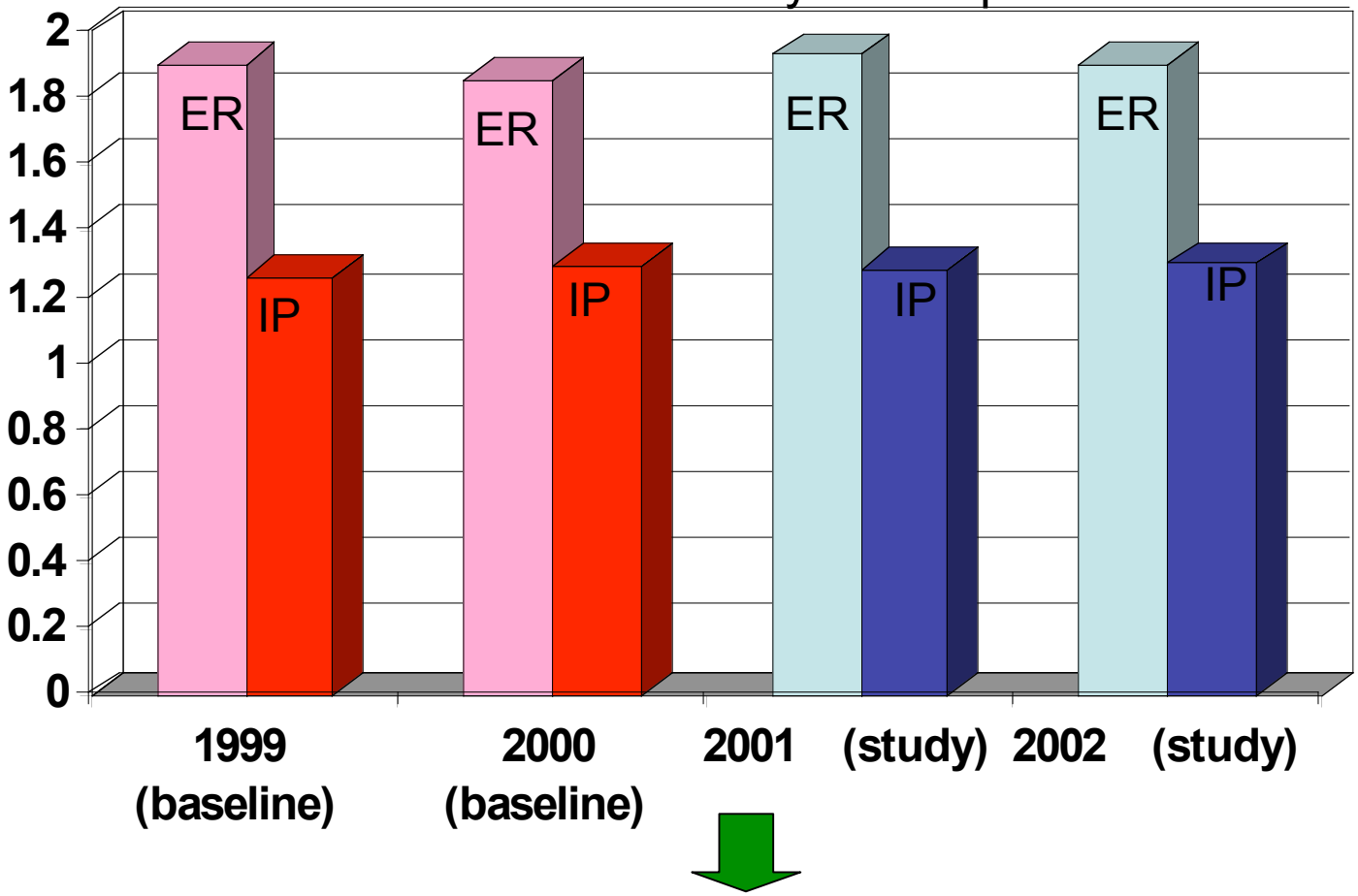
Expectation is something like…

493.xx ER visits and IP stays/1000 planwide



Plausibility Indicator Actual:
Validation for Asthma savings from same plan including ALL CLAIMS for asthma, not just claims from people already known about

493.xx ER visits and IP stays/1000 planwide

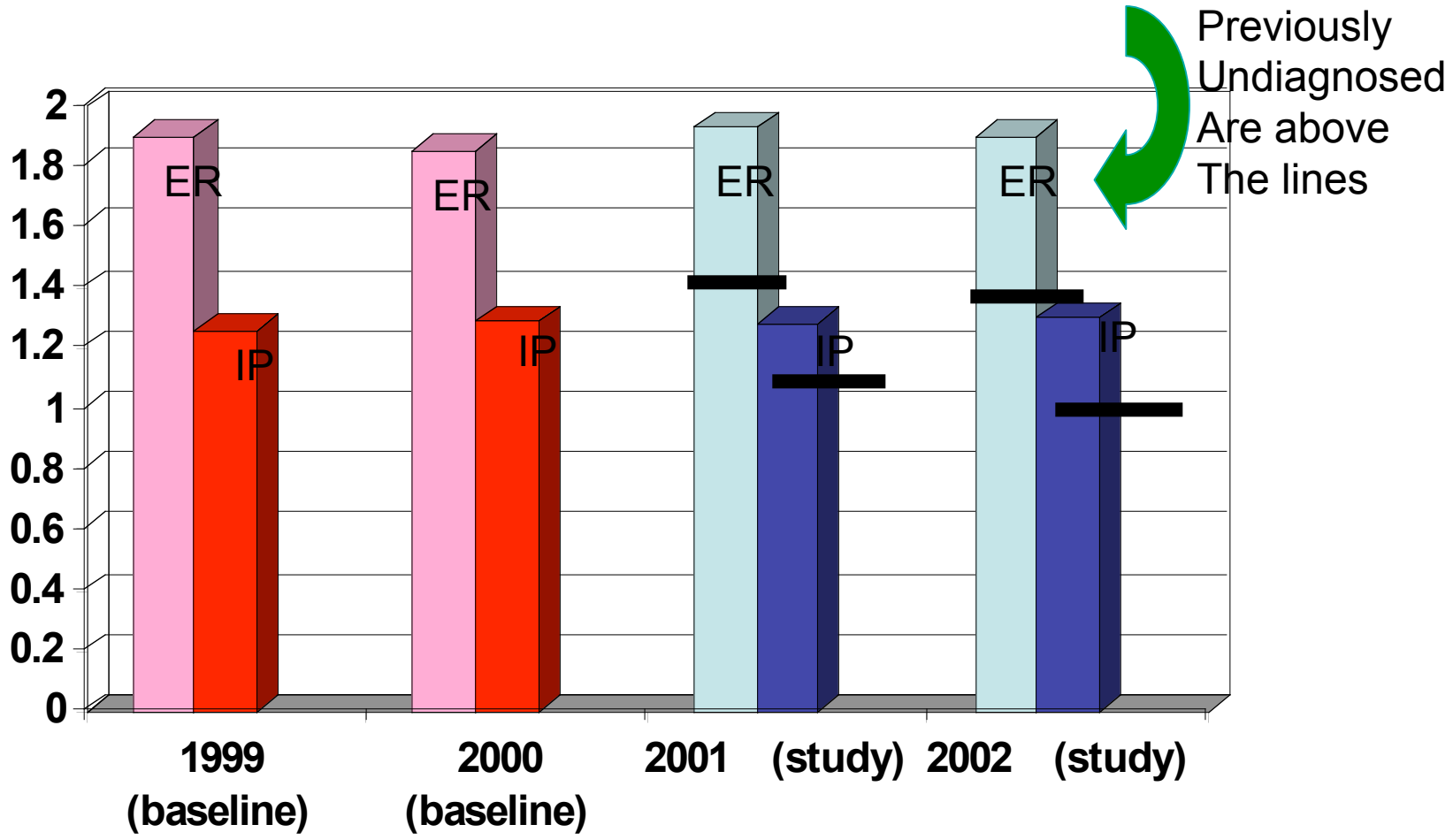


How could the vendor's methodology have been so far off?

We then went back and looked…

- …at *which* claims the vendor included in the analysis…

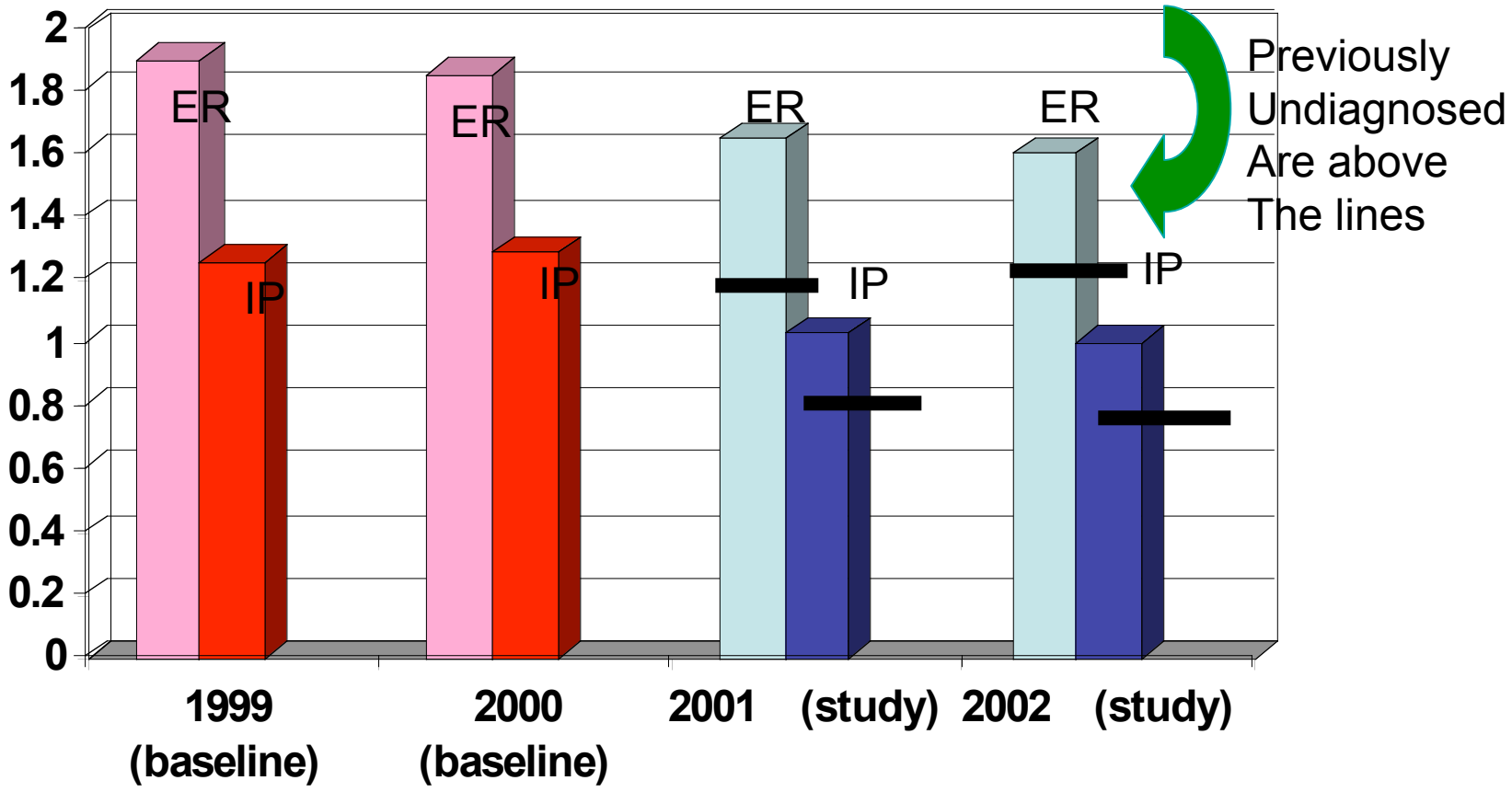
We were shocked, *shocked* to learn that the uncounted claims on previously undiagnosed people accounted for virtually all the “savings”



Is it fair…

- To count the people the vendor didn't know about?

You should be able to reduce visits in the *known* group by enough so that adding back the new group yields the reduction you claimed – otherwise you didn't do anything



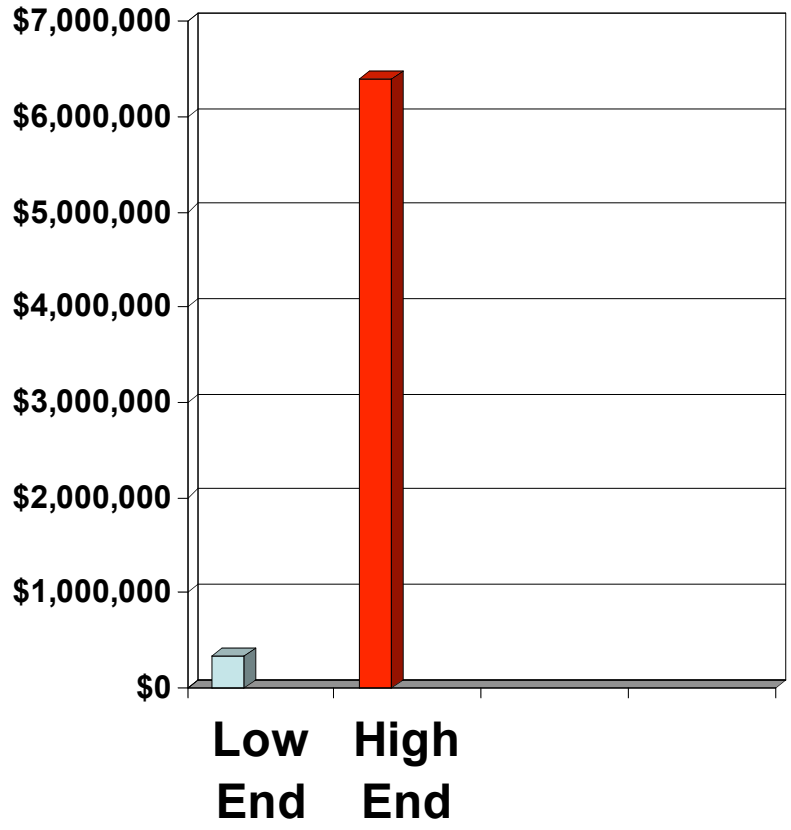
The intersection of Dummy Year and Plausibility

- “You can’t hold us responsible for people we couldn’t have known about.”

- Think about that statement. It says, “We want to ride that RTM curve down but (aside from DMPC contracts, and one vendor) we don’t offer a DYA to see what that RTM curve is

Applying Plausibility to Mercer presentation which found a “range” of possible savings in Respiratory DM

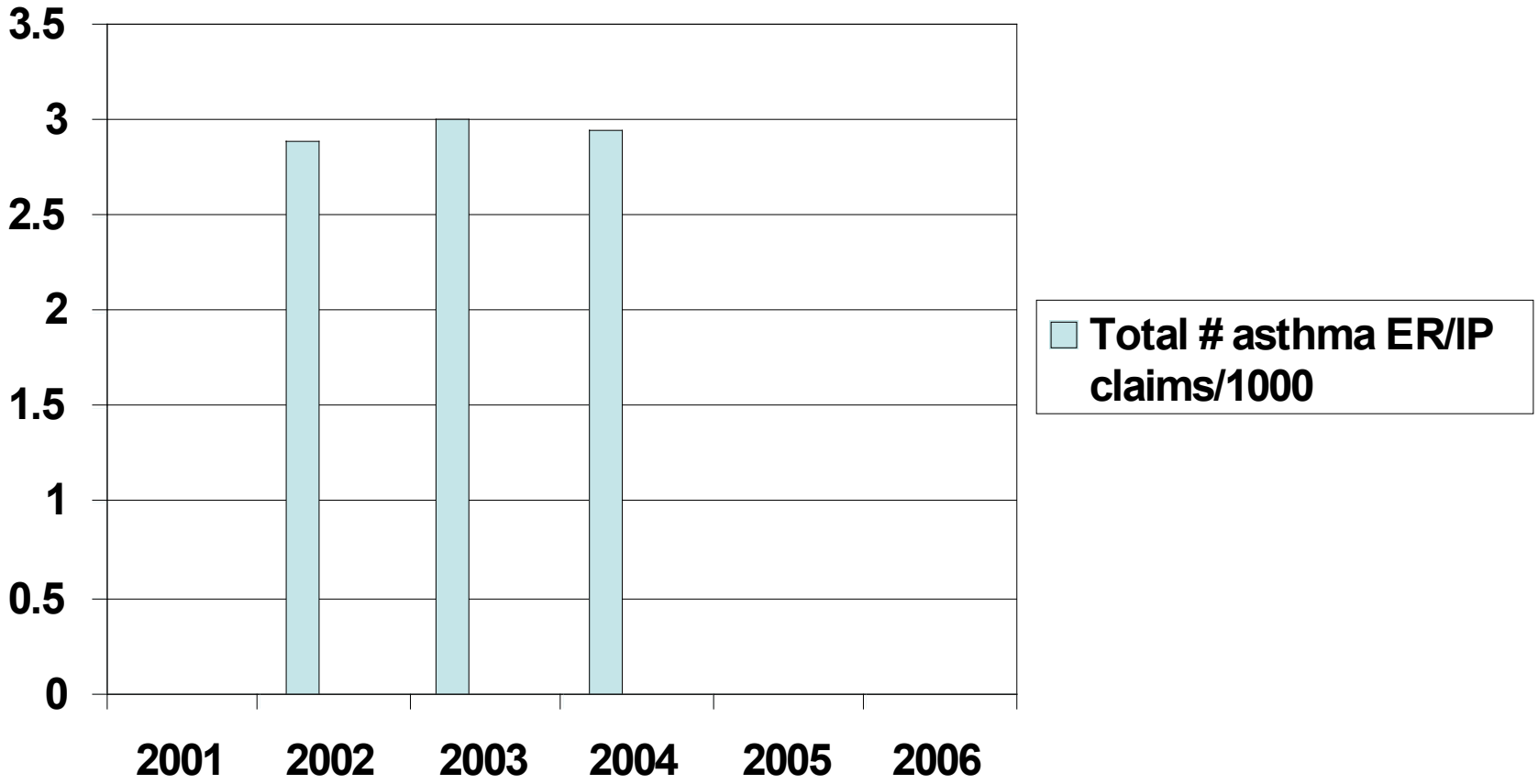
- Mercer’s view: “Varying the methodology has a significant impact on the results” Results “somewhere in that range”
- Our View: There is only one right answer and a Plausibility test will point to it



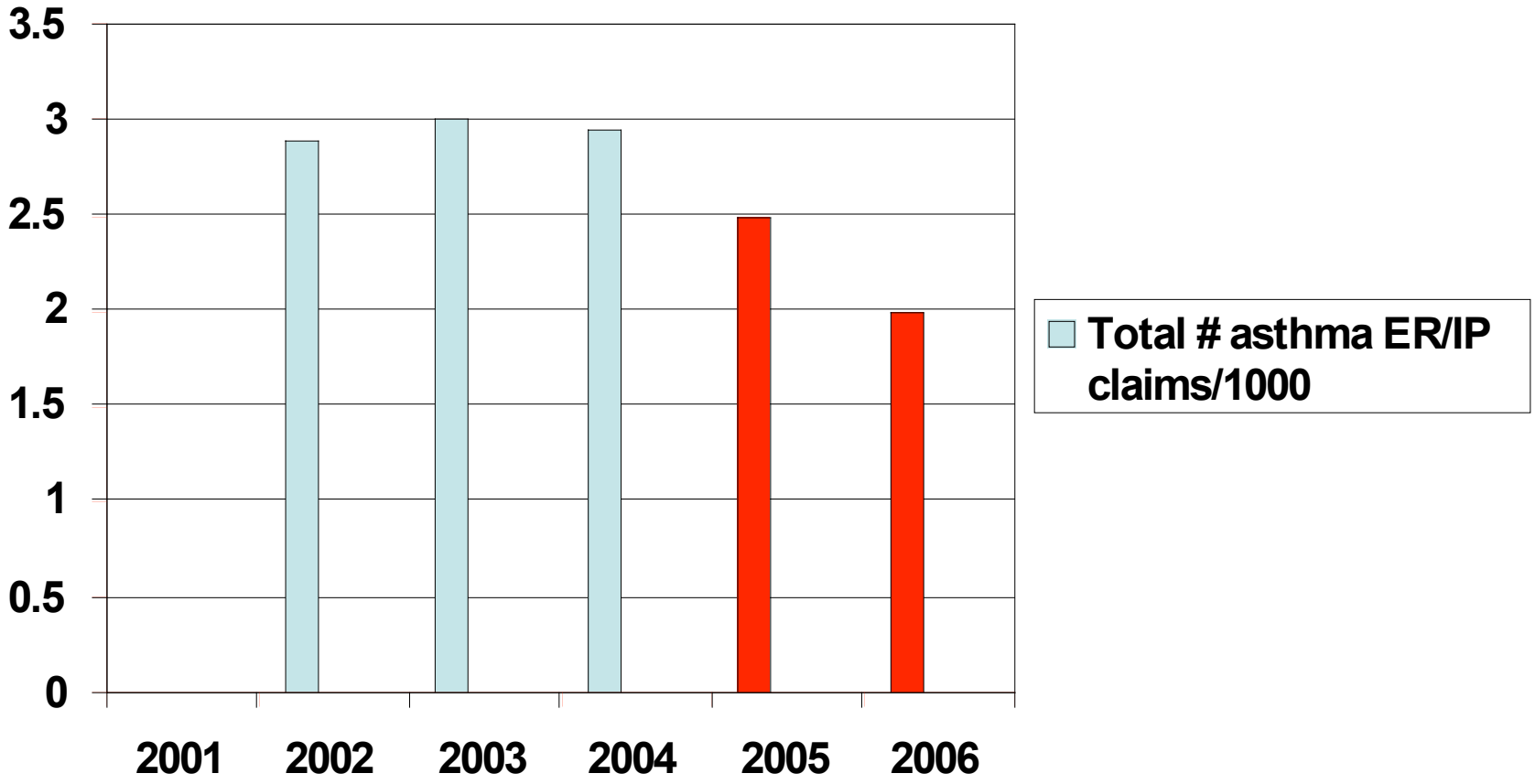
How Mercer could do a plausibility test on asthma

- Take two-three years of claims history in all primary-coded 493.xx claims for ER and IP
- Add together and divide by # of covered lives to get a rate
- Then Ask: What happens in the program year?

Possible trend prior to program



For the program to have saved $6-million, this indicator would have to plunge (it didn't)



Let's Macro-Plausibility-Test Wellness

- The Dummy Year Analysis
- Plausibility Testing
  - For Wellness
- Critical Outcomes Report Analysis

Macro Plausibility for Wellness
Here's how you know wellness reports are inflated or impossible

- Compare all these reported dramatic results in smoking cessation and weight loss to CDC statistics for the US as a whole
  - Even as most large (and many smaller) companies are "producing" these results, obesity continues to climb and the drop in adult smoking rates has stalled

October 26, 2006

Drop in Adult Smoking Rate Stalls

THURSDAY, Oct. 26 (HealthDay News) -- The number of adult smokers in the United States did not change from 2004 to 2005, suggesting that the decline in smoking over the past seven years has stalled, a new federal report found.

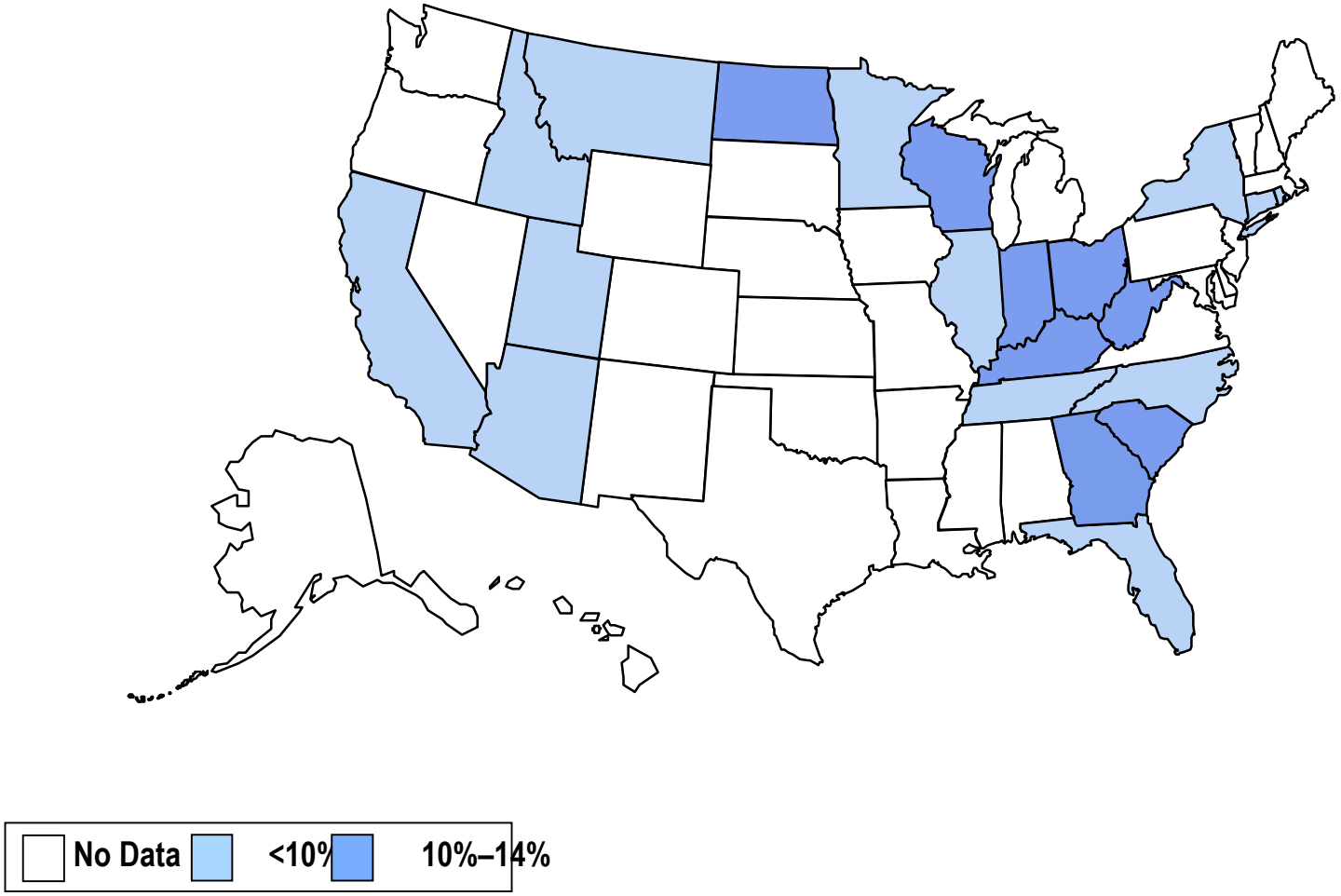
In 2005, 45.1 million adults, or 20.9 percent, were cigarette smokers – 23.9 percent of men and 18.1 percent of women. In addition, 2.2 percent of U.S. adults were cigar smokers and 2.3 percent used smokeless tobacco, according the report.

"After years of progress, what we are seeing is no change in adult prevalence of smoking between 2004 and 2005," said report author Terry Pechacek, the associate director for science at the U.S. Centers for Disease Control and Prevention's Office on Smoking and Health.

Obesity Trends* Among U.S. Adults

BRFSS, 1985

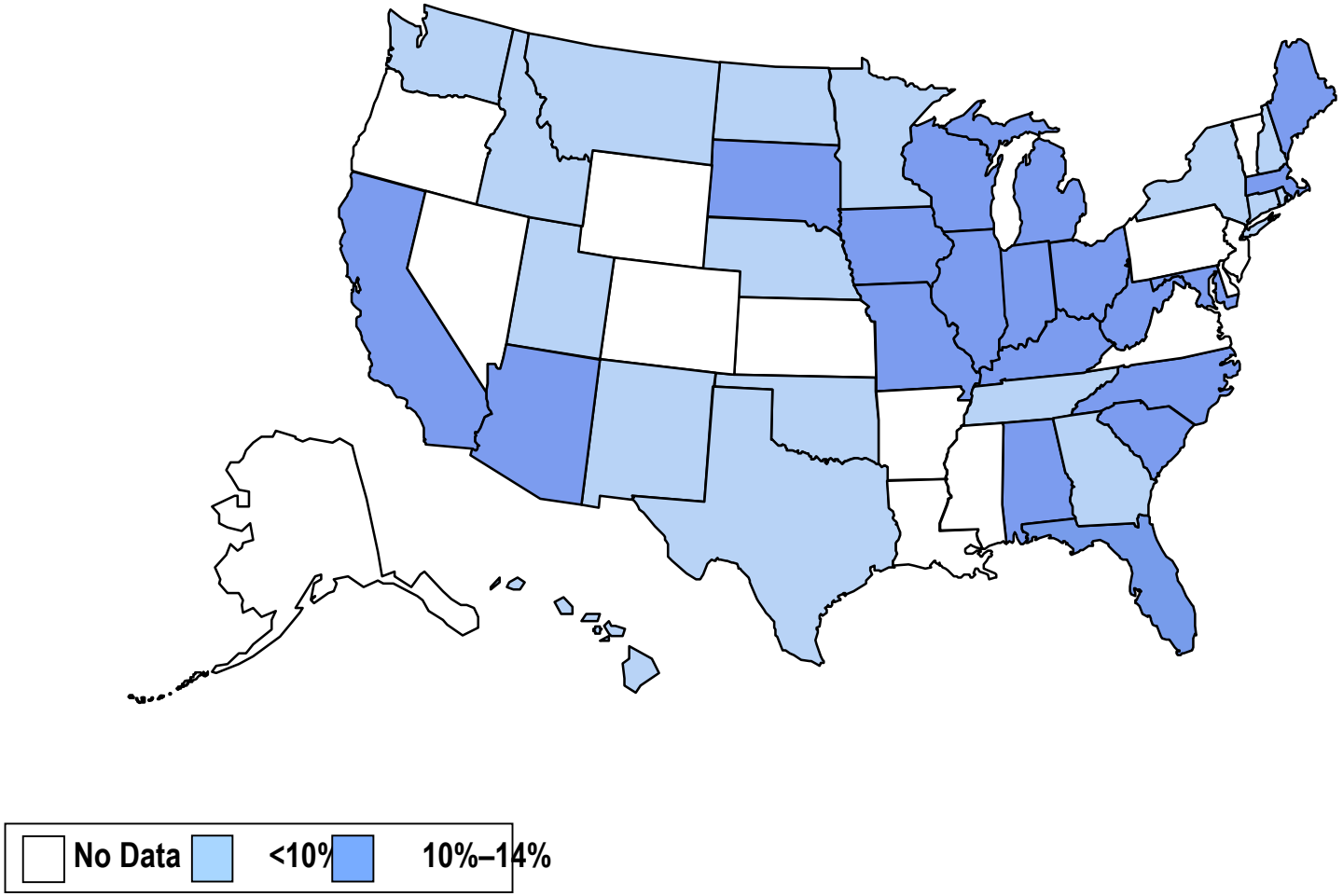
(*BMI ≥30, or ~ 30 lbs. overweight for 5’ 4” person)



Obesity Trends* Among U.S. Adults

BRFSS, 1988

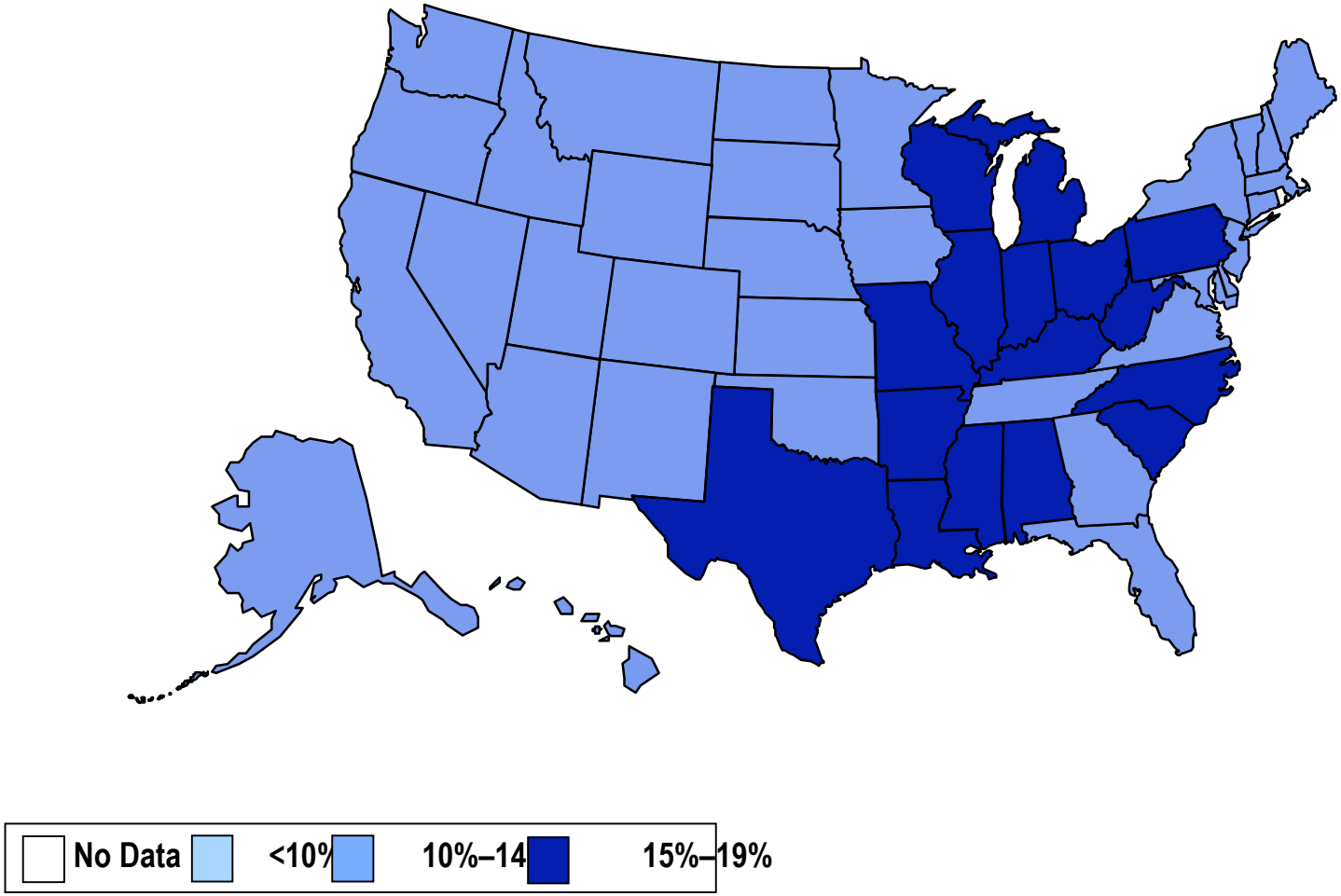
(*BMI ≥30, or ~ 30 lbs. overweight for 5’ 4” person)



Obesity Trends* Among U.S. Adults

BRFSS, 1994

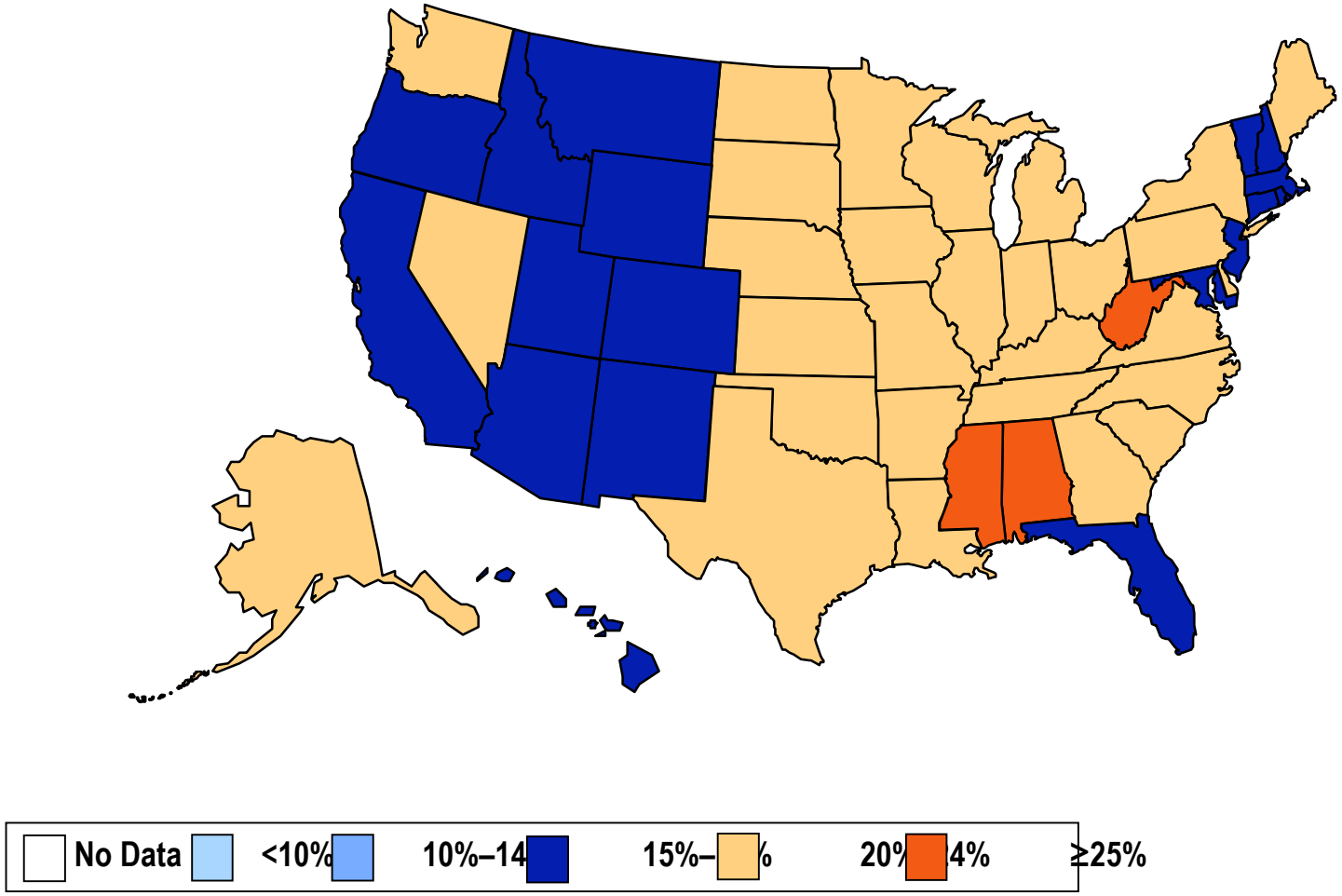
(*BMI ≥30, or ~ 30 lbs. overweight for 5’ 4” person)



Obesity Trends* Among U.S. Adults

BRFSS, 2002

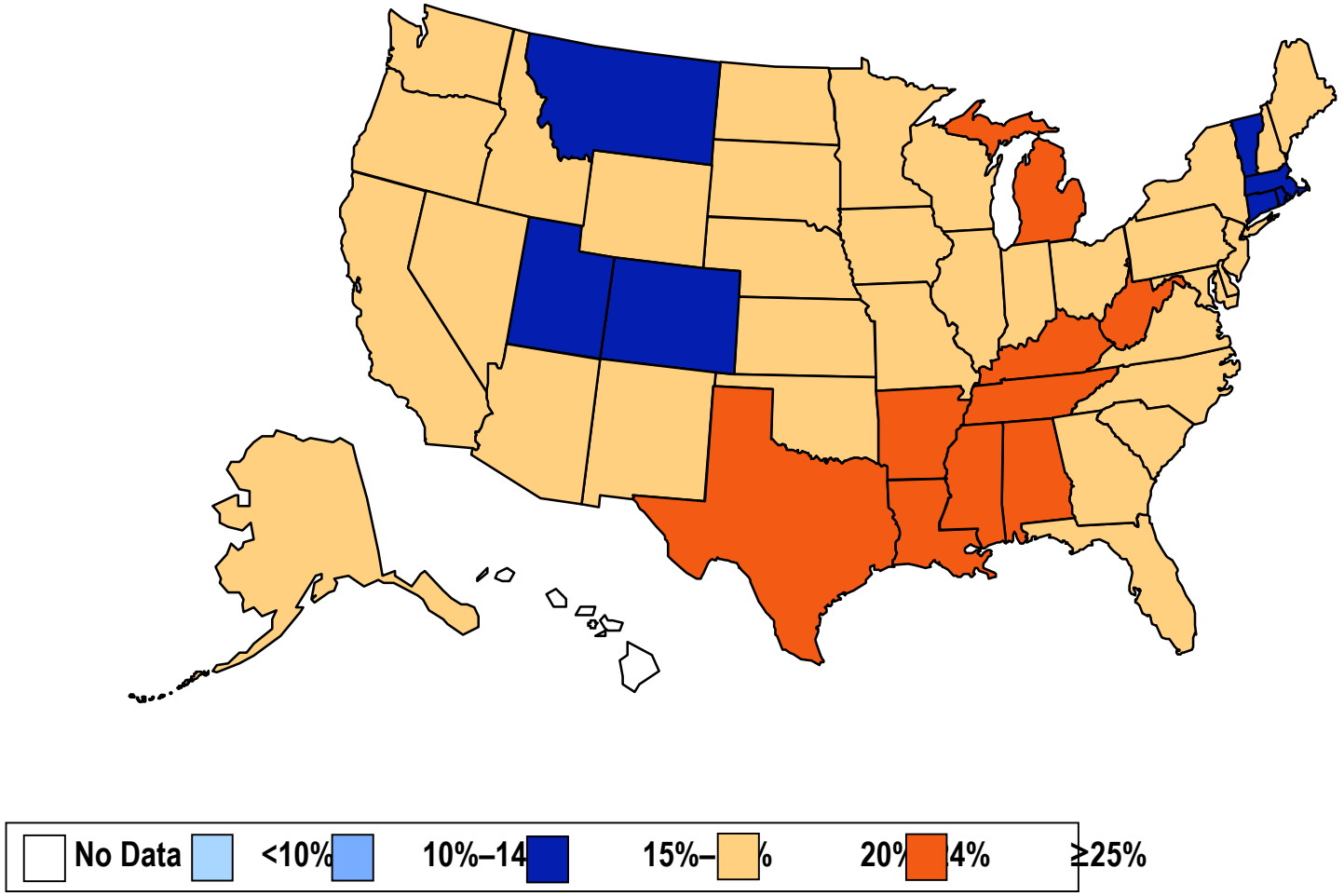
(*BMI ≥30, or ~ 30 lbs. overweight for 5’ 4” person)



Obesity Trends* Among U.S. Adults

BRFSS, 2004

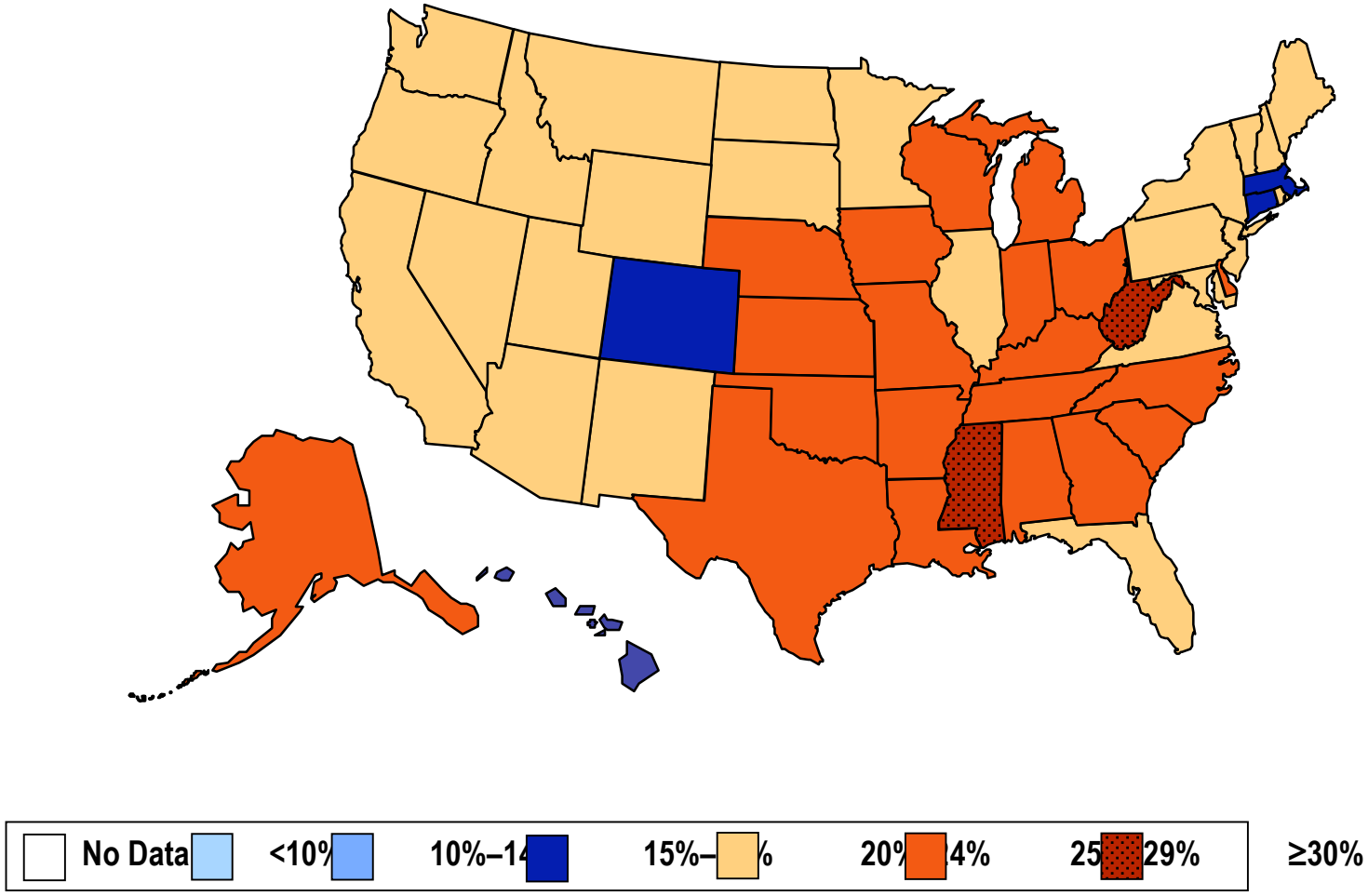
(*BMI ≥30, or ~ 30 lbs. overweight for 5’ 4” person)



Obesity Trends* Among U.S. Adults

BRFSS, 2006

(*BMI ≥30, or ~ 30 lbs. overweight for 5’ 4” person)



Summary of DYA and plausibility

- DYA and plausibility are both ways to check the same thing: Whether your results are due to the measurement or the intervention.
- We recommend checking plausibility first. Often you can be conclusive one way or the other.
  - Plausibility is also fast and inexpensive, and works on long-term programs
  - You can also benchmark it against other health plans performance using DMPC tools!

Questions on DYA and plausibility

- Pre-submitted ones and new ones

Three reasons reports are wrong

1. No one does a Dummy Year Analysis

   The exact same methodology applied to a year in which you did not have disease management

2. No one checks for plausibility
3. No one says, “wait a second – this doesn’t make sense.” This is Critical Outcomes Report Analysis

Why CORA is so important

- Most reports contain major errors, even “controlled studies.”
  - Not just small errors, but major ones easily found by CORA-certified professionals
  - I just got through reading a set of bids where only one sample outcome was even plausible
- If you are a health plan, you want to be only paying for results which you are getting
- Eventually benefits consultants will figure this out. (So far only a few have.)
- When they do, you want to be sending them reports which they can’t easily blow up

After the CORA test…

- You will probably pass this test (60% do)
- HOWEVER, that's because your antennae are now up because you know that 80% of these slides have big mistakes on them or they wouldn't be in the test
- You need to keep those antennae up when you go back to the office

Agenda

1:00	Overview of *why* reports are wrong and how to fix them. This will help somewhat in reading them and in contracting for DM but critical outcomes report analysis is about learning how to read these things generally Sample question and answer
2:00	Test
3:00	Return tests and break
3:15	Going over the answers. Email lines will be open
3:45	Adjournment of formal session. I will be available until 5:00 to answer followup questions privately on phone or email

Sample Question

- Look at each of these slides and both together to find major reporting concerns if any

Table 1: Inpatient Impact of Program (Year One)

Disease	Baseline IP days/1000	Program IP days/1000	Change
Asthma	996	747	-25%
CAD	1897	1391	-27%
CHF	9722	8581	-29%
COPD	2512	2151	-14%
Diabetes	1534	1522	-1%

Table 2: Impact on Physician Visits

Disease	Baseline MD visits/1000	Program MD Visits/1000	Change
Asthma	6990	5907	-15%
CAD	8829	8580	-3%
CHF	7876	7506	-5%
COPD	8481	8090	-4%
Diabetes	7927	7737	-2%

What you might have noticed – first slide

- No plausibility test for very high utilization reduction
- Asthmatics don't have 996 days per 1000
  - Not clear whether they are referring to days per 1000 disease members or days per 1000 overall (either way, it's wrong)
    - Almost certainly it's the first, which means no plausibility check was done
- Nor does CHF have so many days per 1000
- CHF days did not decline 29%

Second slide, and both combined

- Ridiculously high number of doctor visits
- Doctor visits should be going up or staying the same, not going down
  - This suggests strongly that a DYA is needed because they seem to have selected a high-utilizing sample as a baseline
- No correlation between MD-intensity and IP-intensity of diseases

Agenda

1:00	Overview of *why* reports are wrong and how to fix them. This will help somewhat in reading them and in contracting for DM but critical outcomes report analysis is about learning how to read these things generally Sample question and answer
2:00	Test
3:00	Return tests and break
3:15	Going over the answers. Email lines will be open
3:45	Adjournment of formal session. I will be available until 5:00 to answer followup questions privately on phone or email