

# FROM DARWIN TO COCHRANE: EVOLUTION AND EVIDENCE BASED MEDICINE

**American College of Surgeons Clinical Congress  
October 26, 2011**

**R. Lawrence Moss, MD  
Nationwide Children's Hospital  
The Ohio State University College of Medicine**



# GOAL

- 1. To show you how we as individuals, and as a collective group of specialists, are prone to forces that can prevent us from critical thinking that best helps us make good choices in patient care**
- 2. To ask you to commit to consciously overcoming this challenge**



**NATIONWIDE CHILDREN'S**  
*When your child needs a hospital, everything matters.™*



# DISCLOSURE

The speaker has no relevant financial relationship(s) which affect the content of this presentation.



# EXPERIENCE

photo

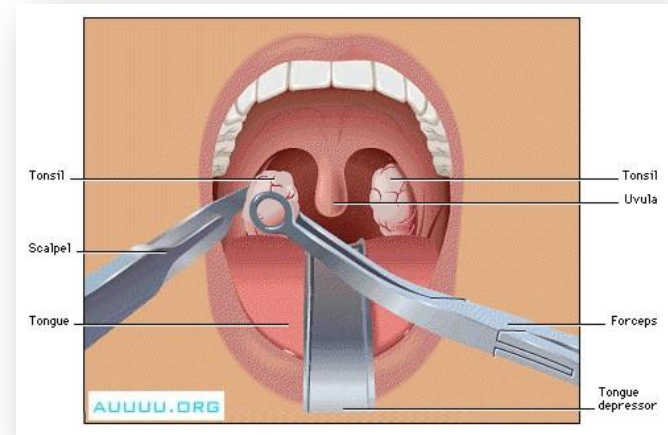
# OBJECTIVES

- To convince you that our basic human cognitive structure gives us a strong tendency to incorrectly learn from our experience
- To show you that even in those rare instances where the evidence exists we often do not choose to use it
- To suggest it is possible for us to overcome these “evolutionary” tendencies

# TONSILLECTOMY IN THE UNITED STATES

- For >50 years done to prevent pharyngitis
- In 1959 - 1.4 million tonsillectomies in US
- Never a scientific study showing that it worked—review of 3765 papers revealed none showing unequivocal effectiveness
- In 2006 – consensus that indications included airway problems and middle ear disease. Estimated <1200 done for “pharyngitis”

# HOW COULD THIS HAPPEN?



- “Experienced” clinicians perceived that patients did better after tonsillectomy
- Subsequent generations of physicians learned from their more “experienced” mentors
- Patients were “protected” from a study of something so obviously proven effective by experience

**(THERE IS) “A SOLID BODY OF  
COGNITIVE EVIDENCE THAT  
PEOPLE DO NOT LEARN  
EFFECTIVELY FROM EXPERIENCE,  
ESPECIALLY NOT WHEN THE  
EXPERIENCE CONSISTS OF A  
SERIES OF CASES”**

Brehmer, B: *Acta Psychologica* 45(223-141). 1980





**THE EFFECTIVENESS OF CLINICIANS' JUDGMENTS:  
THE DIAGNOSIS OF ORGANIC BRAIN DAMAGE FROM  
THE BENDER-GESTALT TEST <sup>1</sup>**

LEWIS R. GOLDBERG <sup>2</sup>  
*University of Michigan*

**COMPARED EXPERIENCED MASTER  
CLINICIANS TO THEIR SECRETARIES IN  
THE DIAGNOSIS OF ORGANIC BRAIN  
INJURIES**



**NATIONWIDE CHILDREN'S**  
*When your child needs a hospital, everything matters.™*



# RESULTS: CLINICIANS V. SECRETARIES

GROUP	% CORRECT	RANGE	% OF SUBJECTS PERFORMING BETTER THAN CHANCE (p<0.05)
Experienced attending clinicians	65%	60-70%	25%
Secretaries	67%	67-73%	62.5%

Goldberg. J Conc Psych, 1969

# FUNDAMENTAL CONCEPTS OF HOW THE HUMAN MIND PROCESSES EXPERIENCE

1. The inherent weakness of inductive reasoning
  2. Our innate tendency to confirm rather than refute hypotheses
  3. Our basic cognitive structure needs “rules” to explain the world
  4. Humans have deterministic minds in a probabilistic world
-

# INDUCTION



## The problem of induction

- **Hume 1765 – how things behave when they go beyond the present**
- **We need to see the world as “regular”**
- **Humans are creatures of emotion**
- **We are conditioned to respond to compelling personal experience – not cold analysis of data**

# THESE SIMPLE TRUTHS SUPPORTED BY MOUNTAINS OF DATA ARE ONLY RECENTLY BEING ACCEPTED

## SMOKING CAUSES CANCER AND DEATH

“my grandfather smoked two packs a day and lived until 88 years old”

## DRUNK DRIVING KILLS

“I have driven home after drinking many times and I have never had an accident”

## “X” IS NOT AN EFFECTIVE OPERATION

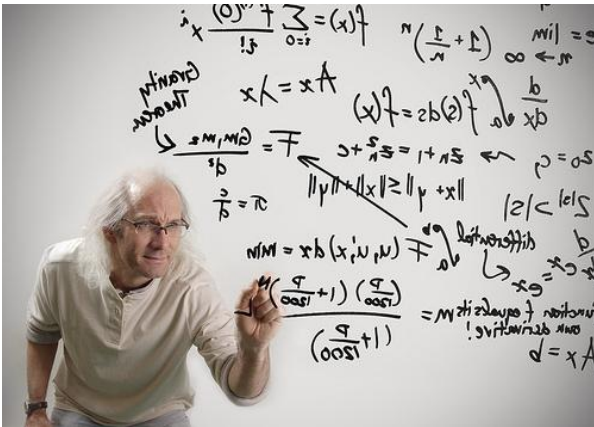
“I was trained to do this operation and in my hands have had good results with it for years”

# WHICH VARIABLES ARE IMPORTANT?

Classic psychology experiment in 1937

Subjects asked to determine a rule explaining a sequence of numbers

A group of 4<sup>th</sup> graders markedly outperformed a group of full professors of mathematics from Ivy League universities



# INDUCTION ASSUMES THAT YOU KNOW WHICH VARIABLES ARE IMPORTANT IN CLASSIFYING DATA

- Genetic predisposition
- Severity of disease
- Original anatomy
- Choice of medications
- Emotional state of the patient
- Random chance
- Technical skill of the surgeon
- Co-existent disease that is unrecognized
- Quality and type of anesthetic
- Choice and rate of IV fluids
- Innate immune defenses
- Choice of operation
- Virulence of the individual patient's condition
- Quality of post-operative care
- Co-existent disease that is recognized
- Quality and type of anesthetic
- Choice and rate of IV fluids
- Innate immune defenses
- Virulence of the individual patient's condition
- Quality of pain management
- Variability in natural history of disease

# OUR INNATE TENDENCY TO CONFIRM RATHER THAN REFUTE HYPOTHESES

- Human beings create positive hypotheses (we set out to prove the care we provide our patients is good not that it is inferior)
- As we test these “positive” hypotheses with our experience we inherently overvalue confirming experience and undervalue refuting experience



# TPN CHOLESTASIS AND FAT

or “my personal failure”



**NATIONWIDE CHILDREN'S**  
*When your child needs a hospital, everything matters.™*

# OUR BASIC COGNITIVE STRUCTURE NEEDS “RULES” TO EXPLAIN THE WORLD

When confronted with a problem of understanding and predicting complex stimuli, human beings will search for a rule and apply it

We are more likely to cling to a rule that is wrong than accept there is not a rule

# A LITTLE LEARNING ... DECISION RESEARCH

**Fischhoff and Slovic, 1978**

- Examined how subjects processed limited experience with complex stimuli
- Analyzed how subjects created rules and how they used them
- Tasks were picking stocks for profit and handwriting analysis

# RESULTS

- Even with very limited experience, subjects rapidly found a “rule”
- This led to a highly unwarranted confidence in the validity of the rule
- When presented with circumstances where the rule failed they tended to believe that the circumstance was an exception and the rule was still valid. Even when the exceptions exceeded the evidence for the rule.

This is a great operation, it's unusual that my last three patients haven't done so well”

# RESULTS

The mere fact that a rule could be applied to a limited experience and work allowed the subjects to believe they could make appropriate determinations that would predict future events.



# HUMANS HAVE DETERMINISTIC MINDS IN A PROBABILISTIC WORLD

DETERMINISM:

If A then B

PROBABALISM:

If A then a 70% chance of C, a 20% chance of D, and a 10% chance of E

THERE IS OVERWHELMING  
PSYCHOLOGICAL EVIDENCE THAT  
HUMANS ARE “HARD-WIRED” FOR  
DETERMINISM

# DETERMINISM V. PROBABILISM

Subjects will develop a deterministic rule to explain any series of observations (Brehmer 1974)

Increasing the complexity of the problem does not reduce the tendency to develop deterministic rules (Kuylenstierna 1979)

When the selected rule fails they will develop another deterministic rule (Dawes 1974)

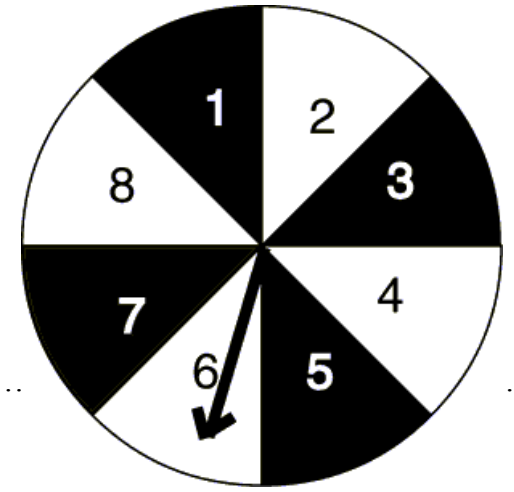
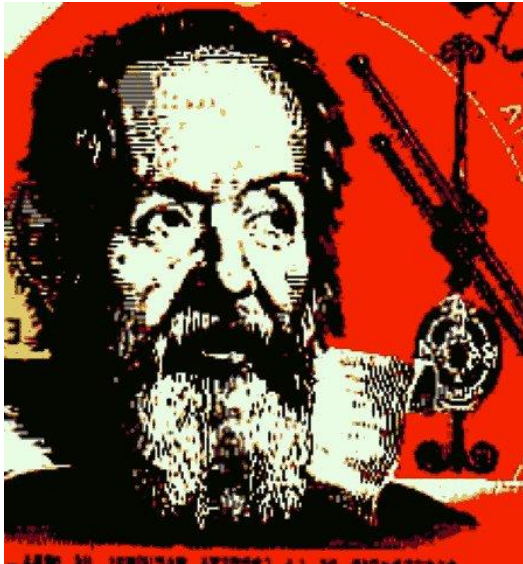
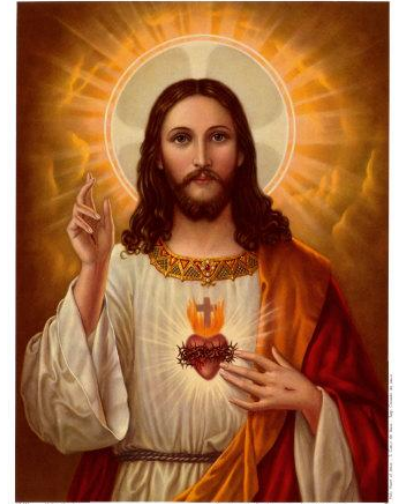


# DETERMINISM V. PROBABILISM

When a series of deterministic rules fail, subjects will give up and guess rather than try a probabilistic rule (Tversky 1978)

Even when subjects are specifically told that the outcome is probabilistic, they will still develop and persistently try deterministic rules. (Johansson 1978)

# PROBABILITY IN HUMAN HISTORY (1820's)

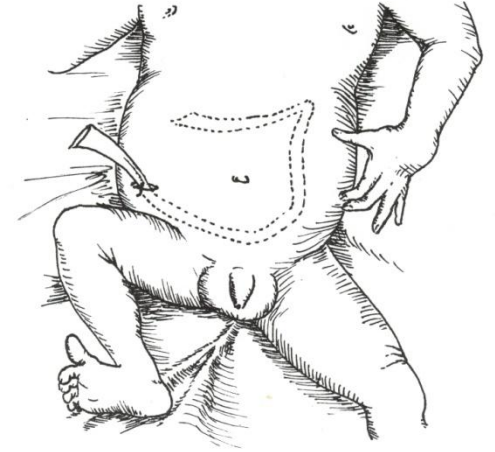


N'S

*When your child needs a hospital, everything matters.™*



# WHAT IS THE RIGHT OPERATION ?



1. There actually is a double secret rule that predicts which babies will do better with one or the other operation
2. I, Larry Moss, actually know this rule and haven't told anybody yet

This may all be true within the confines of arcane psychological theory but everyone knows that if a doctor is more experienced she is smarter and has better judgment.

## Systematic Review: The Relationship between Clinical Experience and Quality of Health Care

Niteesh K. Choudhry, MD; Robert H. Fletcher, MD, MSc; and Stephen B. Soumerai, ScD

- Systematic review to determine the relationship between clinical experience and quality of care
- 59 studies on 62 groups of physicians: thousands of physicians & millions of patients
- 4 dimensions of outcomes: knowledge, adherence to standards for diagnosis, adherence to standards of treatments, actual patient outcomes

Ann Int Med 2005, 142:260-273

# DOES EXPERIENCE IMPROVE PERFORMANCE?

## RESULTS

- 32/62 studies (52%) showed negative association in all four dimensions
- 13/62 (21%) showed negative association in some dimensions
- 2 (3%) showed initial improvement followed by decline
- 13 (21%) showed no relationship



This may be true for obscure book  
knowledge of rare problems but does not  
apply to everyday patient care?

# DOES EXPERIENCE IMPROVE PERFORMANCE?

## RESULTS

- Mortality in 39,007 patients with acute MI
- 4546 cardiologists, internists, GPs
- Controlled for severity of disease, hospital location, practice type, physician specialty, board certification, physician volume
- 0.5% increase in mortality for every year since doctor graduated from medical school





This is probably true for doddering old  
internists but can't possibly apply to  
surgeons

# OUTCOMES BY SURGEON EXPERIENCE IN CARDIAC SURGERY

- Mortality rate for 275 cardiac surgeons in 83,547 patients
- Multivariate adjustment for both physician and patient covariates
- More years in practice significantly associated with higher mortality rates ( $p < 0.001$ )



# EXPERIENCE: BOTTOM LINE

It is NOT worthless

It is much less valuable than we think it is

We are hard-wired to overvalue our experience

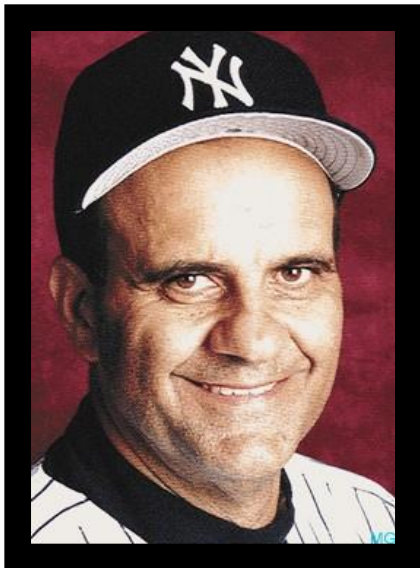
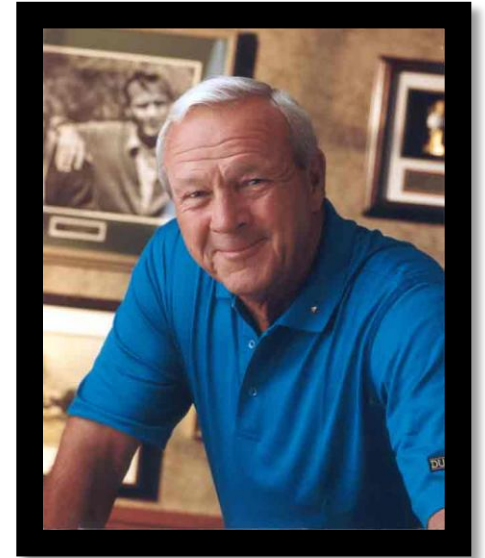
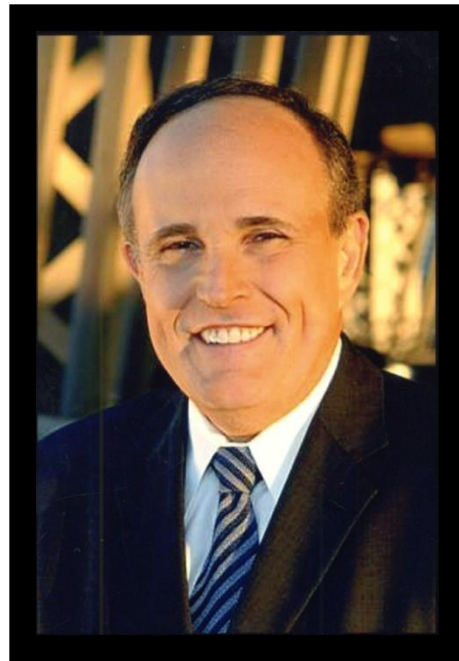
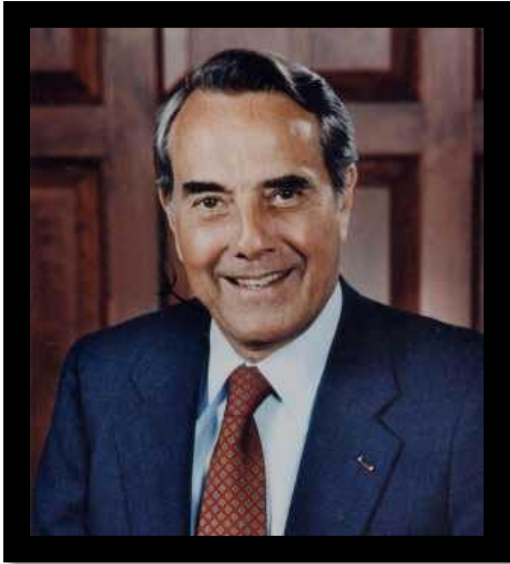
We can rise above this

This requires a conscious effort to do so

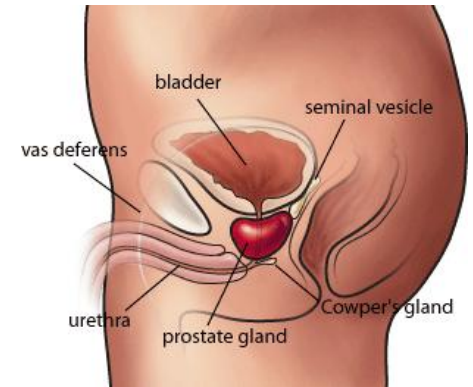
# OBJECTIVES

- To convince you that our basic human cognitive structure gives us a strong tendency to incorrectly learn from our experience
- To show you that even in those rare instances where we know the correct answer we often do not choose to use it
- To suggest it is possible for us to overcome these “evolutionary” tendencies

# WHAT DO THESE MEN HAVE IN COMMON?



# THE PROSTATE



- In middle age benign nodules become disorganized and develop CIS
- 50% at age 50    100% at age 80
- Slow process. Most men die of other causes before prostate CA
- Cells secrete PSA – can be detected

# A LOGICAL SEQUENCE OF EVENTS

**PSA IS HIGH**



**FIND AND BIOPSY THE SOURCE**



**TAKE OUT THE CANCER**



**LIFE IS GOOD**

# PROSTATECTOMY FOR PROSTATE CANCER

- Thousands of “experienced” urologists championed its benefits
- Overwhelming evidence from single institution case series studies that it saved lives
- Could not possibly “subject” these patients to a randomized trial



# The New England Journal of Medicine



## A RANDOMIZED TRIAL COMPARING RADICAL PROSTATECTOMY WITH WATCHFUL WAITING IN EARLY PROSTATE CANCER

LARS HOLMBERG, M.D., PH.D., ANNA BILL-AXELSON, M.D., FRED HELGESEN, M.D., JAAKKO O. SALO, M.D., PH.D.,  
PER FOLMERZ, M.D., MICHAEL HÄGGMAN, M.D., PH.D., SWEN-OLOF ANDERSSON, M.D., PH.D., ANDERS SPÅNGBERG, M.D.,  
CHRISTER BUSCH, M.D., PH.D., STEG NORDLING, M.D., PH.D., JUNI PALMGREN, PH.D., HANS-OLOV ADAMI, M.D., PH.D.,  
JAN-ERIK JOHANSSON, M.D., PH.D., AND BO JOHAN NORLÉN, M.D., PH.D.,  
FOR THE SCANDINAVIAN PROSTATIC CANCER GROUP STUDY NUMBER 4\*

- 695 men - newly diagnosed prostate CA
- Prostatectomy versus watchful waiting
- Median follow-up 6.2 years

# RADICAL PROSTATECTOMY (RP) VERSUS WATCHFUL WAITING RESULTS

- RP group 50% less likely to be dead of prostate CA but overall survival was identical to watchful waiting
- RP changed the cause of death but not the date

Holmberg et al NEJM 347:781-9, 2002

# **RADICAL PROSTATECTOMY (RP) VERSUS WATCHFUL WAITING QUALITY OF LIFE**

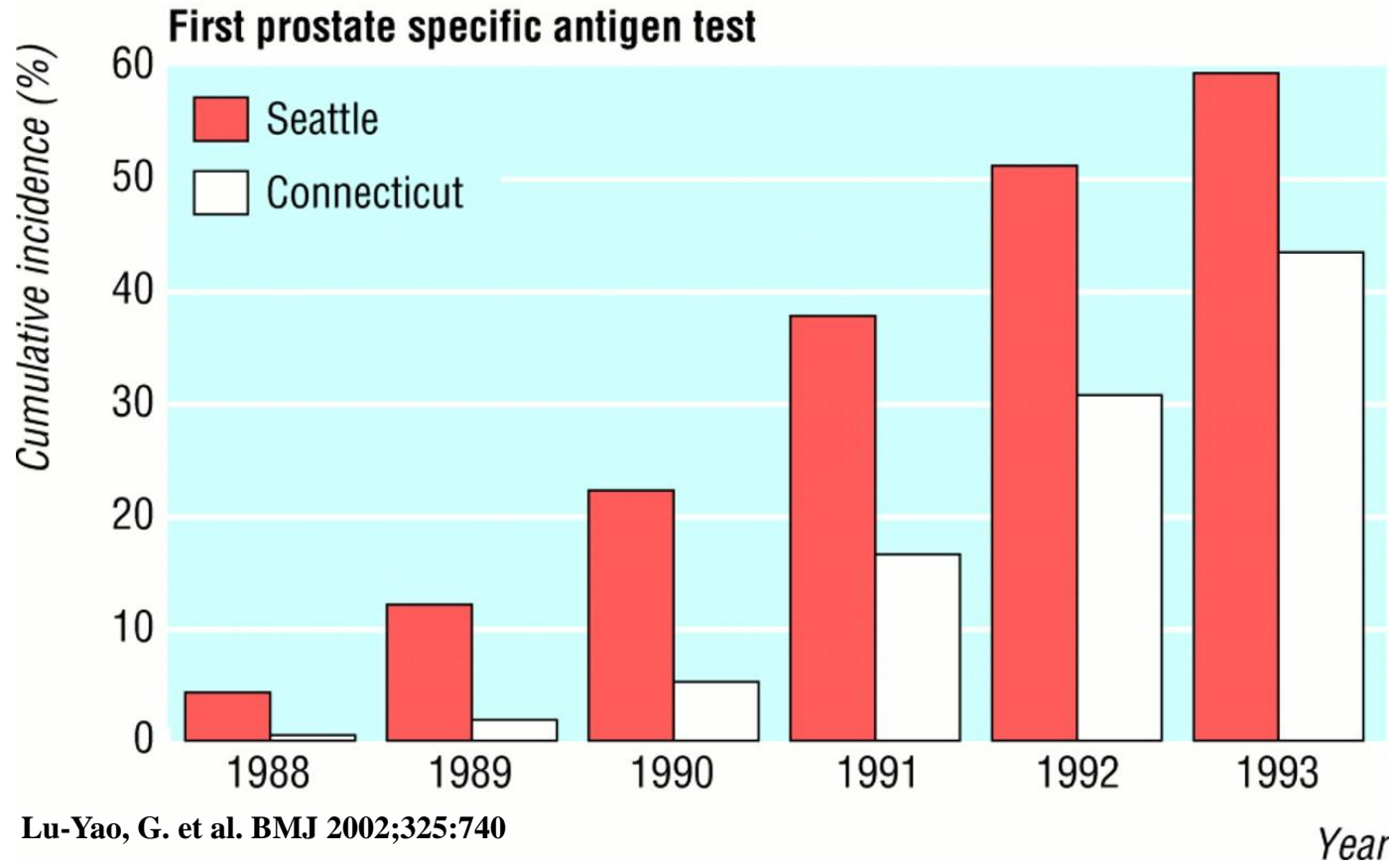
<b>FUNCTION</b>	<b>RP</b>	<b>WATCHFUL WAITING</b>
<b>Greatly distressful erectile dysfunction</b>	<b>30%</b>	<b>17%</b>
<b>Moderately distressful urinary leakage</b>	<b>29%</b>	<b>9%</b>
<b>Regular dependence upon diaper</b>	<b>14%</b>	<b>1%</b>

## Natural experiment examining impact of aggressive screening and treatment on prostate cancer mortality in two fixed cohorts from Seattle area and Connecticut

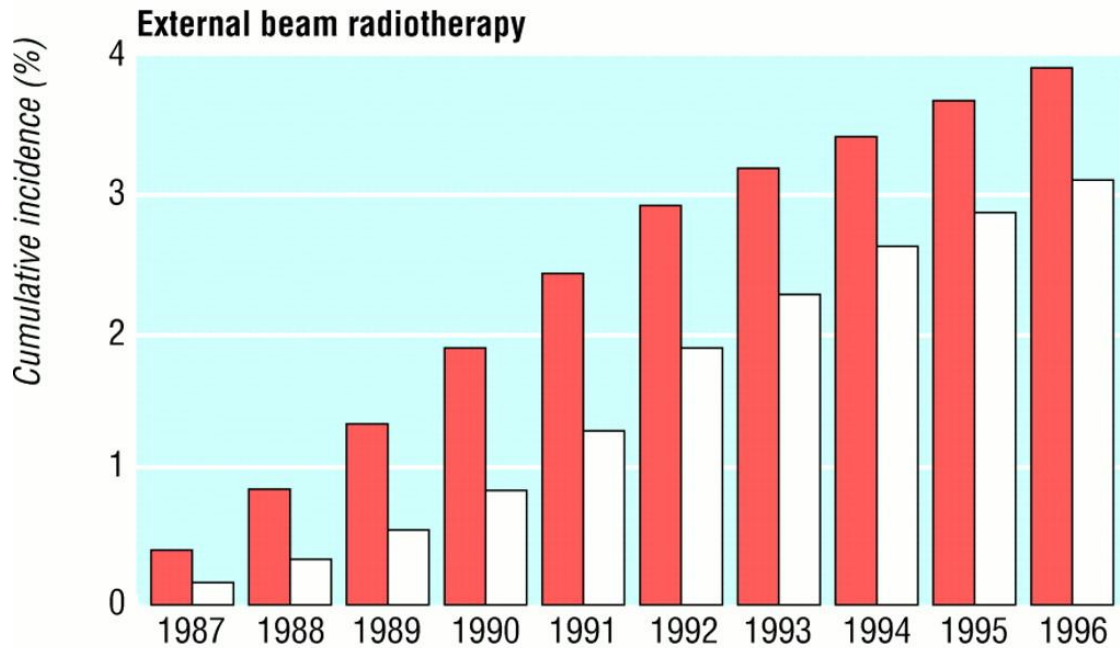
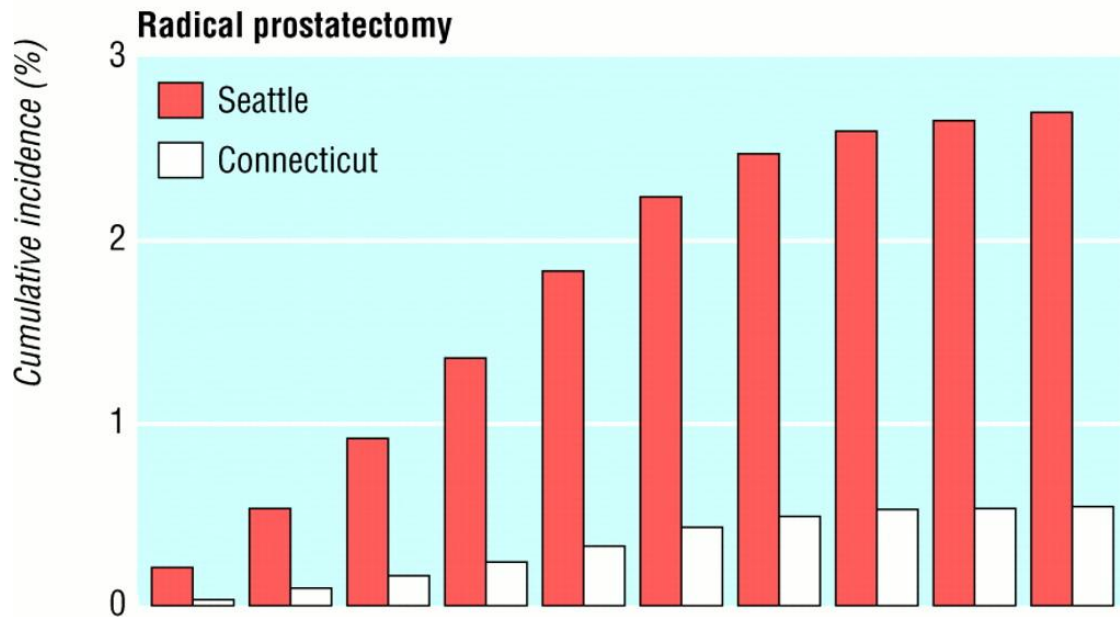
Grace Lu-Yao, Peter C Albertsen, Janet L Stanford, Therese A Stukel, Elizabeth S Walker-Corkery, Michael J Barry

- **Medicare beneficiaries from two regions**
- **PSA screening was 5 times more likely in Seattle than in Connecticut**
- **RP was done 5 fold more often in Seattle**

# PSA SCREENING

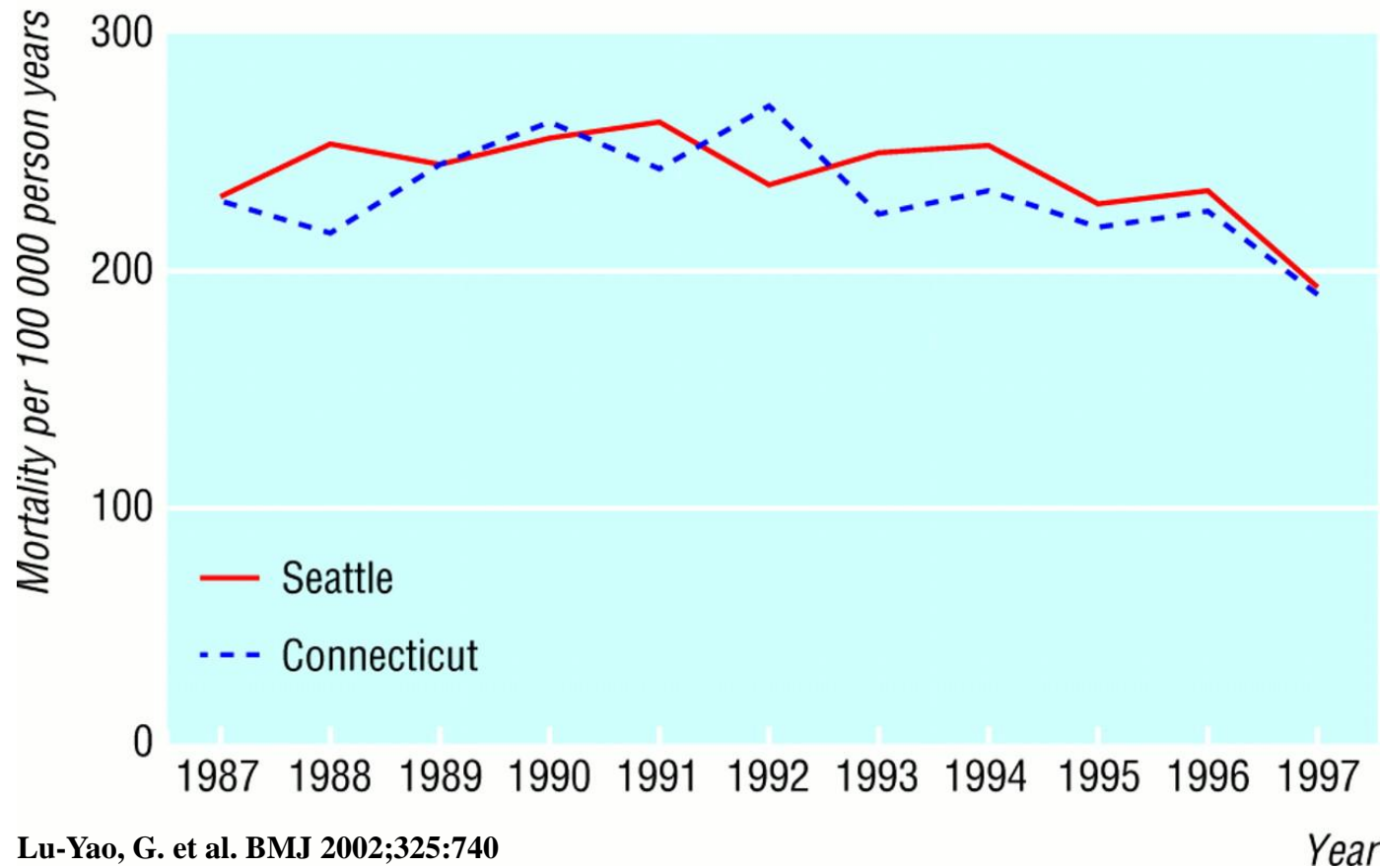


Lu-Yao, G. et al. BMJ 2002;325:740



Year

# Age Adjusted Prostate CA Mortality



Lu-Yao, G. et al. BMJ 2002;325:740

Year

BMJ

# PSA SCREENING: THE TIMELINE

- **2009 - US Preventative Services Task Force (UPSTF) creates internal document recommending against PSA screening**
- **2009 – Uproar regarding mammography recommendation**
- **2010 Chairman sends PSA rec “back for review”**
- **2010 Subcommittee chair resigns over this decision**
- **2011 UPSTF recommends against PSA screening**

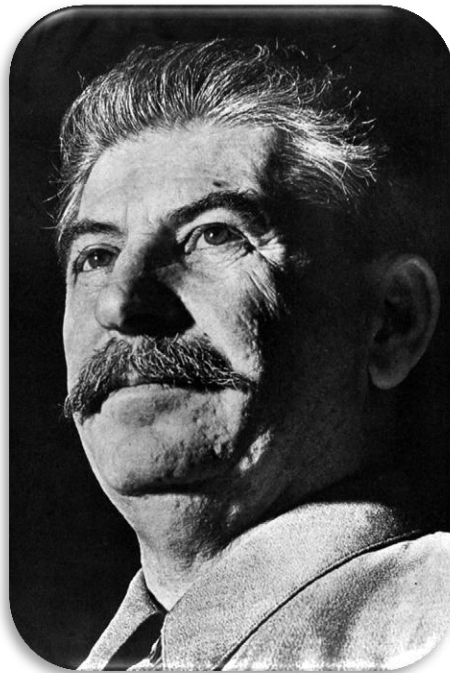


# DISSEMINATION AND ADOPTION OF MEDICAL EVIDENCE

- 6 years to achieve 25 citations in lit.
- Adopted into practice 3.2% per year
- 10-12 years to reach clinical guidelines
- 17 years to reach 50% clinical practice rate



# WHAT DOES THIS GUY HAVE TO DO WITH PEDIATRIC SURGERY?



**“One death is a tragedy,  
one million deaths is a statistic”**

# THE ASPECT OF OUR NATURE WE MUST FIGHT EVERY DAY

- One surgeon's personal anecdotal experience with an operation is a compelling human drama
- Evidence based outcomes data regarding the same operation is a boring statistic
- When we walk up to the OR table, our human tendency is to value the anecdote and ignore the data.

# WE CAN ACHIEVE THIS GOAL

- The human mind has the capacity to understand and rise above its “hard-wired” evolutionary tendencies
- All dimensions discussed have shown improvement in the past 10 years
- There is no group of individuals with a greater capacity for self improvement than pediatric surgeons

