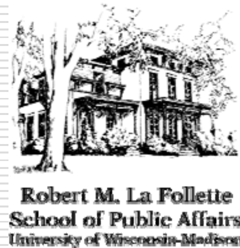
Transforming the National Health System in England: Any Lessons for the United States?

Health Care Payment Reform and Pay-for-Performance in Wisconsin: How to Promote System Transformation (and What Not to Do)

Tuesday, April 29, 2008
Monona Terrace Convention Center

Gwyn Bevan



Transforming NHS in England: Lessons for United States?

- P4P
  - Why the interest?
  - Alternative: reputational damage?
  - Potential common problems?
- NHS in England
  - Transformation through reputational damage
  - Unintended & perverse outcomes
- Reflections

Transforming NHS in England: Lessons for United States?

- P4P
  - Why the interest?
  - Alternative: reputational damage?
  - Potential common problems?
- NHS in England
  - Transformation through reputational damage
  - Unintended & perverse outcomes
- Reflections

Why the interest in P4P? Knights/knaves in health care?

- People don't work in health care for the money
  - Incentives necessary?
- Knights?
  - Choose health care as motivated to meet needs & do best possible job from resources available
- Knaves?
  - Choose health care for security & because failure results not in sanctions but rewards

Systematic reviews of literature on performance assessment

- Dominated by US evalutions
  - scale of activity & rigorous evalution
- Two hypotheses of pathways for provider action
  i. *Change* (knights): act when information identifies scope for improvement
  ii. *Selection*: act on threat to market share from patients using information as consumers
- Evidence
  i. weak &
  ii. absent

Sources: Marshall et al (2000) & Fung et al (2008)

Incentives are necessary: *Reports are not self executive*

- We do need to design a system of rewards & sanctions
- Normal market does not work for *providers* of health care
- Hence attraction of P4P

Alternative: third pathway from performance assessment

- Reputation damage
  - providers respond to information that threatens to damage their public reputation
- Requires system of performance measurement that is
  - ranking system
  - published & widely disseminated,
  - easily understood by public (see which providers performing well & poorly)
  - followed up by future reports (showing whether performance improved or not)
- Examples: Quality counts & CSRS New York

Sources: Hibbard (2008), Hibbard et al. (2003, 2005)

Controlled experiment in Wisconsin

- Two summary indices of adverse events (deaths and complications):
  - broad categories surgery / nonsurgery
  - cardiac, maternity, & hip/knee
- System characteristics
  - Ranking, easily understood, followed up
- Three groups of hospitals
  - public-report: published & widely disseminated
  - private-report
  - no-report

Source: Hibbard et al. (2003, 2005)

Controlled experiment in Wisconsin: public-report set

- significantly greater efforts to improve quality than other two sets
- no significant changes in market share
- managers of hospitals shown to have been performing poorly took action, because of concerns over impacts on hospitals' reputations

Source: Hibbard et al. (2003, 2005)

Cardiac Surgery Reporting System (CSRS) New York

- Selection ☒: market forces played no role.
  - Managed care companies did not use data way to reward better performing hospitals or drive patients toward them.
  - Nor did patients avoid high-mortality hospitals or seek out those with low mortality ...
- Change ☒ or Reputation ☑
  - impetus to use data to improve limited to hospitals named as outliers with poor performance ... [others] largely failed to use rich performance data to find ways to lift themselves from mediocrity to excellence

Source: Chassin (2002)

Impacts of different pathways in US

	Wisconsin	New York	
Change	☒	☒	
Selection	☒	☒	
Reputation	☑	☑	

Performance measures & incentives in health care

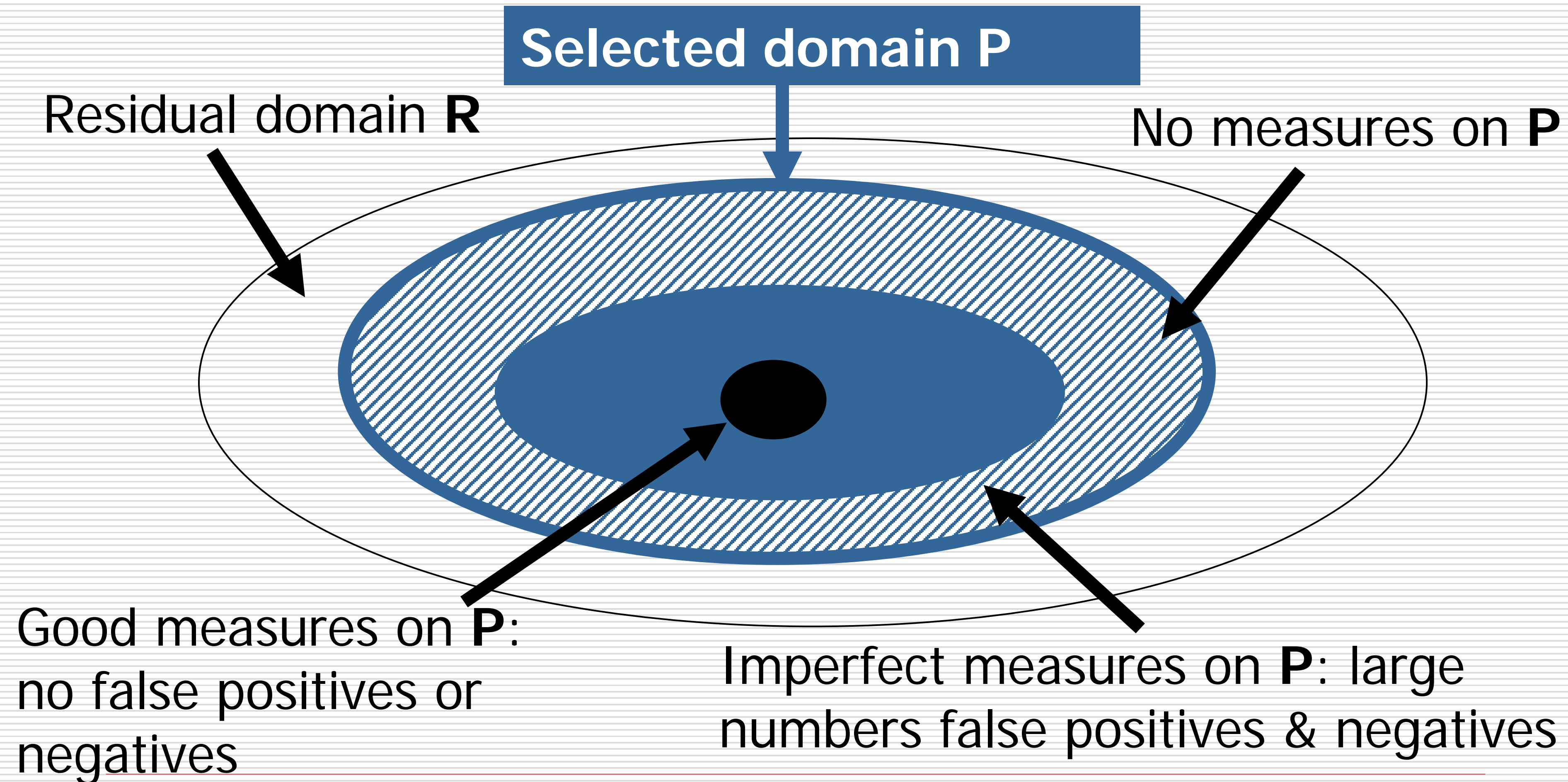
- Need incentives
- All systems imperfect
- P4P & reputational damage, common problems:

  High-powered incentives

  + imperfect measures

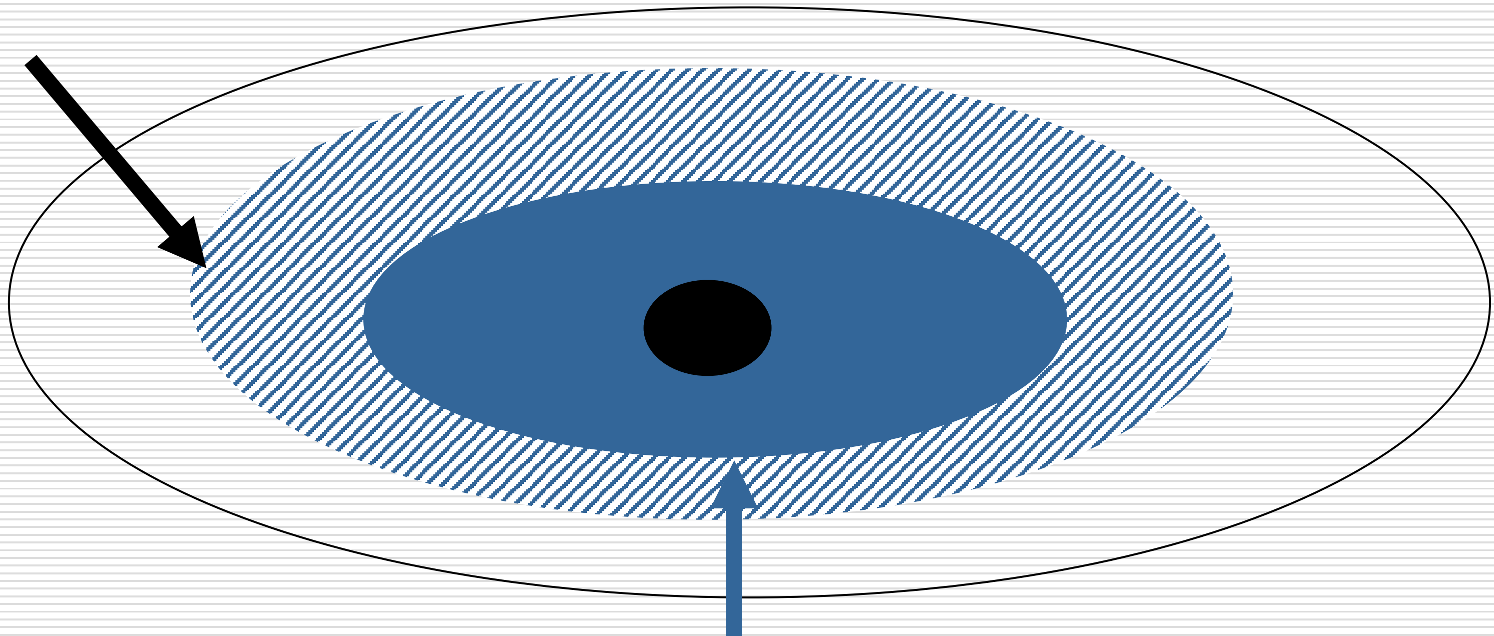
  = unintended & perverse outcomes

Problems of selection & measurement



Problem of synecdoche: part measured represent whole?

Excluded domain **N**: unimportant or cannot be measured



Selected domain T: incentives apply to measures M[T]

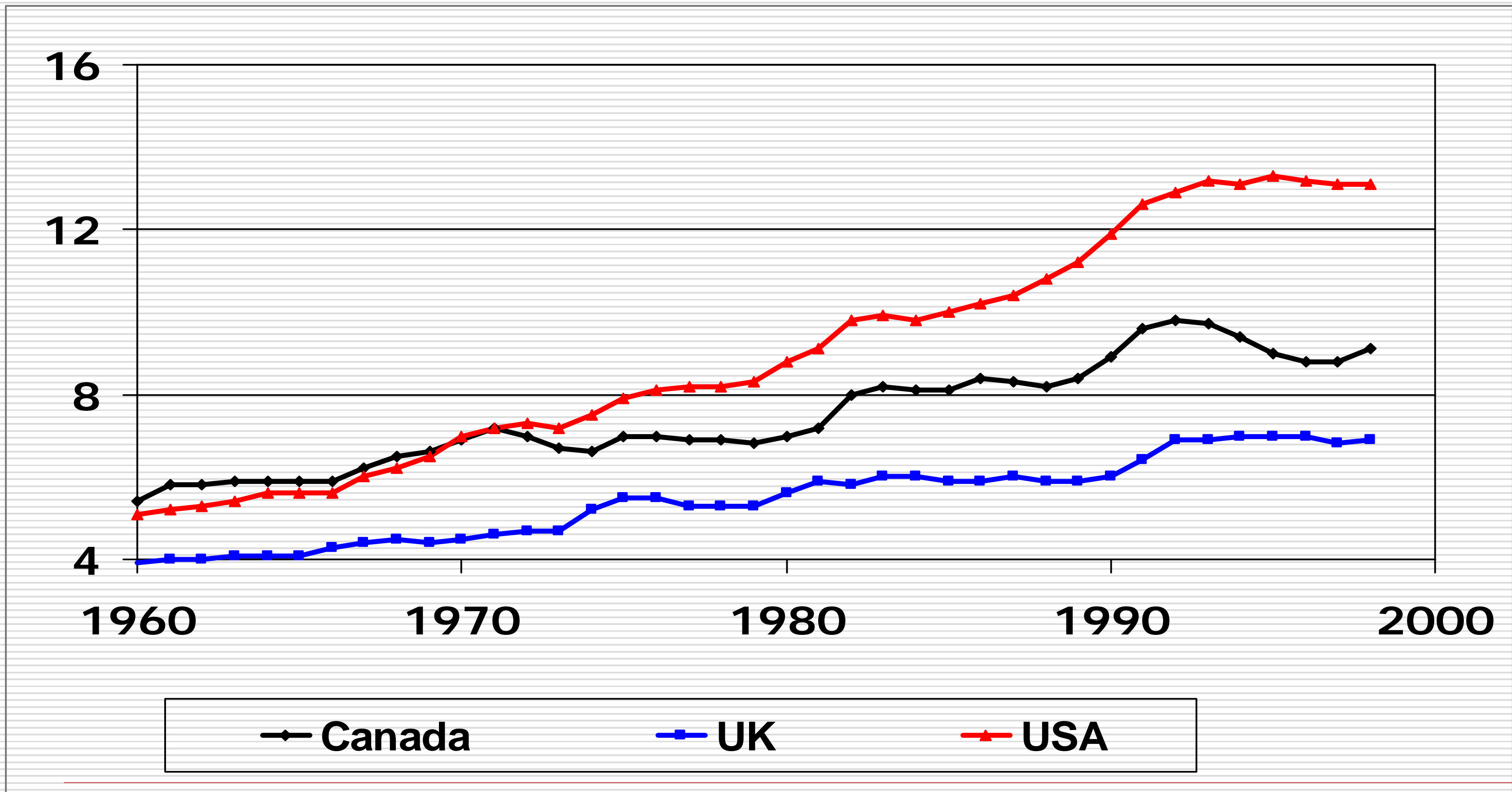
Did CSRS improve outcomes for cardiac patients in New York?

- Chassin (2002)
  - ‘By 1992 New York lowest risk-adjusted mortality rate of any state in nation & the most rapid rate of decline of any state with below-average mortality’
- Dranove et al. (2003)
  - ‘mandatory reporting mechanisms inevitably give providers incentive to decline to treat more difficult & complicated patients’

Transforming NHS in England: Lessons for United States?

- P4P
  - Why the interest?
  - Alternative: reputational damage?
  - Potential common problems?
- NHS in England
  - Transformation through reputational damage
  - Unintended & perverse outcomes
- Reflections

Total health care spend (% GDP)

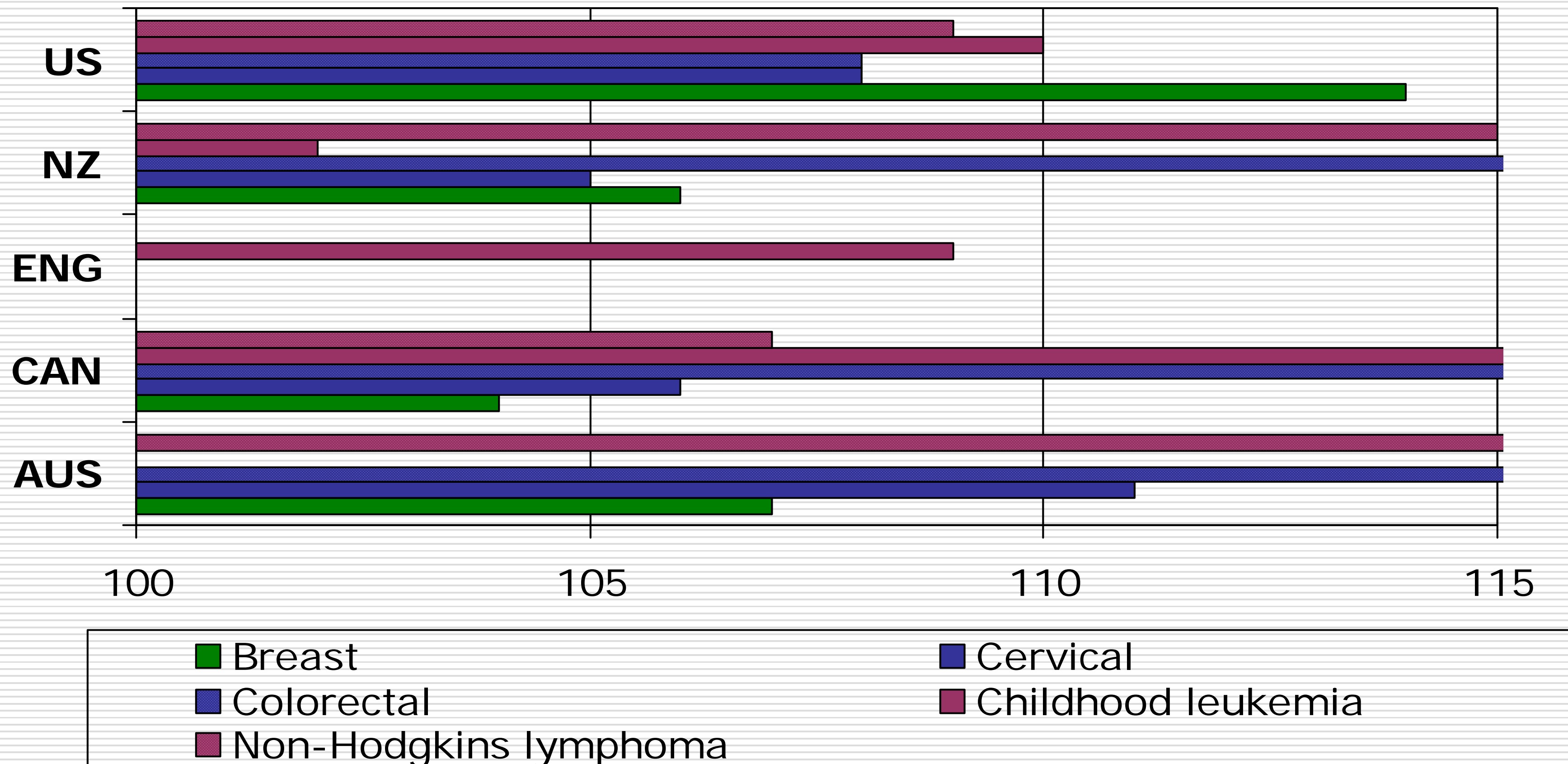


Source: OECD (2007)

NHS: Cheap, spartan, poor patient experience, long waiting times

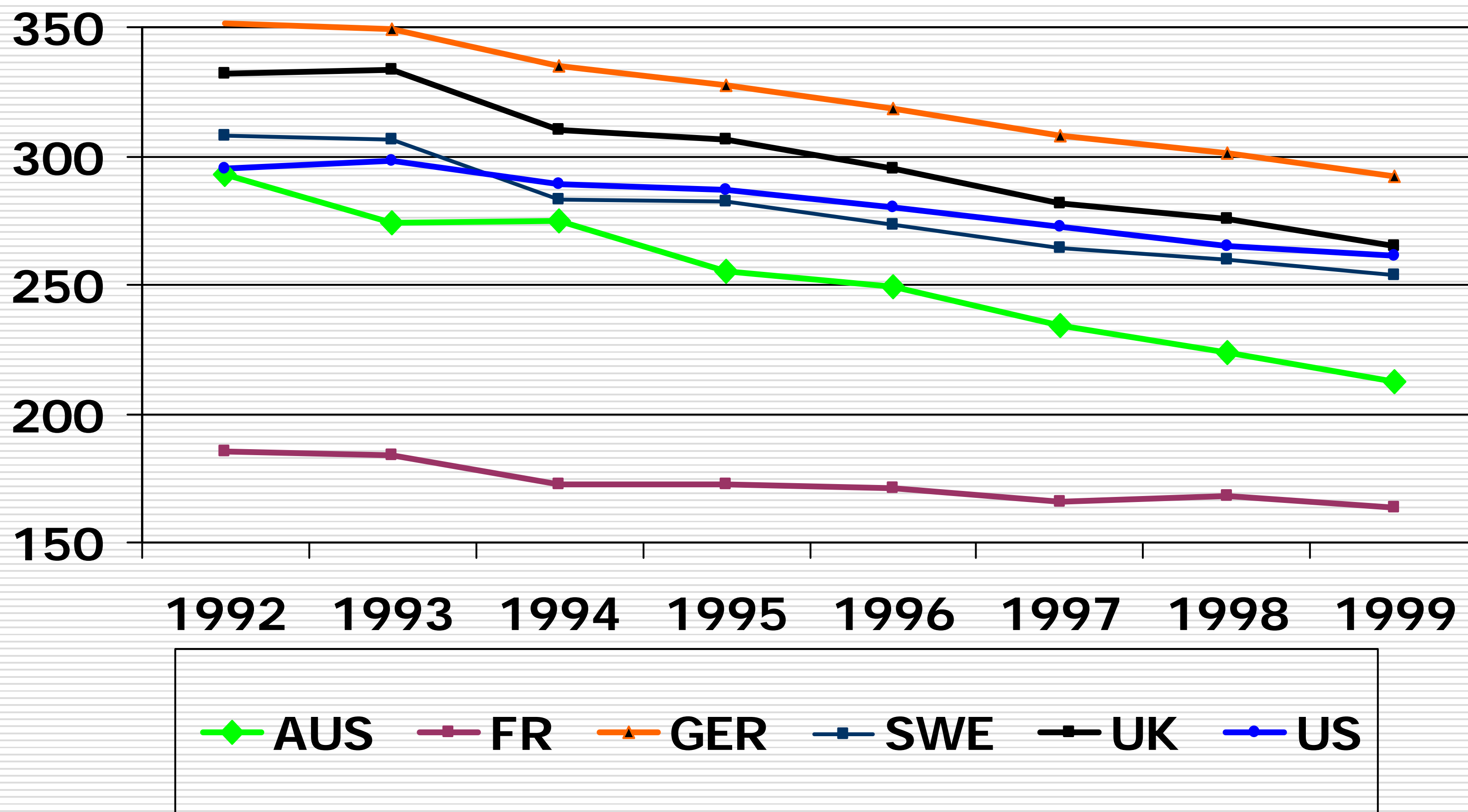
- Lord Winston (17 May 2000):
  - 87-year-old mother waited 13 hours in casualty before getting a bed in a mixed-sex ward drugs not given on time … missed meals … lying on the floor when morning staff came on.. caught an infection … ulcer on her leg … nothing unusual for NHS
- But resources directed at good clinical outcomes from dedicated clinical staff?

Survival rates cancers (diagnosed 1991-94: worst =100)



Source: Hussey et al (2004)

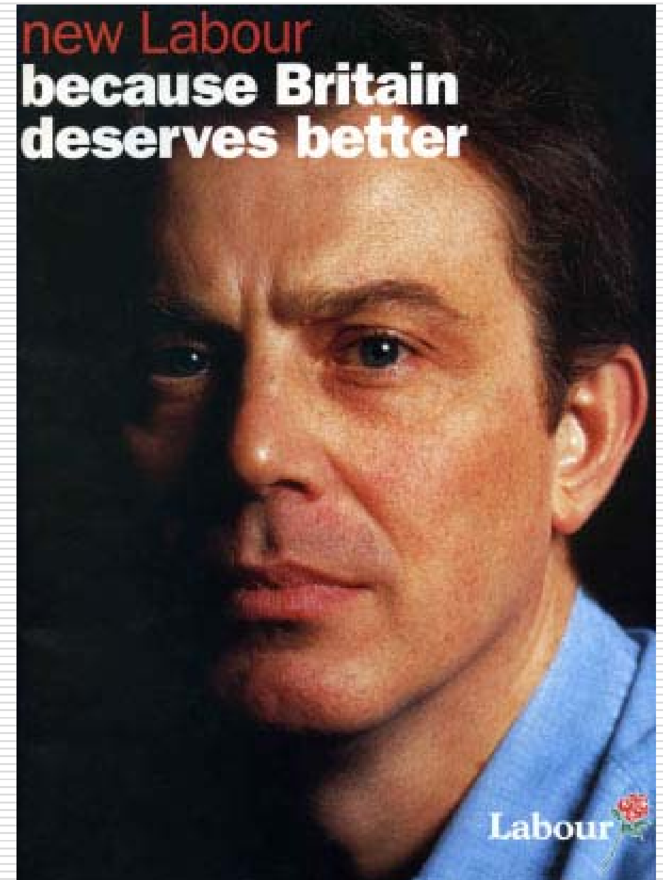
Circulatory disease (deaths / 100,000)



Source: OECD (2007)

1997 new Labour manifesto

- We will save the NHS
  - End the Tory internal market
  - End waiting for cancer surgery
- Five manifesto pledges
  - NHS: 100,000 people off waiting lists



Lord Winston's verdict on NHS under Labour (17 January 2000)

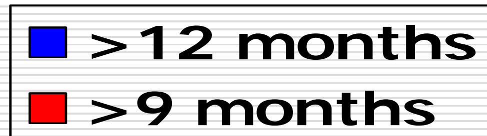
- NHS "much the worst in Europe"
- want NHS steadily deteriorate rationed & inferior for heart disease & cancer?
- gave categorical promises would abolish internal market … not done that

medical expert & Labour peer

Source: *The New Statesman*, interview, 17 January 2000

Failed manifesto promise on waiting times for elective admission (England)

Numbers waiting elective admissions ('000s)



200
150
100
50
0

1997 1998 1999 2000 2001 2002 2003 2004

Source: Chief Executive's Report to the NHS – Statistical Supplement (2004)

Response: 20 January 2000 Evidence-based policy making?

Most expensive breakfast in British history. Tony Blair: *We've decided to raise NHS spend to European average*

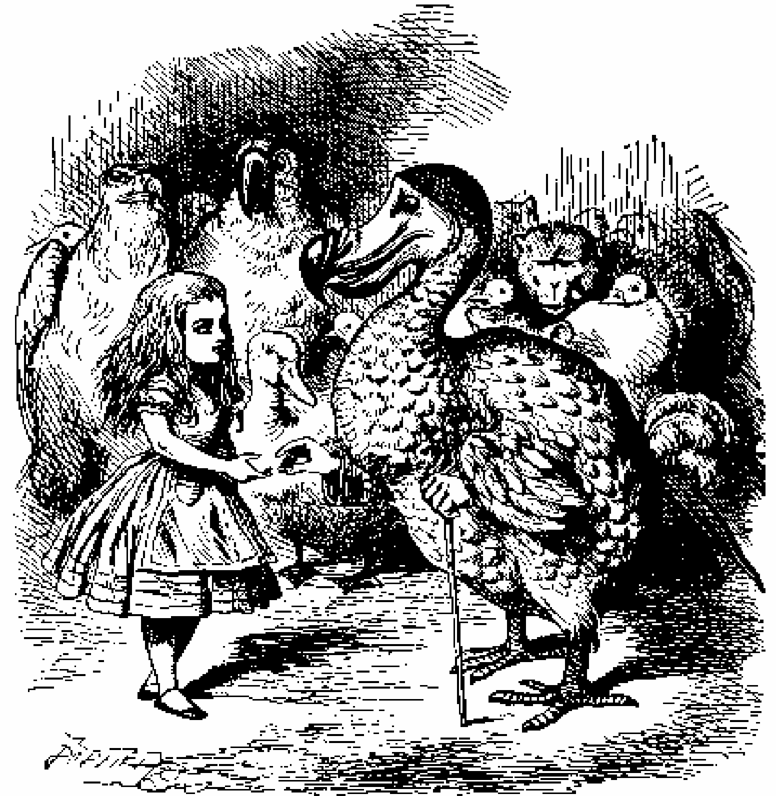
Response to crisis in practice: process & outcome

NHS spend (% GDP) before & after 20/1

10
8
6
4
2
0
1948 1958 1968 1978 1988 1998 2008

Failed systems

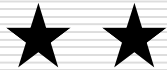
- No incentives
- Reward poor performers
- Markets



caucus race: everyone must have prizes

Performance assessment by star ratings: separating knights from knaves

Zero



9 Key targets

'balanced scorecard'

- patient surveys
- clinical outcomes
- capability & capacity

Natural experiment: performance assessment UK countries

	England	Scotland, Wales & Northern Ireland
ranking system	☑	☒
published & widely disseminated	☑	☒
easily understood by public	☑	☒
future reports	☑	☒

Reputational damage by star rating: ‘naming & shaming’

- Brutally redefined success & failure through national & local publicity: huge pressure on NHS to deliver
- Mannion et al (2005):
  - hospital staff highly engaged with information
  - effective communication & dissemination strategy
  - comprehensibility & appeal of such a stark & simple way of presenting data

Zero-rating: Traumatising impact on CEO, Board & staff

- 'devastating ...hit right down to the workforce – whereas bad reports usually hit senior management upwards ...nurses demanding changing rooms .. because being accosted in streets

2001: *the dirty dozen*

Source: Mannion et al (2005)

Zero-rated 2004

THE Sun

Thursday 21 October 2004

You make us sick! Scandal of Bosses running Britain's worst hospitals

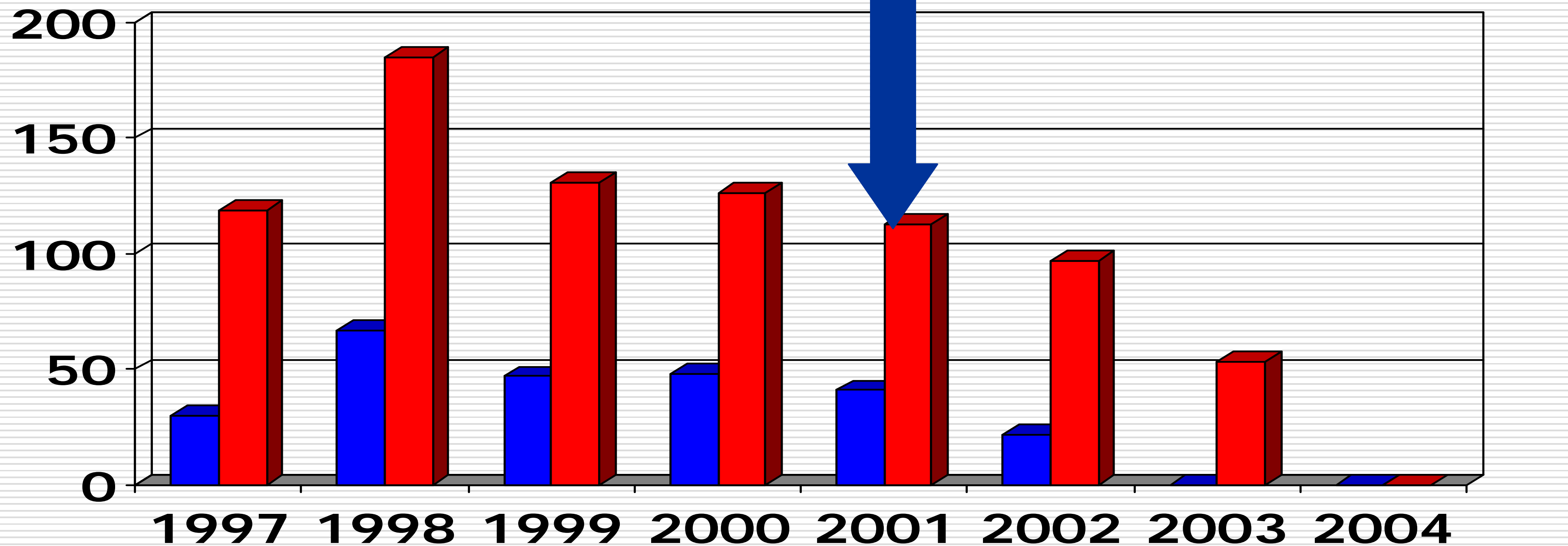
'squalid wards, long waiting times for treatment & rock-bottom staff morale'

if working in private sector 'sacked long ago'

Other effects

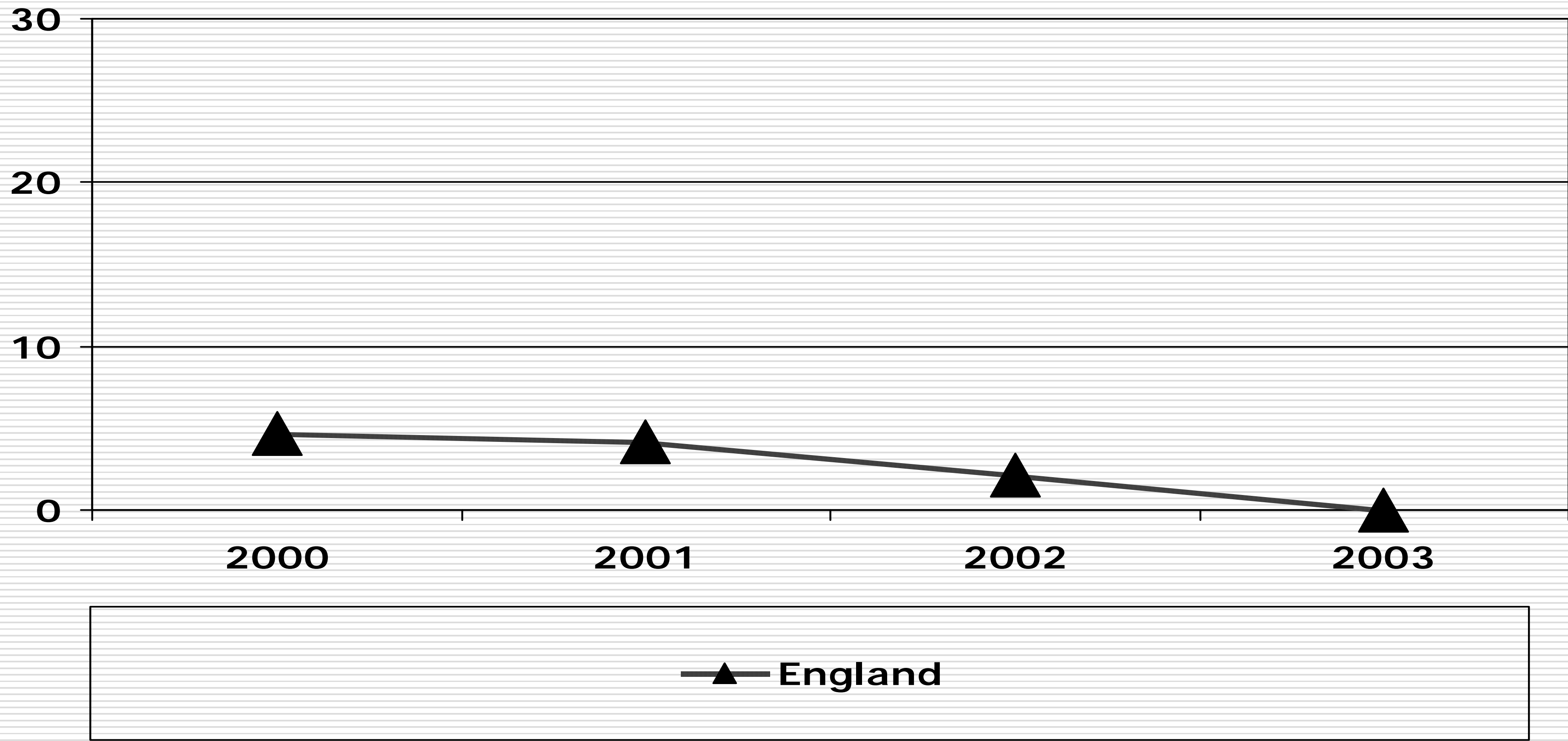
- one-star
  - working for a third class organisation
- recruitment
  - three-star: attractive & good organisation
  - Zero / one star: reluctant to join organisation publicly classified as under-performing
- ‘falling star’
  - shock to CEO & palpable effect on staff

Waiting times elective admission before & after star ratings (England)



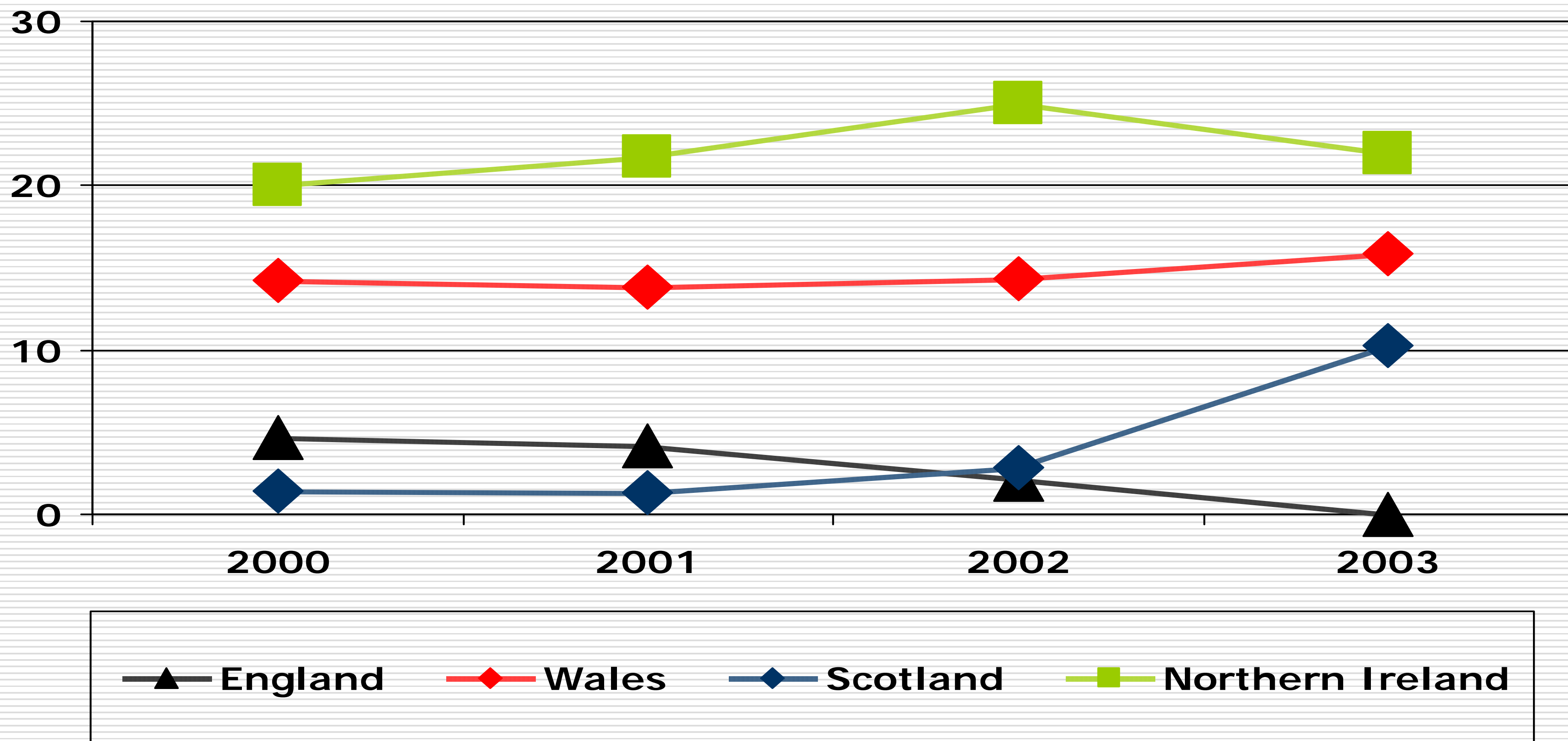
Source: Chief Executive's Report to the NHS – Statistical Supplement (2004)

Natural experiment: % patients waiting > 12 months hospital admission



Source: http://www.statistics.gov.uk National Health Service hospital waiting lists by region: Regional Trends 35, 36, 37 & 38

Natural experiment: % patients waiting > 12 months hospital admission



Source: http://www.statistics.gov.uk National Health Service hospital waiting lists by region: Regional Trends 35, 36, 37 & 38

Waiting time targets (December 2005)

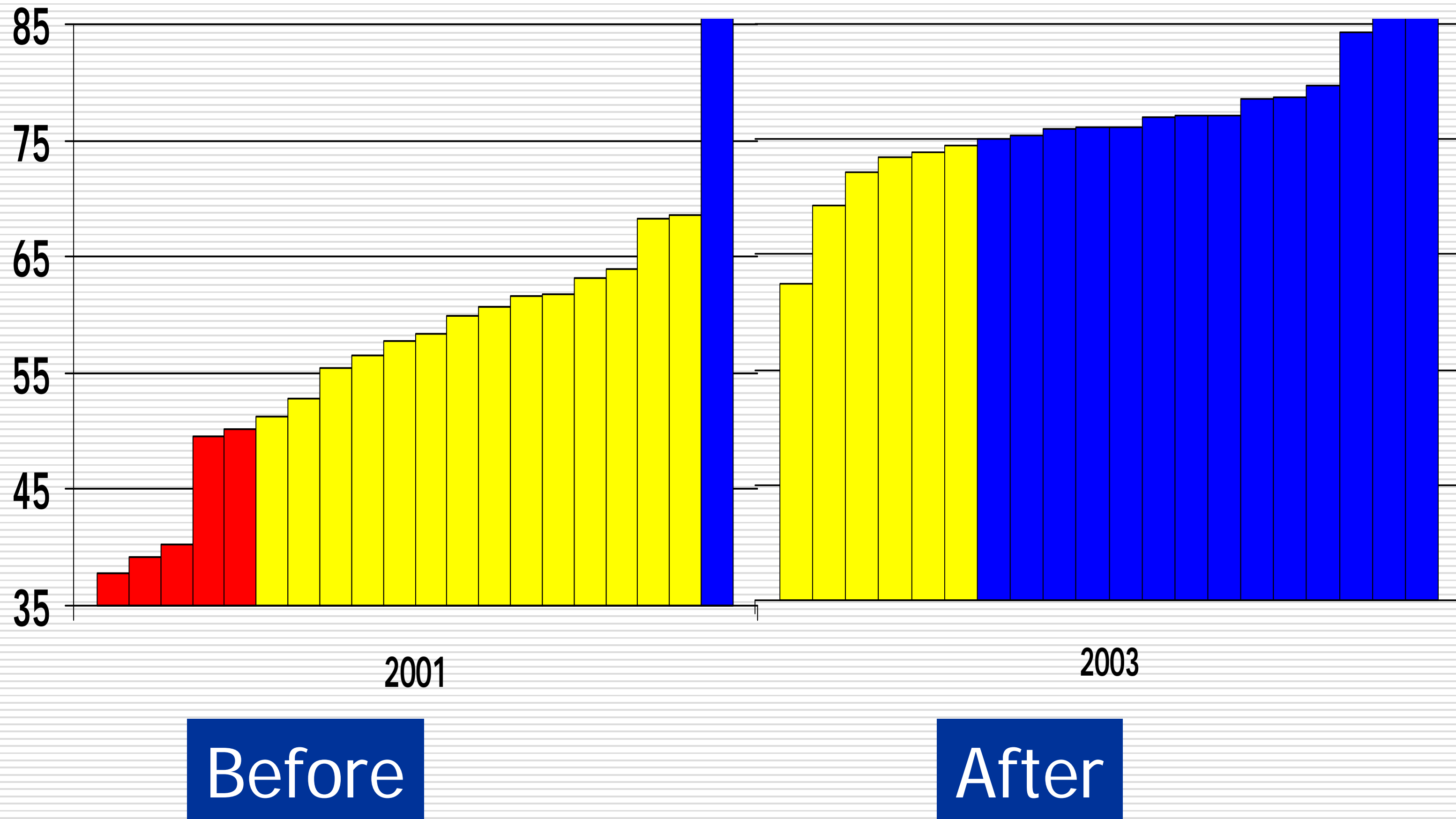
	England* (weeks)	Wales (weeks)	Scotland (weeks)
First outpatient appointment	13	78	26
Inpatient/day case treatment	26	78	26

* 2008 target: 18 weeks from GP referral to admission

Source: Auditor General for Wales (2005)

Ambulance response times life-threatening emergencies 75%< 8 minutes *England*

Target →

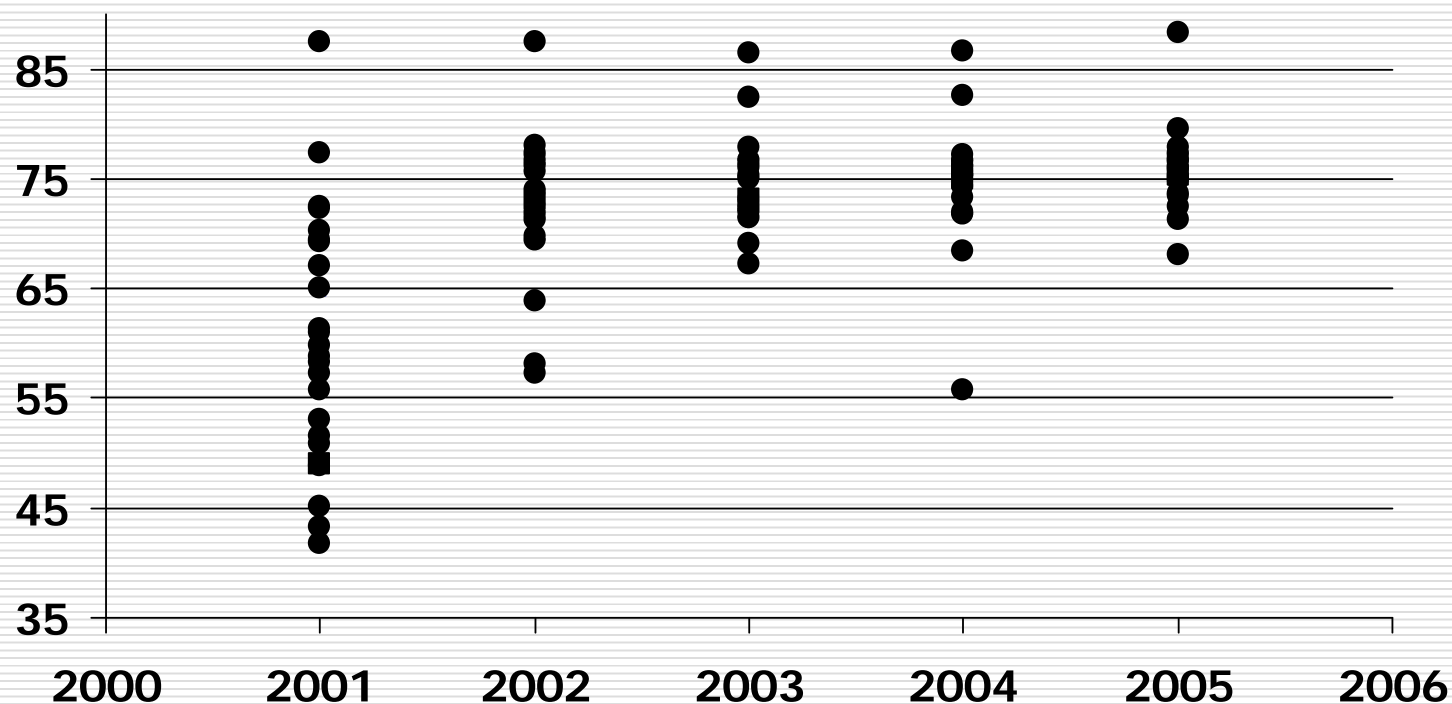


Ambulance response times life-threatening emergencies < 8 minutes *England*

% < 8 minutes

Target →

Zero-rated→



Star ratings published

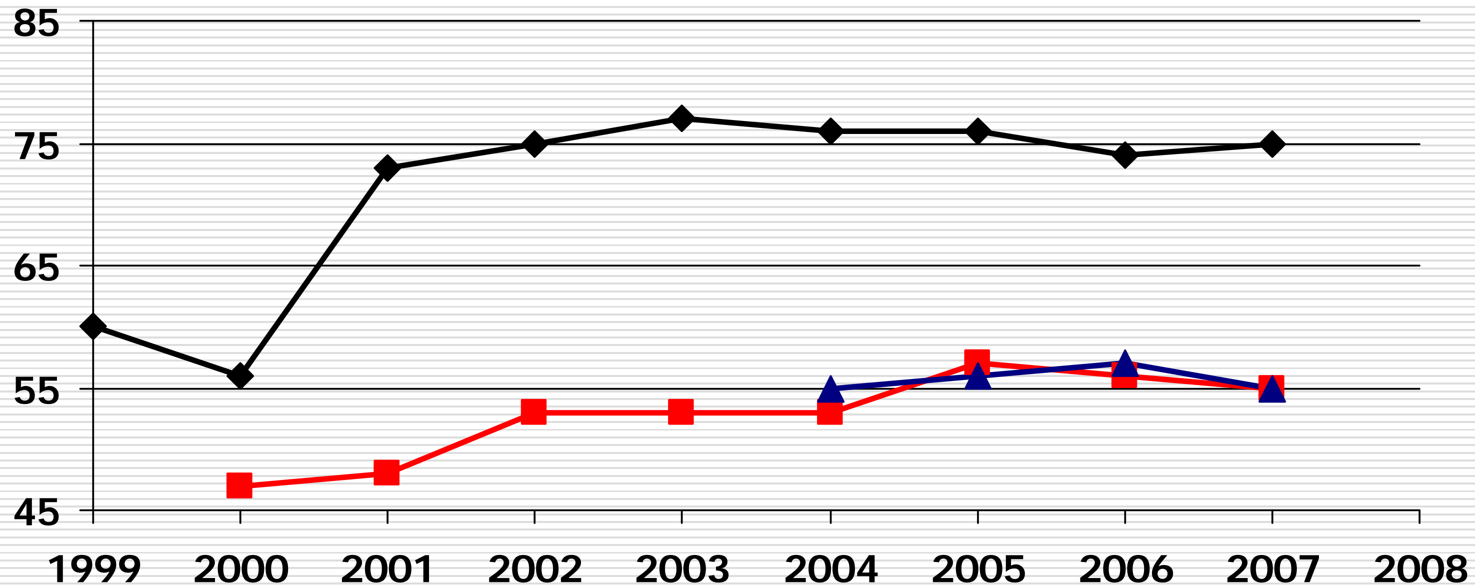
Source: http://www.dh.gov.uk/en/Publicationsandstatistics/Statistics/StatisticalWorkAreas/Statisticalhealthcare/DH_4086490

Ambulance response times life-threatening emergencies < 8 minutes UK countries

% < 8 minutes

Target →

Zero-rated →



Star ratings published

England

Wales

Scotland

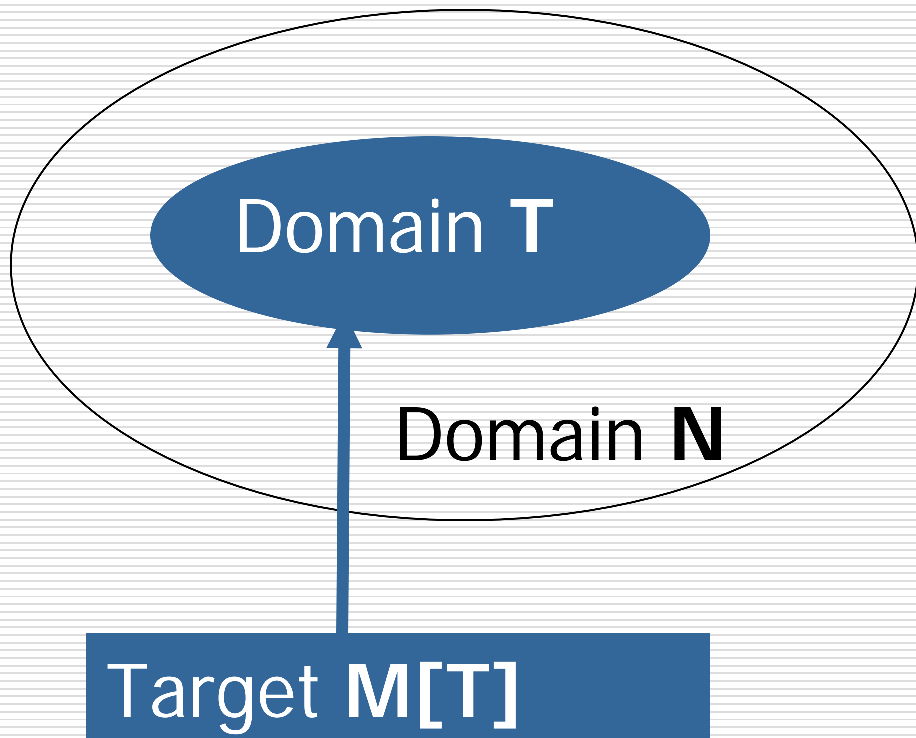
Impacts of different pathways in US & UK

	Wisconsin	New York	UK targets
Change	☒	☒	☒
Selection	☒	☒	☒
Reputation	☑	☑	☑

Transforming NHS in England: Lessons for United States?

- P4P
  - Why the interest?
  - Alternative: reputational damage?
  - Potential common problems?
- NHS in England
  - Transformation through reputational damage
  - Unintended & perverse outcomes
- Reflections

Problems from synecdoche & gaming: hit targets **M[T]**?



- Synecdoche (**N**)?
  - Worsen performance?
- Gaming **M[T]**:?
  - Hit the target & miss the point
  - 'Reporting error'
  - Trade-offs between targets

Synecdoche: Value for money

- More pay no improved performance
  - Cost £715 million (27% > original estimate) not yet delivering value for money to NHS & patients (March 2006)
- Recommendations
  - Before negotiating a new policy, Department should ensure that it has analysed sufficient contemporaneous evidence from relevant stakeholders
  - All possible scenarios for new policies should be fully financially modelled before they are implemented

Source: NAO (2007)

Problems of measurement & gaming in A&E (4 hour target)

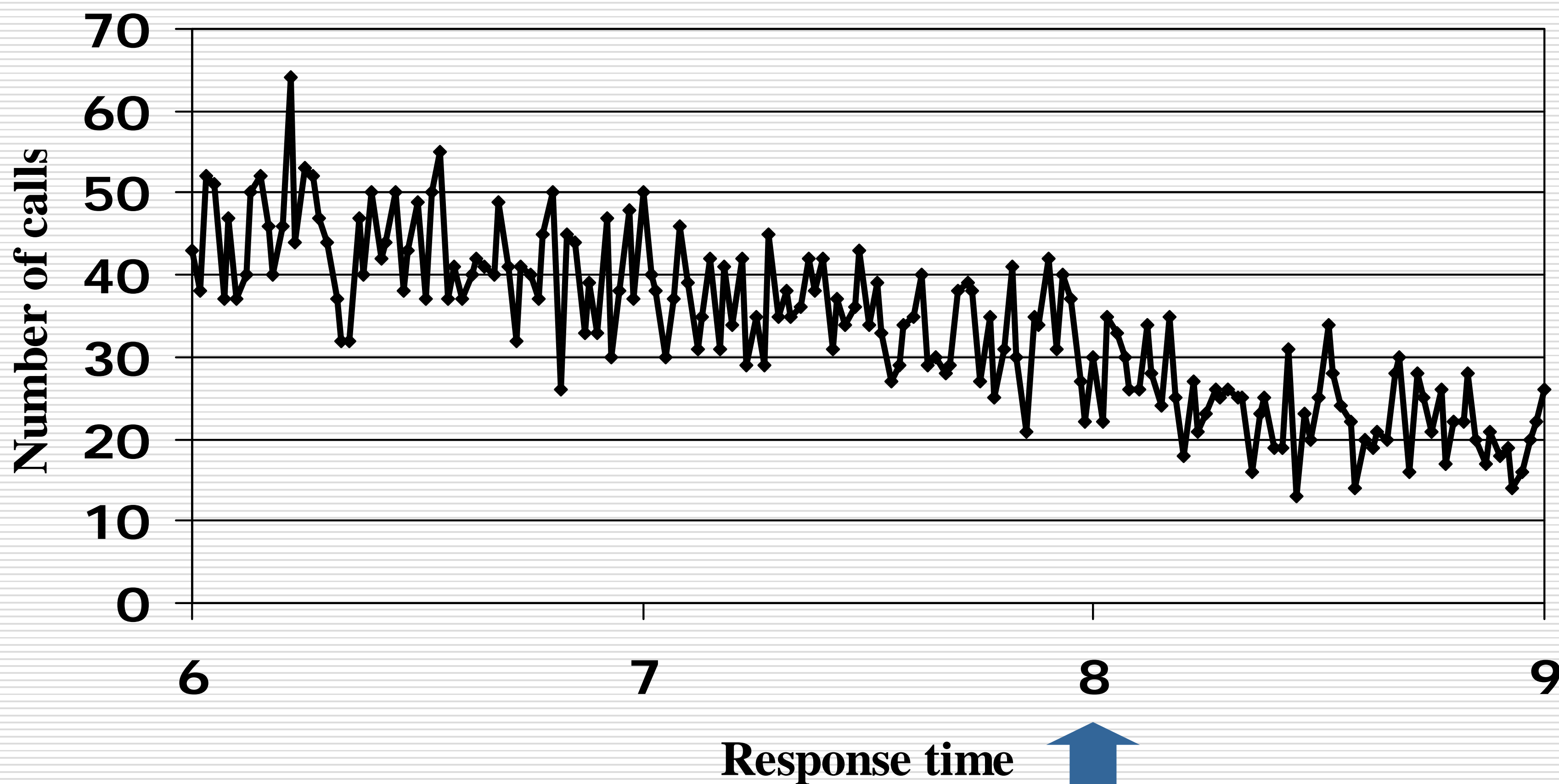
- Problems of measurement (2004)
  - Hospitals report 96%
  - Survey 55,000 patients: 77%

Problems of gaming

M[T] unreliable Ambulance response times

- Definition of life threatening emergency call?
  - *Fivefold* variation in percentage
- Measurement error?
  - 3 to *8* minutes!
- 'Reporting errors'
  - Manually 'correcting' times of calls taking > 8 minutes

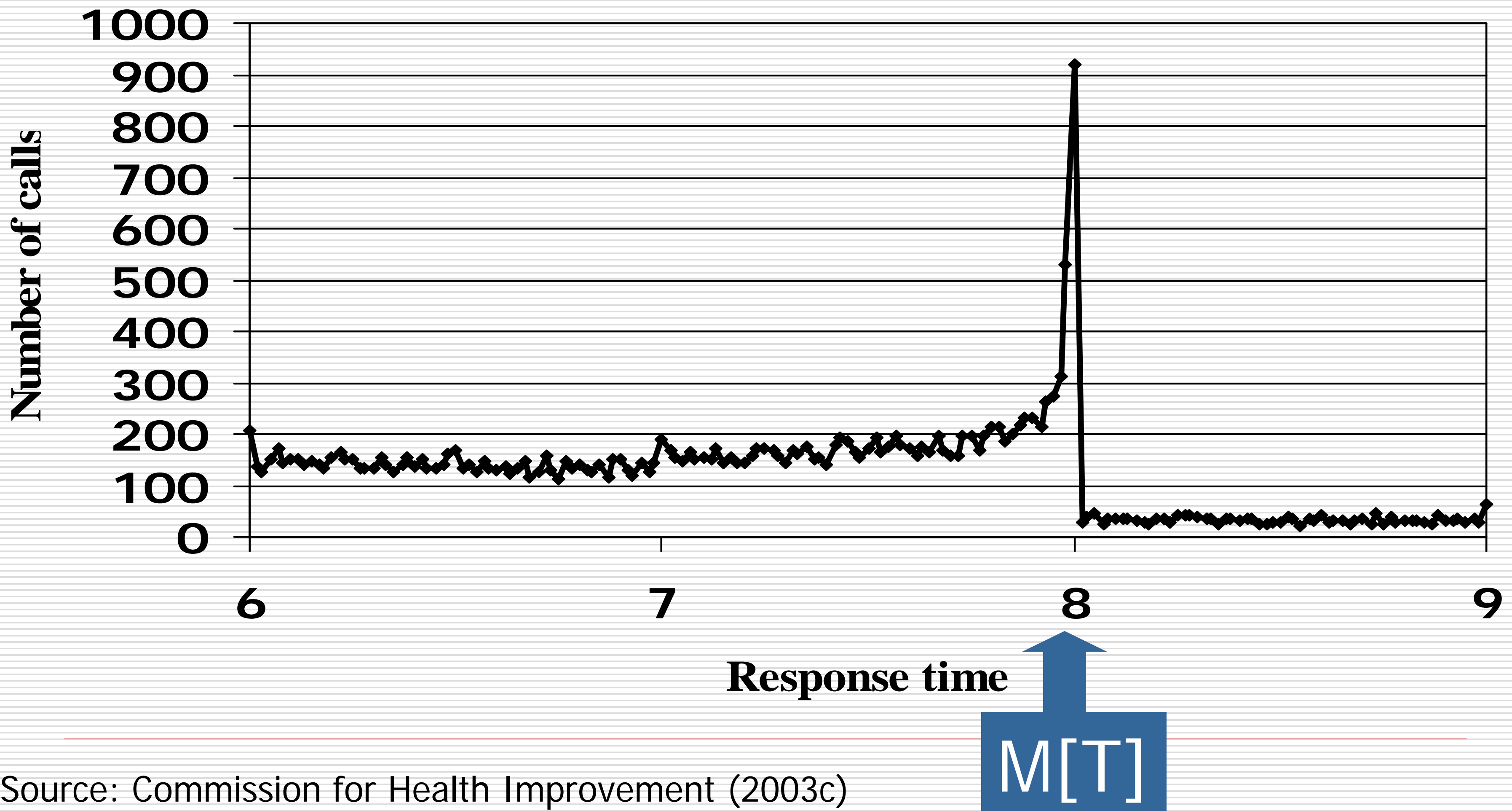
Ambulance response times ‘Uncorrected’: Noisy decline



M[T]

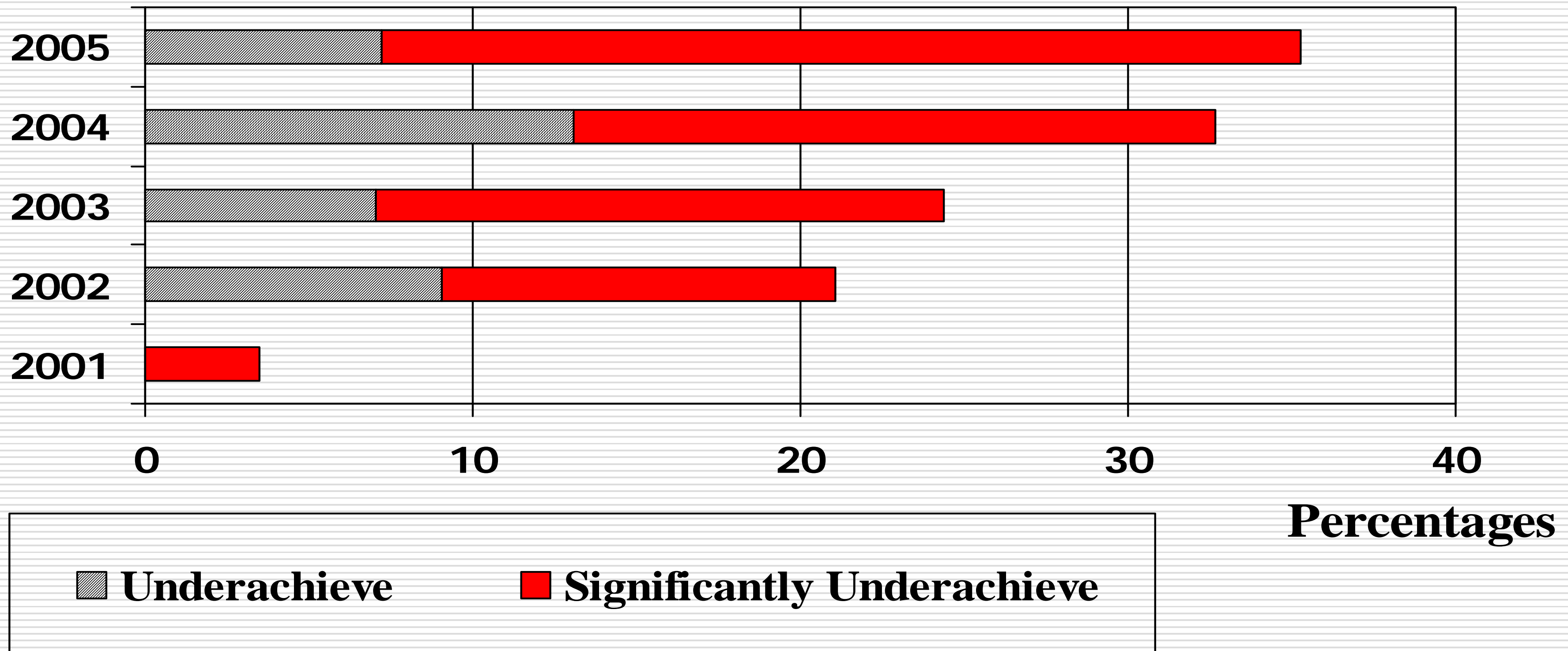
Source: Commission for Health Improvement (2003c) What CHI Has Found in: Ambulance Organisations.

Ambulance response times: an example of a 'Correction'



Source: Commission for Health Improvement (2003c)
What CHI Has Found in: Ambulance Organisations.

Trading-off wait time targets & financial deficits



Source: star ratings publications for each year

Consequences of hospital financial deficits

- 2005-06: NHS ‘£750m in the red’
- Ministers denied NHS chief Sir Nigel Crisp quit over figures

Sir Nigel announced he was stepping down, saying "not everything has gone well"

Transforming NHS in England: Lessons for United States?

- P4P
  - Why the interest?
  - Alternative: reputational damage?
  - Potential common problems?
- NHS in England
  - Transformation through reputational damage
  - Unintended & perverse outcomes
- Reflections

The wrath of Achilles: Public humiliation by Agamemnon

Briseis taken to Agamemnon

Reputation vs Financial incentives?

- NHS internal market → Patients & populations → managers
  - Patients & populations benefit from good & suffer twice from bad
- P4P
  - Why pay more for what people ought to do?
- Which is more powerful in public services
  - Relating perfomance to reputation / income?
- Evidence from another sector with asymmetry of information?

Evidence of impact of P4P where it really matters?

- $70m (stock, options and cash)
  - largest bonus ever given to Wall Street CEO (Christmas 2007)
- Wall Street cash bonuses (% of base salary)
  - top executives 40 to 100%
  - senior managers 15 to 30%
  - entry-level employees 10 to 20%
- Consequences for the real economy?
  - Sub-prime mortgage market

Lloyd Blankfein

Goldman Sachs

Transforming NHS in England: Lessons for United States?

Thank you

LSE

References

- Auditor General for Wales (2005) NHS waiting times in Wales. Volume 2 - Tackling the problem. <http://www.wao.gov.uk/reportsandpublications/2005.asp>
- Bevan (2006) Setting Targets for Health Care Performance: lessons from a case study of the English NHS. National Institute Economic Review, 197, 67-79.
- Bevan & Hood (2006) Have targets improved performance in the English NHS? British Medical Journal, 332, 419-422.
- Bevan & Hood C (2006) What's Measured is What Matters: Targets and Gaming in the English Public Health Care System. Public Administration 84, 3, 517-38.
- Chassin (2002) Achieving and sustaining improved quality: Lessons from New York State and cardiac surgery. Health Affairs, 21(4), 40-51.

References

- CRAG (2002) Clinical Outcome Indicators. A report from the Clinical Outcomes Working Group. <www.show.scot.nhs.uk/crag>
- Fung et al (2008) Systematic review: The evidence that publishing patient care performance data improves quality of care. Annals of Internal Medicine, 148, 111-123.
- Hibbard (2008) What can we say about the impact of public reporting? Inconsistent execution yields variable results. Annals of Internal Medicine, 148, 160 - 161.
- Hibbard et al (2003) Does publicizing hospital performance stimulate quality improvement efforts? Health Affairs, 22 (2), 84-94.
- Hibbard et al (2005b) Hospital performance reports: impact on quality, market share, and reputation. Health Affairs, 24(4), 1150-60.

References

- Mannion et al (2005a) Impact of star performance ratings in English acute hospital organisations. Journal of Health Services Research & Policy, 10 (1), 18–24.
- Mannion & Goddard (2001) Impact of published clinical outcomes data: case study in NHS hospital organisations. BMJ, 323, 260 – 263
- Mannion & Goddard (2003) Public disclosure of comparative clinical performance data: lessons from the Scottish experience. Journal of Evaluation in Clinical Practice, 9 (2), 277–286
- Marshall et al (2000) The Public release of performance data: What do we expect to gain? A review of the evidence. Journal of the American Medical Association, 283(14), 1866-1874.