Department of Statistics public lecture

Trying to Quantify Uncertainty

Professor David Spiegelhalter
Winton Professor of the Public Understanding of Risk, University of Cambridge and senior scientist, MRC Biostatistics Unit

Professor Anthony Atkinson
Chair, LSE

Trying to quantifying uncertainty

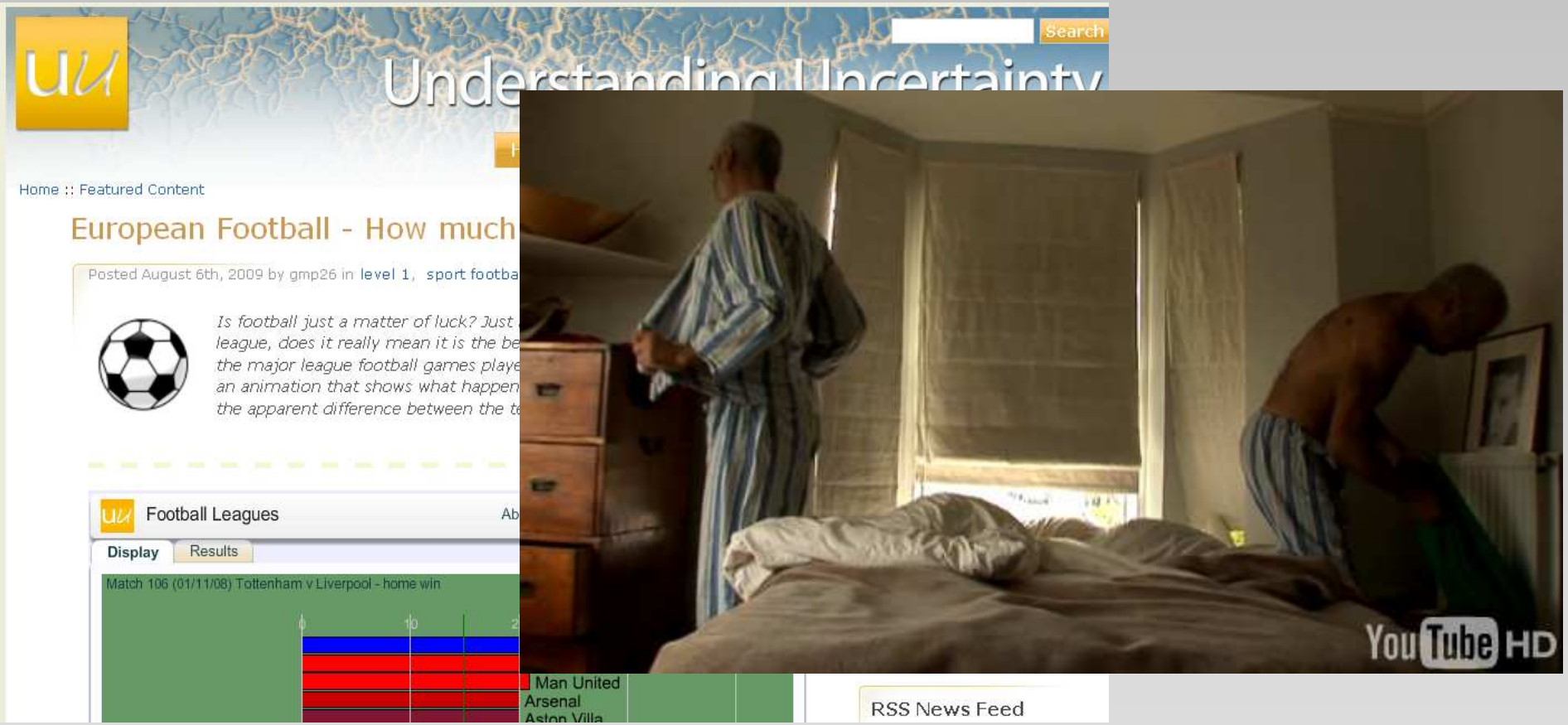
David Spiegelhalter

winton professor of the
public understanding of risk,
university of cambridge

LSE, November 2010

With thanks to Mike Pearson, Ian Short, etc

www.understandinguncertainty.org



'*Professor Risk*' on Youtube

so surprised?

Problems

Why not send us your solutions?

Sociable Cards

Stage: 3 Challenge Level: ★

Move your counters through this snake of cards and see how far you can go. Are you surprised by where you end up?

What Does Random Look Like?

Stage: 3 Challenge Level: ★

Engage in a little mathematical detective work to see if you can spot the fakes.

At Least One...

Stage: 3 and 4 Challenge Level: ★

Imagine flipping a coin a number of times. Can you work out the probability you will get a head on at least one of the flips?

Mathsland National Lottery

Stage: 3 and 4 Challenge Level: ★

Can you work out the probability of winning the Mathsland National Lottery? Try our simulator to test out your ideas.

Featured Solution

Your Number Was...

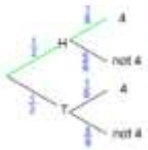
We received a variety of good strategies for solving this problem.

Go to last month's problems to see more solutions.

Articles & Games

Why Do People Find Probability Unintuitive and Difficult?

Uncertain about the likelihood of unexpected events? You are not alone!



An Introduction to Tree Diagrams

This article explains how tree diagrams are constructed and helps you to understand how they can be used to calculate probabilities.

Lottery Simulator

Use this animation to experiment with lotteries. Choose how many balls to match, how many are in the carousel, and how many draws to make at once.

Why try to quantify uncertainty?

People should have an idea of the magnitudes of

- how likely something is
- how good or bad it might be

Risk communication using numbers?

A recent population survey asked

- *Which of the following numbers represents the biggest risk of getting a disease:*

  1 in 100, 1 in 1000, 1 in 10 ?

% with incorrect answer:

Germany 28%

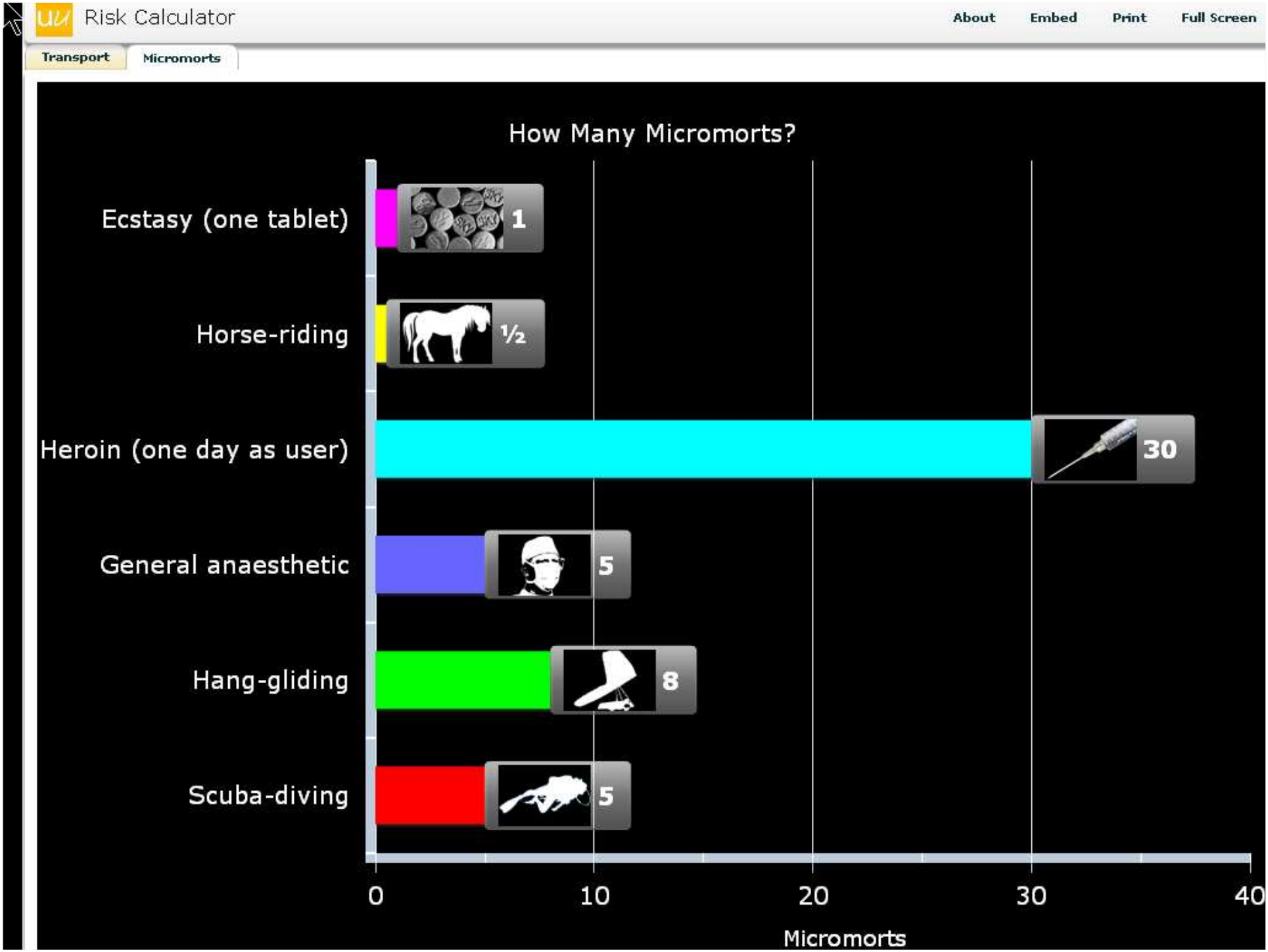
USA 25%

Statistical Numeracy for Health

A Cross-cultural Comparison With Probabilistic National Samples

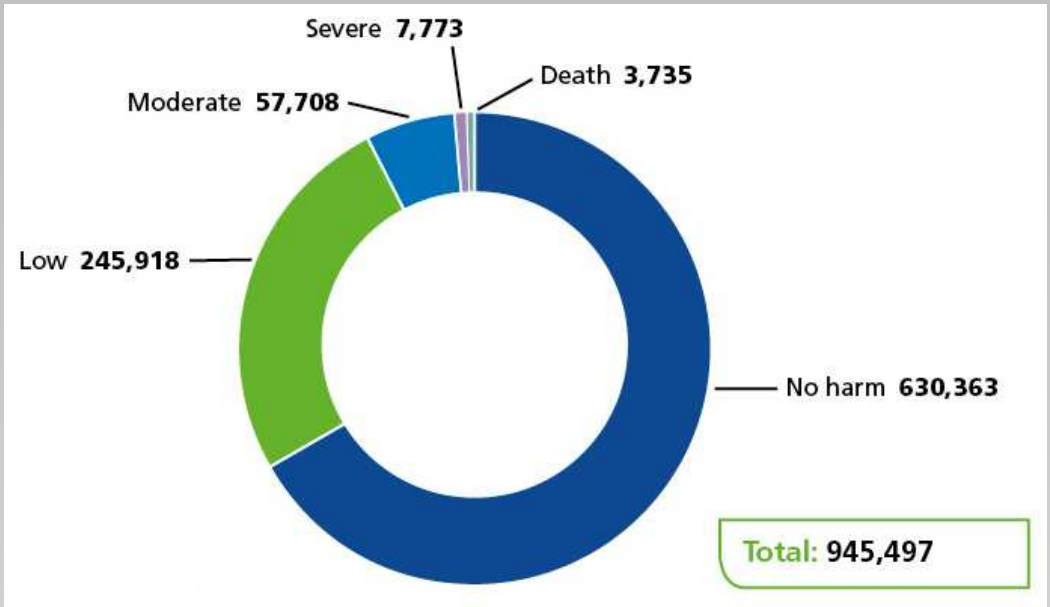
Mirta Galesic, PhD; Rocio Garcia-Retamero, PhD

Arch Intern Med. 2010;170(5):462-468


Risk Calculator
About
Embed
Print
Full Screen
Transport
Micromorts
How Many Micromorts?
Ecstasy (one tablet)
1
Horse-riding
½
Heroin (one day as user)
30
General anaesthetic
5
Hang-gliding
8
Scuba-diving
5
0
10
20
30
40
Micromorts

Going into hospital?

Safety incidents in English hospitals reported to NHS National Patient Safety Agency July 2008 to June 2009



Deaths: **3735**

Average number of beds occupied each day in English hospitals: **135,000**

Average Micromorts per day: **75**

War or peace?

UK Deaths in Afghanistan:

12th July to 19th September 2010:

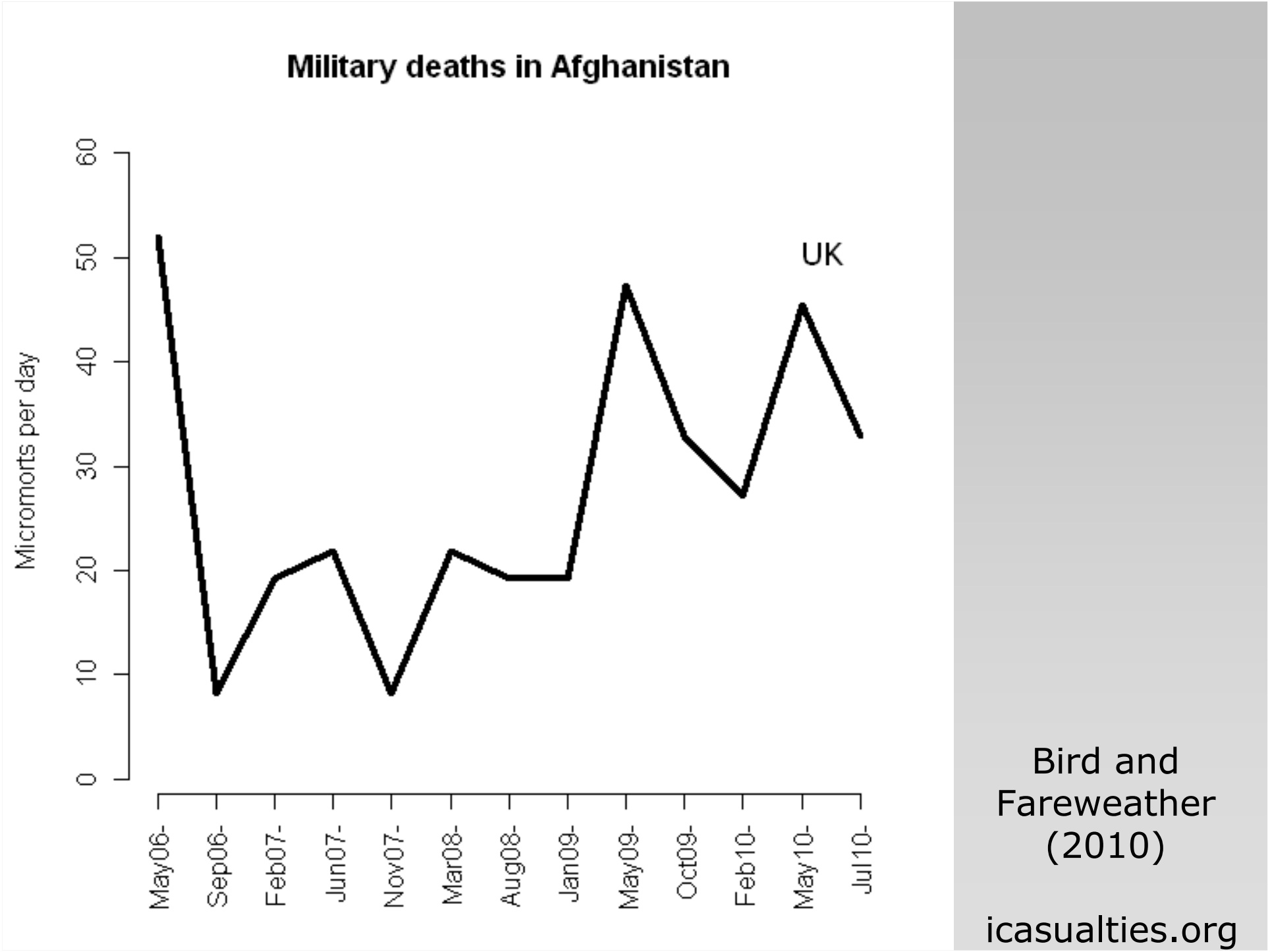
23 deaths

Average service-personnel per day:

10,000

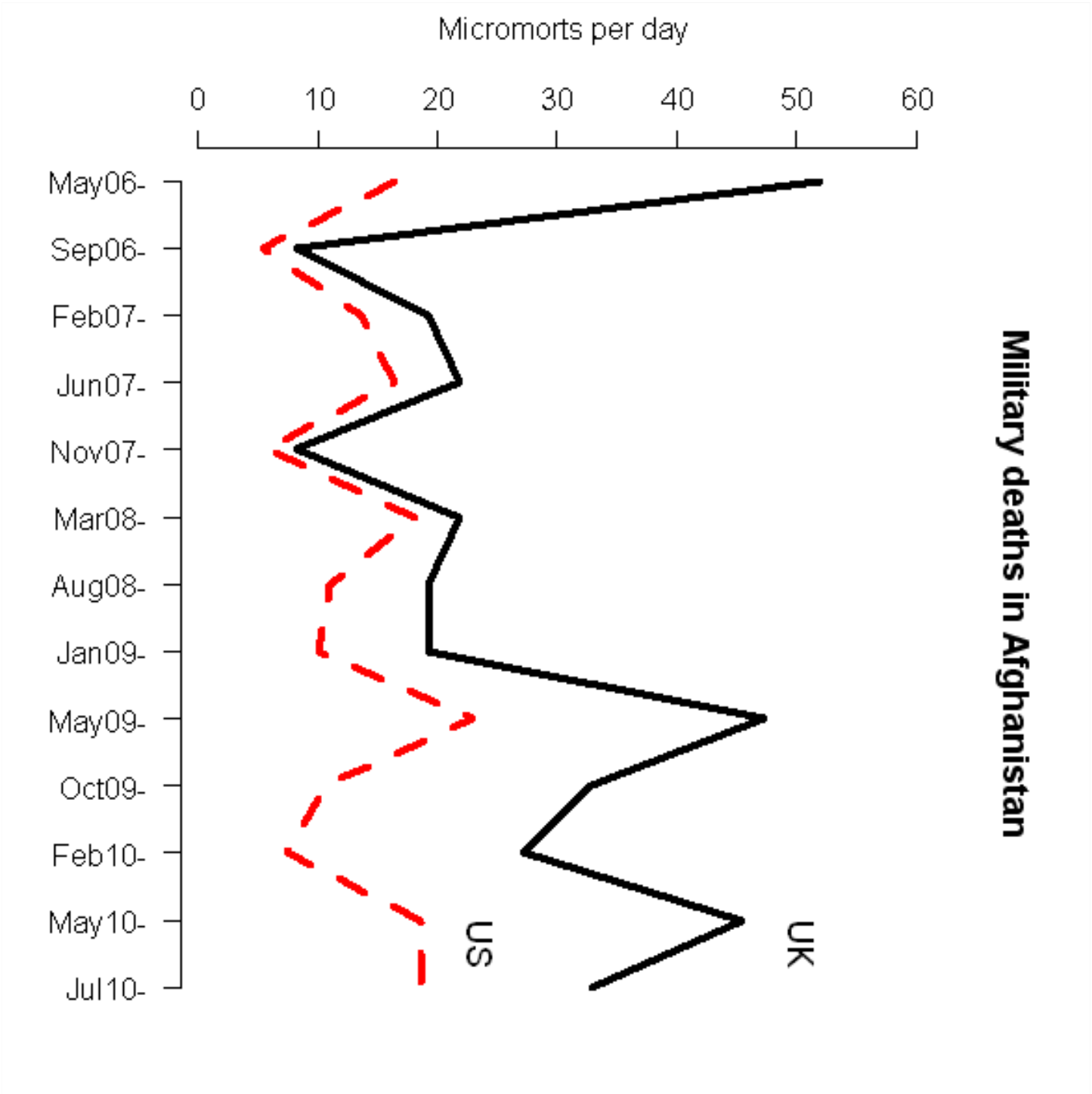
Average micromorts per day:

33



Bird and Fareweather (2010)

icasualties.org


Micromorts per day
0
10
20
30
40
50
60
May06-
Sep06-
Feb07-
Jun07-
Nov07-
Mar08-
Aug08-
Jan09-
May09-
Oct09-
Feb10-
May10-
Jul10-
US
UK
Military deaths in Afghanistan

Expressing benefits and harms

Cochrane Collaboration "*Summary of findings table*" Uses GRADE scale for quality of evidence (in addition to confidence interval)

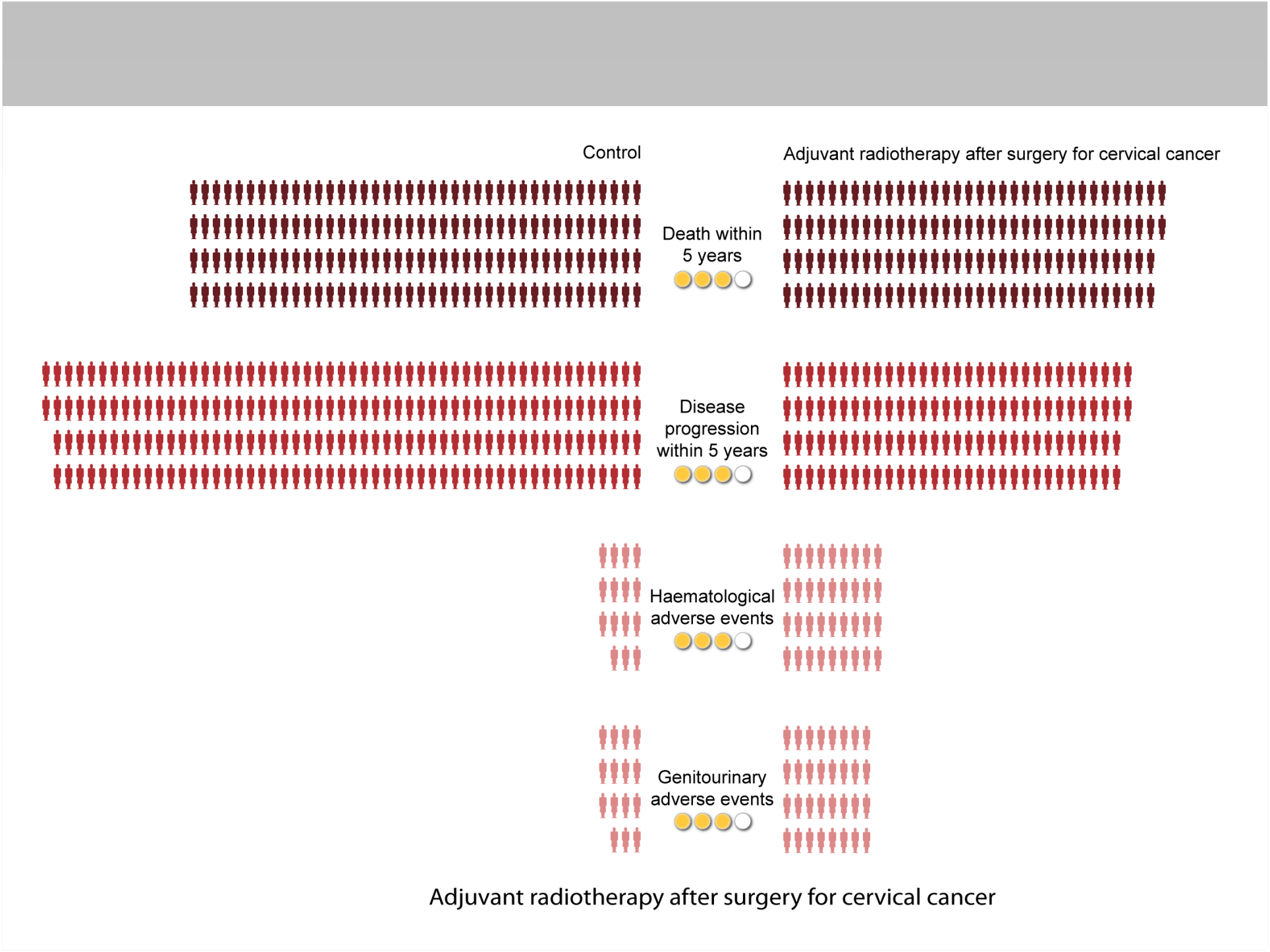
SUMMARY OF FINDINGS FOR THE MAIN COMPARISON *[Explanation]*

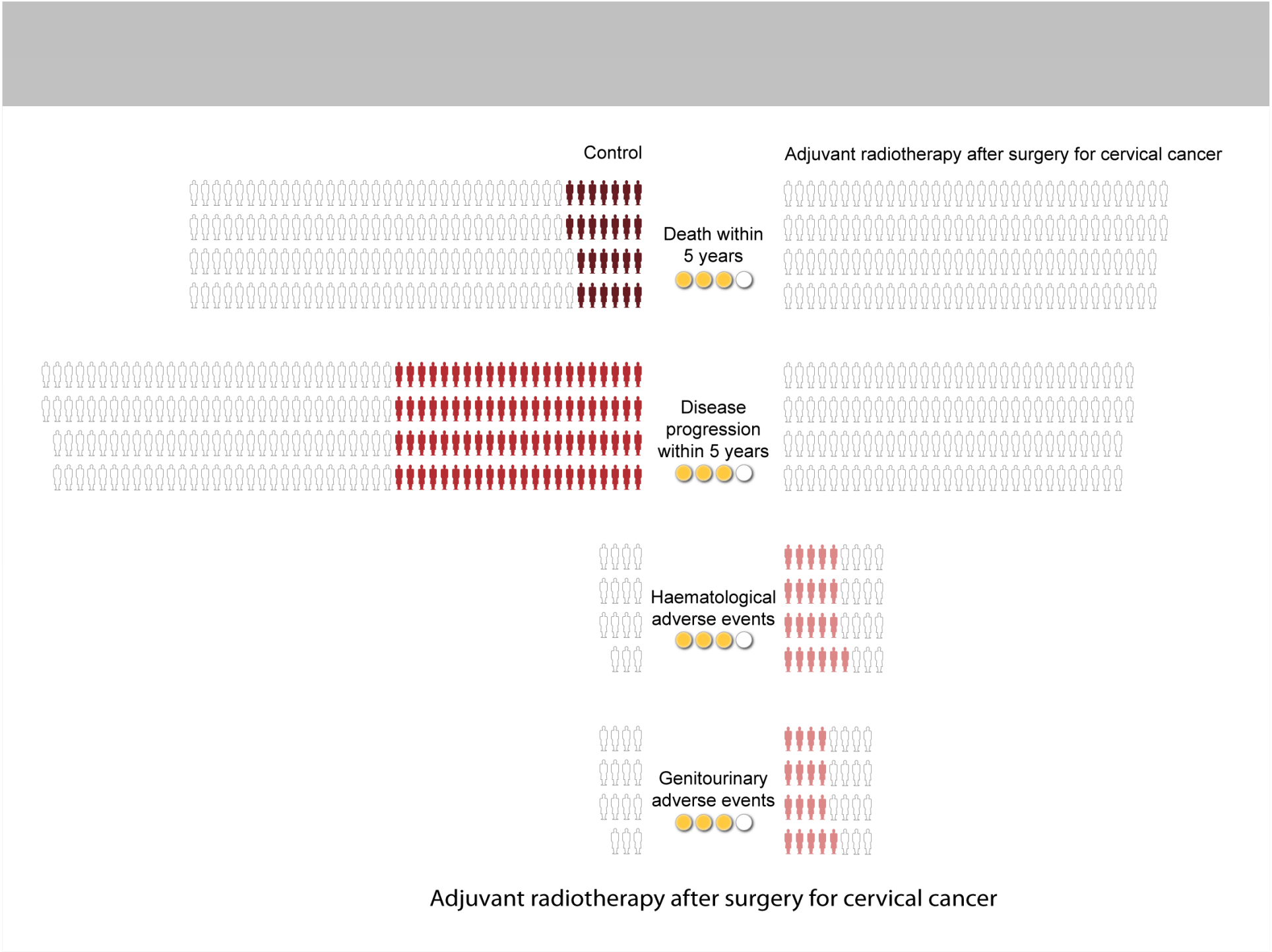
Adjuvant radiotherapy after surgery for cervical cancer

Patient or population: patients with early stage cervical cancer (FIGO stages IB1, IB2 or IIA)
Settings: Inpatient or outpatient
Intervention: Adjuvant radiotherapy after surgery

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	C
	Assumed risk	Corresponding risk				
	Control	Adjuvant radiotherapy after surgery				
Death within 5 years	Study population		RR 0.84 (0.3 to 2.36)	397 (2 studies)	⊕⊕⊕○ moderate[1]	
	160 per 1000	134 per 1000 (48 to 378)				
	Medium risk population					
	124 per 1000	104 per 1000 (37 to 293)				
Disease progression within 5 years	Study population		RR 0.58 (0.37 to 0.91)	397 (2 studies)	⊕⊕⊕○ moderate[2,3]	
	210 per 1000	122 per 1000 (78 to 191)				
	Medium risk population					
	164 per 1000	95 per 1000 (61 to 149)				

One thousand


Control
Adjuvant radiotherapy after surgery for cervical cancer
Death within 5 years
Disease progression within 5 years
Haematological adverse events
Genitourinary adverse events
Adjuvant radiotherapy after surgery for cervical cancer


Control
Adjuvant radiotherapy after surgery for cervical cancer
Death within 5 years
Disease progression within 5 years
Haematological adverse events
Genitourinary adverse events
Adjuvant radiotherapy after surgery for cervical cancer

BMJ

RESEARCH

Unintended effects of statins in men and women in England and Wales: population based cohort study using the QResearch database

Julia Hippisley-Cox, professor of clinical epidemiology and general practice, Carol Coupland, associate professor in medical statistics

Division of Primary Care, University Park, Nottingham NG2 7RD

Correspondence to: J Hippisley-Cox
julia.hippisley-cox@ntlworld.com

Cite this as: *BMJ* 2010;340:c2197
doi:10.1136/bmj.c2197

ABSTRACT

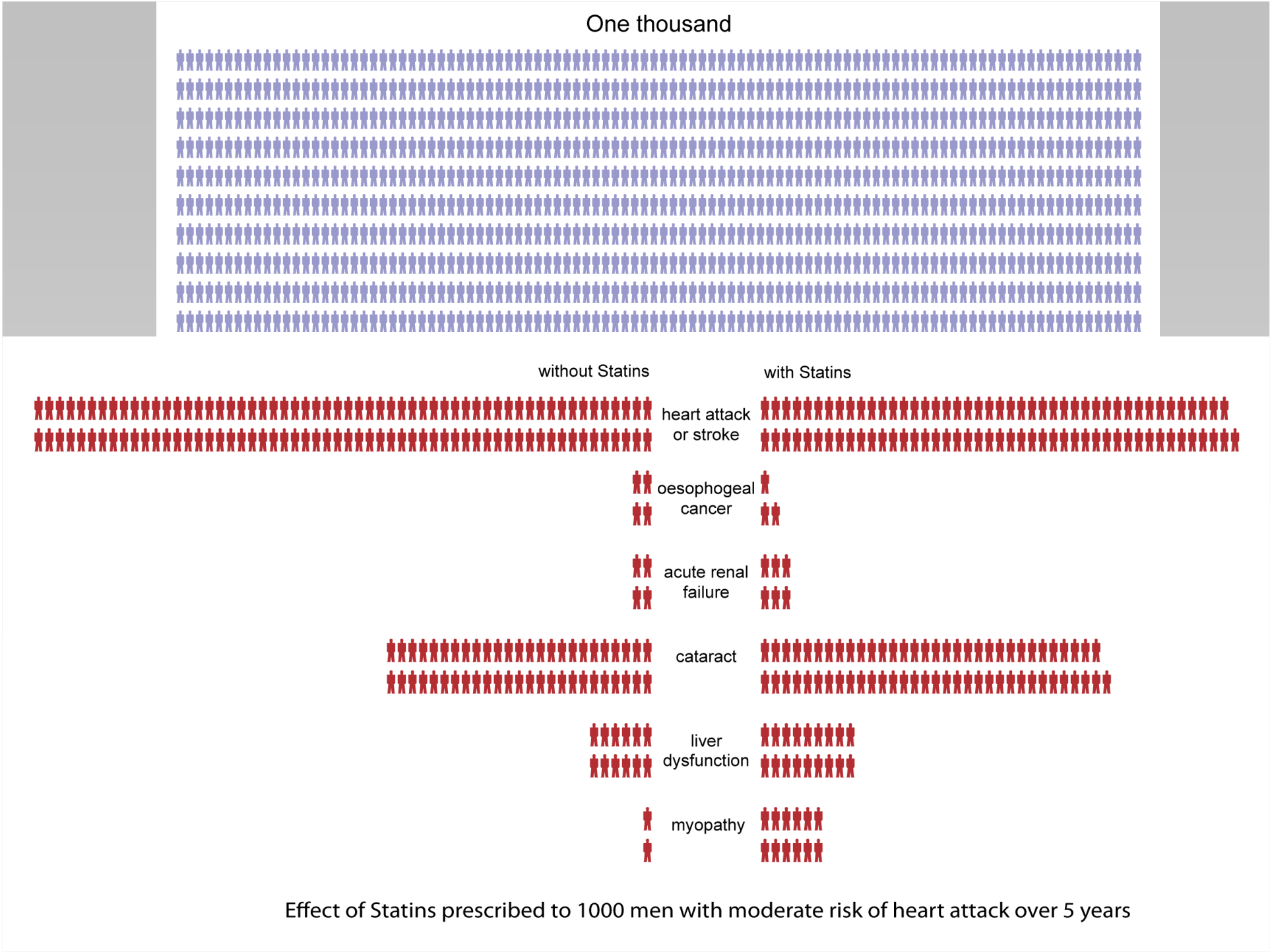
Objective To quantify the unintended effects of statins according to type, dose, and duration of use.

Design Prospective open cohort study using routinely collected data.

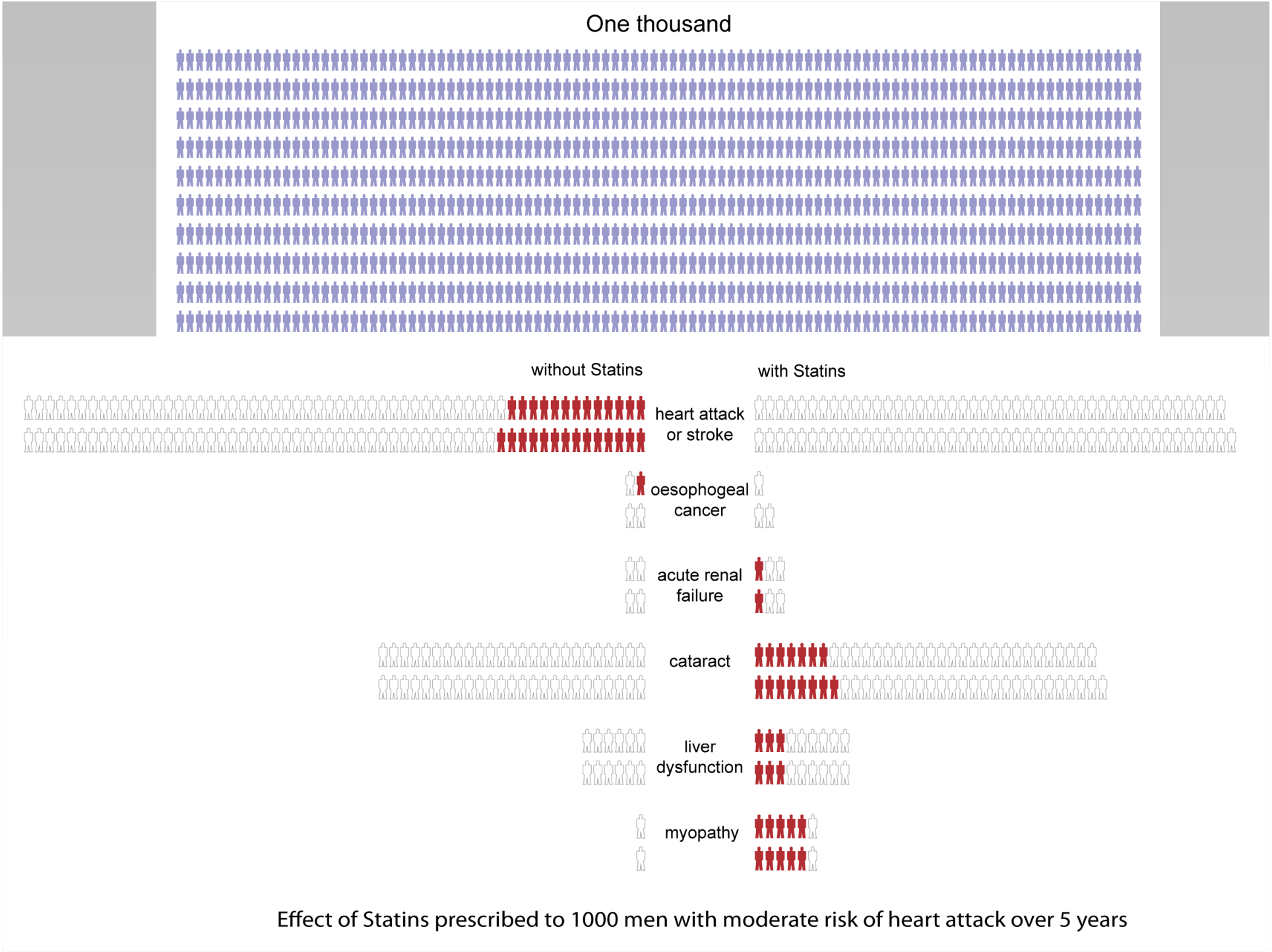
Setting 368 general practices in England and Wales supplying data to the QResearch database.

Participants 2 004 692 patients aged 30-84 years of whom 225 922 (10.7%) were new users of statins:

of acute renal failure returned to normal within 1-3 years in men and women, and liver dysfunction within 1-3 years in women and from three years in men. Based on the 20% threshold for cardiovascular risk, for women the NNT with any statin to prevent one case of cardiovascular disease over five years was 37 (95% confidence interval 27 to 64) and for oesophageal cancer was 1266 (850 to 3460) and for men the respective values were 33 (24 to 57) and 1082 (711 to 2807). In women the NNH for an additional case of



Effect of Statins prescribed to 1000 men with moderate risk of heart attack over 5 years



Effect of Statins prescribed to 1000 men with moderate risk of heart attack over 5 years

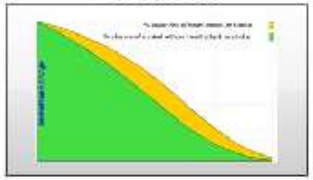
- Recently I went to see my GP …
- He told me I had a 12% chance of a heart attack or stroke over the next 10 years
- But I could take statins!

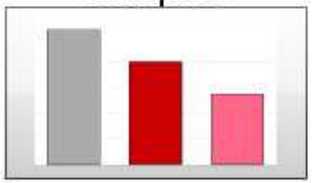
Cardiovascular Risk Assessment

Full Scre

Profile | Heart Age | Outlook | Outcomes | Compare







Profile

Date of Birth (DD MM YYYY): 16 8 1953

Gender: (•) male ( ) female

Total Cholesterol: 5.20 mmol/L

HDL Cholesterol: 1.45 mmol/L

Systolic Blood Pressure: 130 mm Hg

Tick if you have received blood pressure treatment [ ]

Tick if you suffer from diabetes [ ]

Tick if you currently smoke [ ]

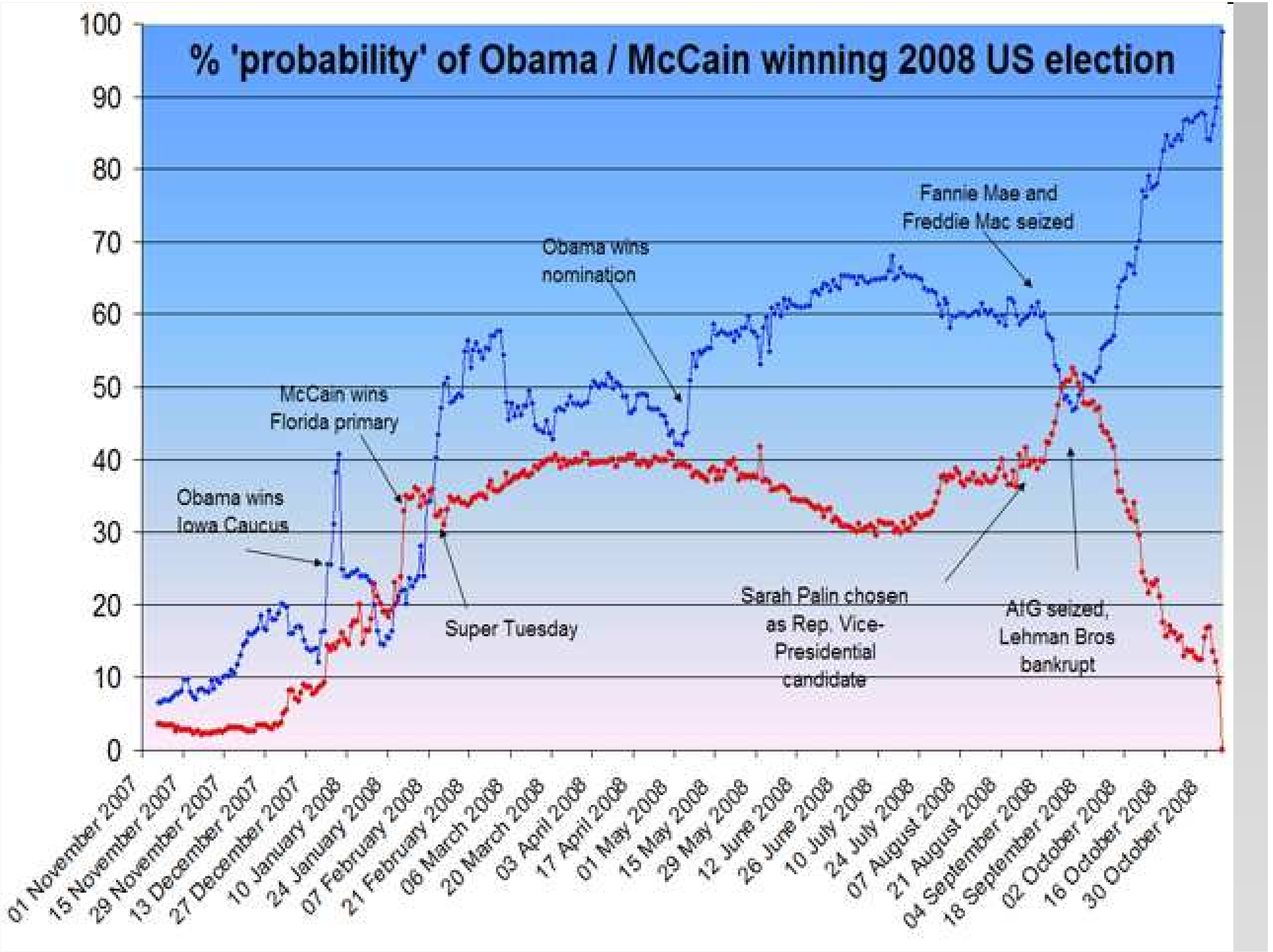
Tick if you used to smoke [ ]

Tick if physically active [x]

Save | Load | Next

Using history

- History is not always a reliable guide


% 'probability' of Obama / McCain winning 2008 US election
100
90
80
70
60
50
40
30
20
10
0
Obama wins Iowa Caucus
McCain wins Florida primary
Super Tuesday
Obama wins nomination
Sarah Palin chosen as Rep. Vice-Presidential candidate
Fannie Mae and Freddie Mac seized
AIG seized, Lehman Bros bankrupt
01 November 2007
15 November 2007
29 November 2007
13 December 2007
27 December 2007
10 January 2008
24 January 2008
07 February 2008
21 February 2008
06 March 2008
20 March 2008
03 April 2008
17 April 2008
01 May 2008
15 May 2008
29 May 2008
12 June 2008
26 June 2008
10 July 2008
24 July 2008
07 August 2008
21 August 2008
04 September 2008
18 September 2008
02 October 2008
16 October 2008
30 October 2008

Current odds on Sarah Palin being the next President of the United States?

14:1 from William Hill, 17th Nov

Flipping coins

Two types of uncertainty

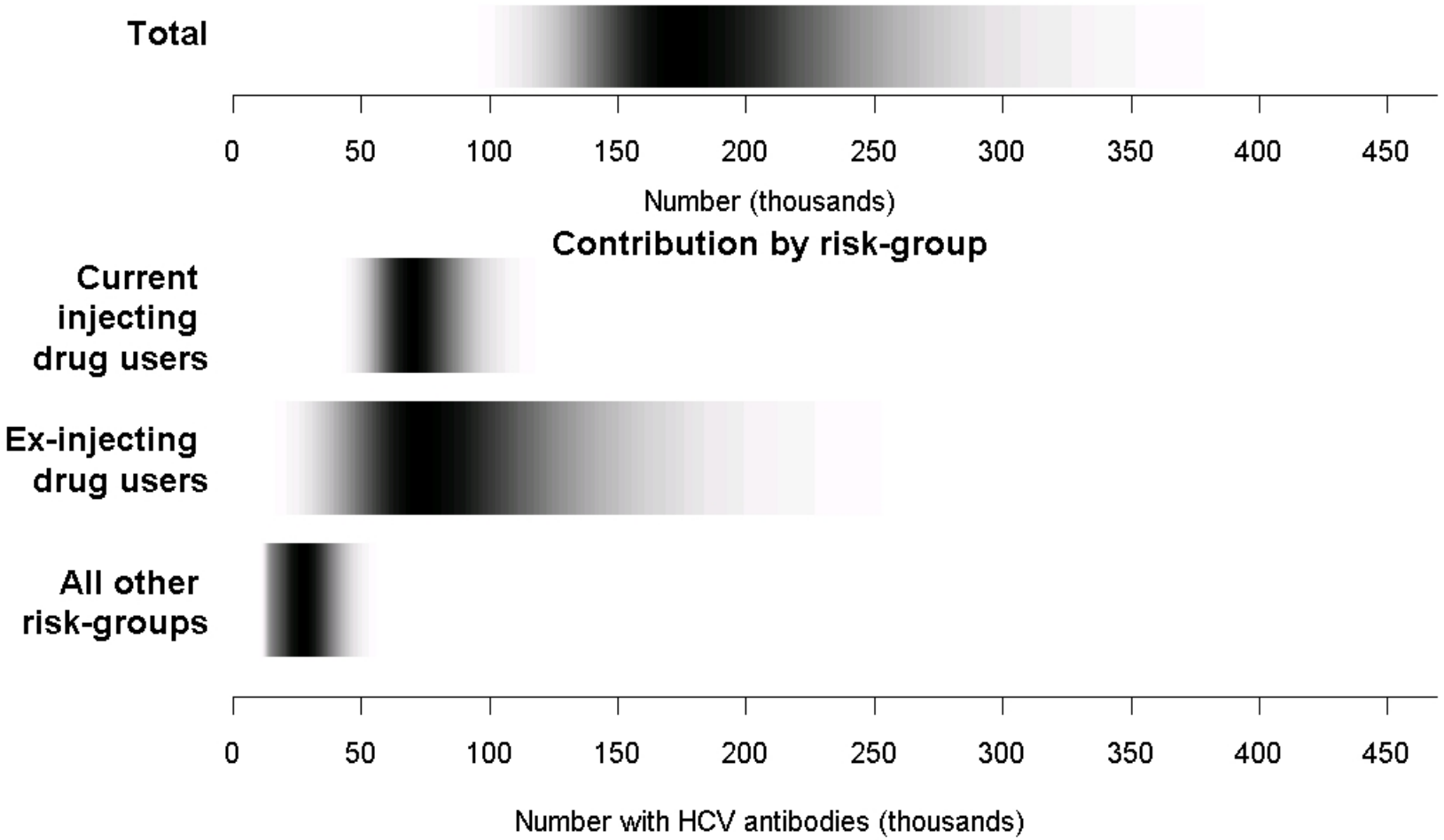
Aleatory

– chance, unpredictable

Epistemic

– lack of knowledge, ignorance

Hepatitis C prevalence in UK



Quantifying your ignorance

- Think whether you prefer (A) or (B) for each question
- Then think of how confident you are with your answer
- Give your confidence a number 5 to 10
- Score yourself when you hear the correct answer

A short quiz

1. Which is higher, A) the Eiffel tower, B) Canary Wharf?

   A (324m vs 235m)

2. Who is older, A) Prince William or B) Kate Middleton ?

   B (Born 21/6/82 vs 9/1/82)

3. Which is older, A) LSE or B) Imperial College?

   A (1895 vs 1907)

4. Which is larger, A) Belgium or B) Switzerland?

   B (30 vs 41 000 sq km)

5. Which is bigger, A) Venus B) Earth?

   B (6051 vs 6371 km radius)

6. Who died first, A) Beethoven or B) Napoleon?

   B (1827 vs 1821)

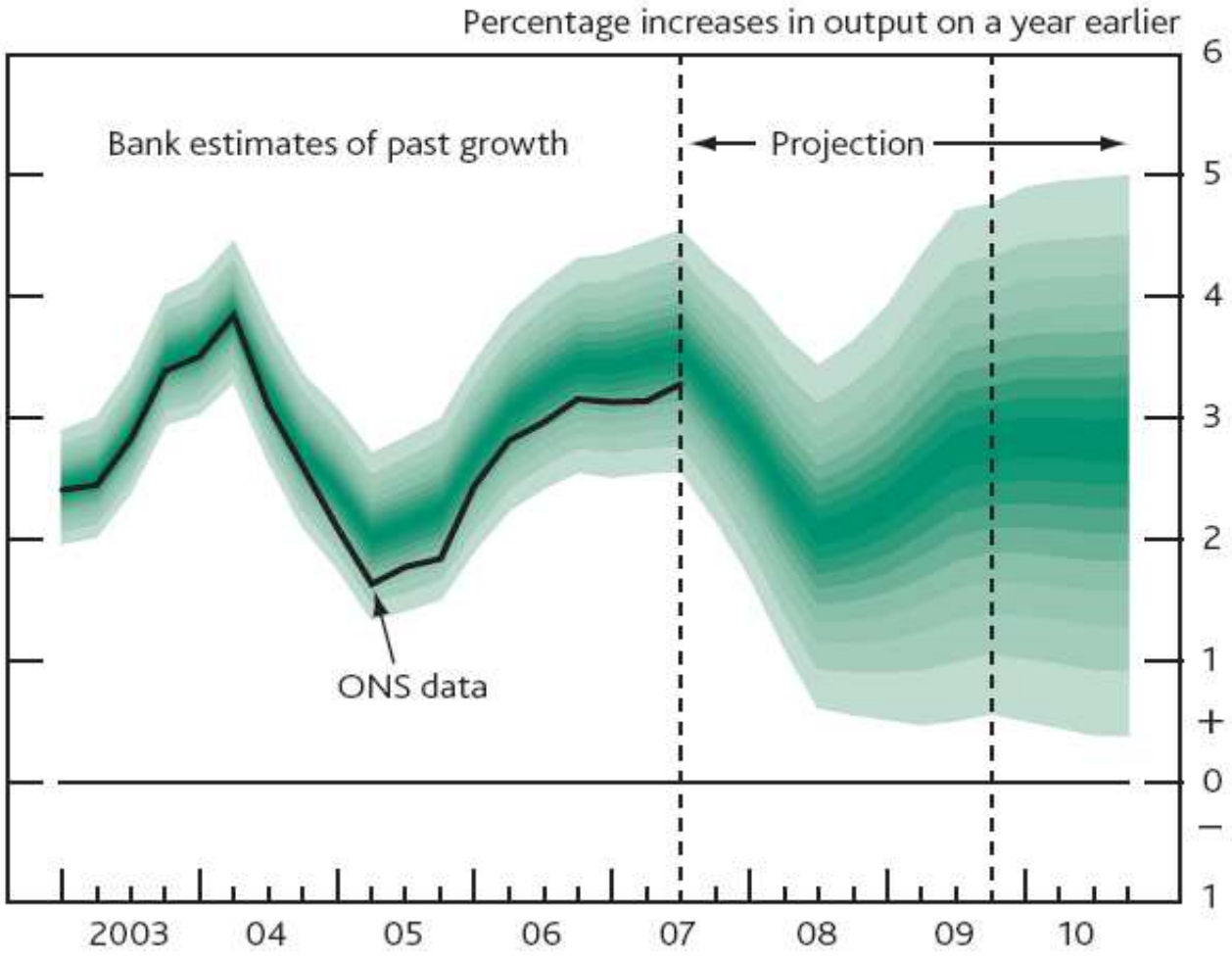
Your 'confidence' in your answer	***5***	***6***	***7***	***8***	***9***	***10***
Score if you are ***right***	0	9	16	21	24	25
Score if you are ***wrong***	0	-11	-24	-39	-56	-75

- Seems harsh on errors
- 25 – (error)^2
- A 'proper' scoring rule
- Encourages honesty

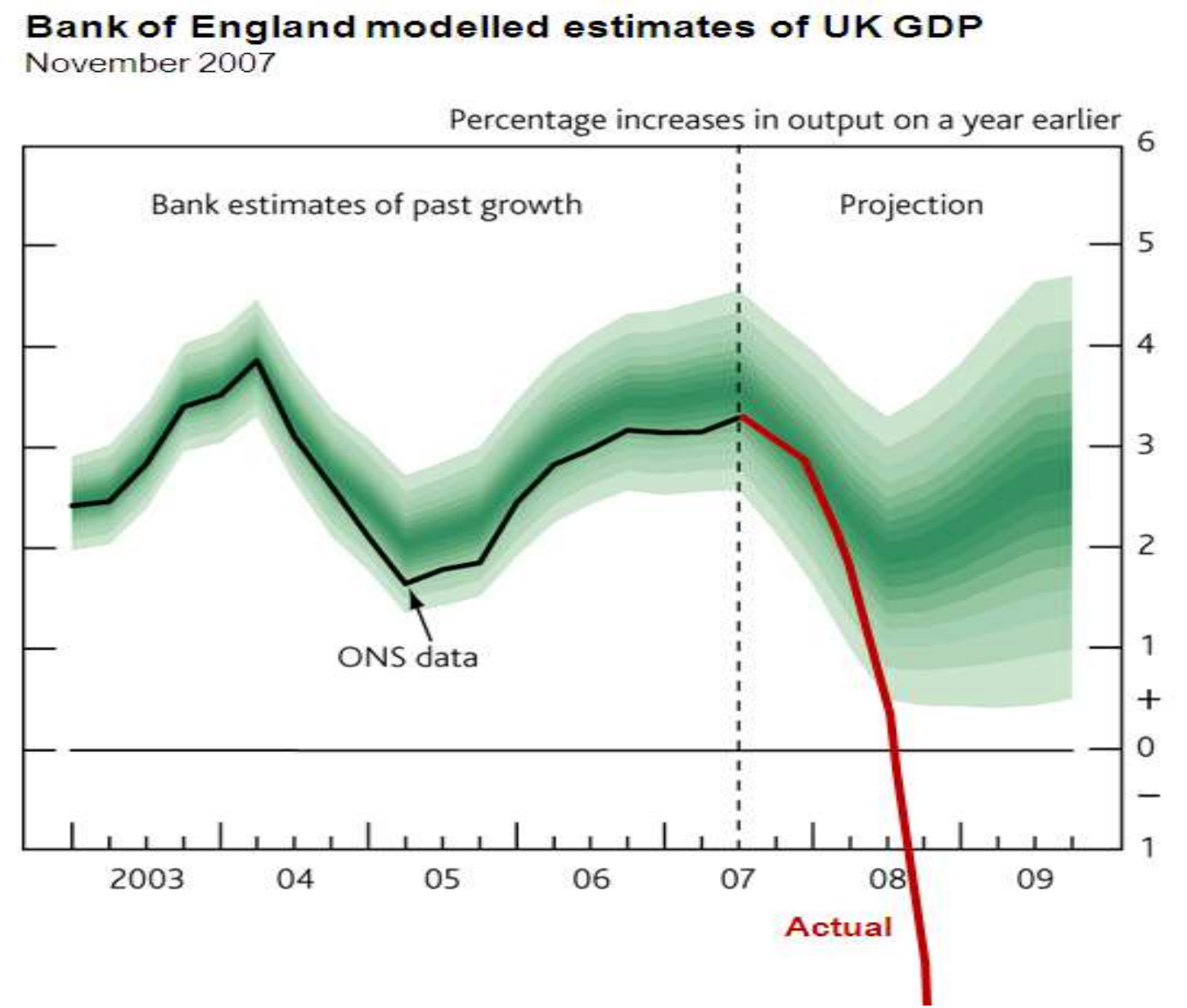
Bank of England Fan Charts

- *If economic circumstances identical to today's were to prevail on 100 occasions …*
- *Consequently, GDP growth is expected to lie somewhere within the entire fan on 90 out of 100 occasions*

Chart 5.1 GDP projection based on market interest rate expectations



Can compare with what happened



"But there are also unknown unknowns. There are things we do not know we don't know"

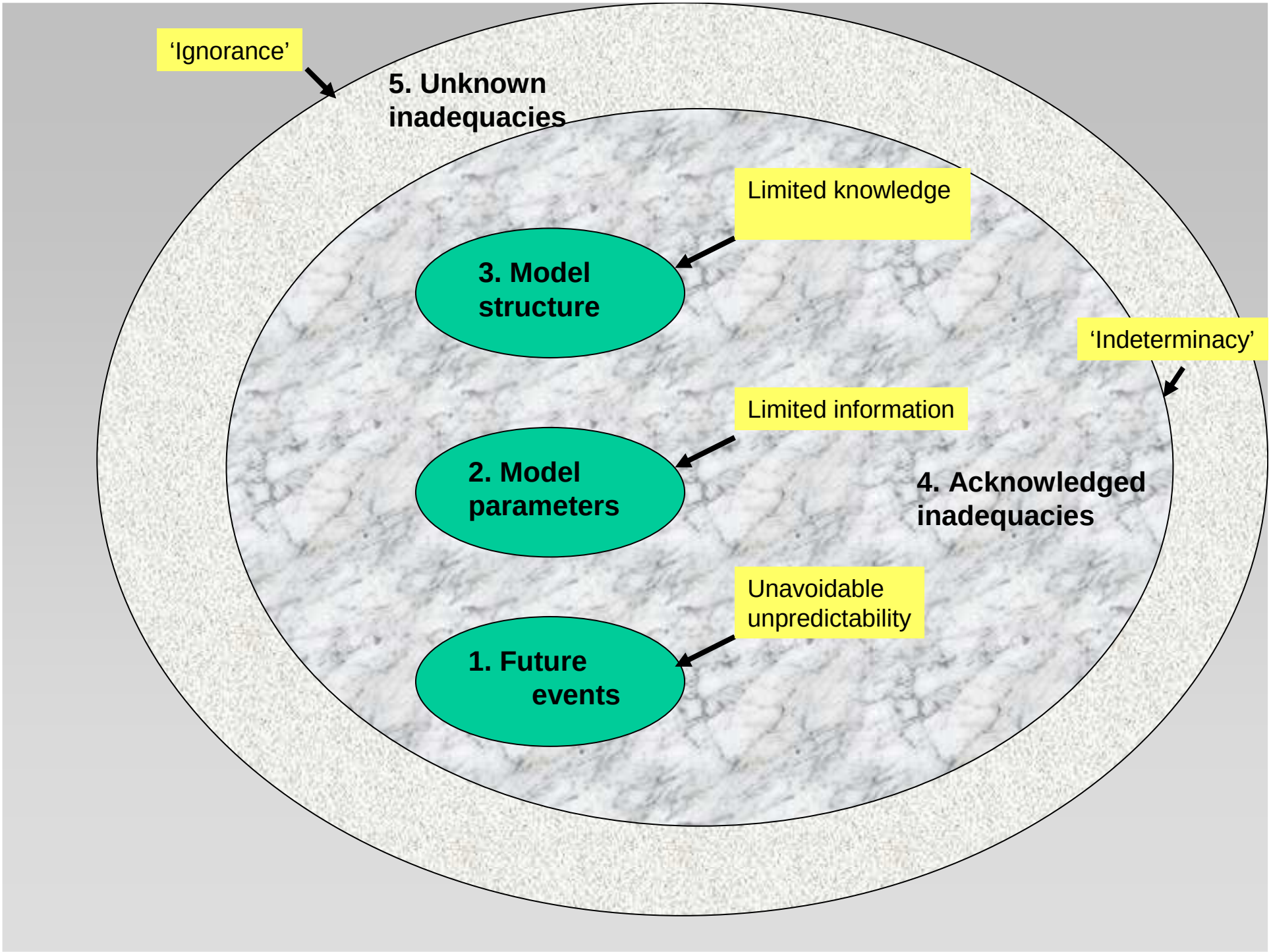
Acknowledging deeper uncertainties

- Frank Knight 1921*:*

  Risk*: quantifiable*

  vs Uncertainty *– not susceptible of measurement*

- Keynes 1937. *"About these matters there is no scientific basis on which to form any calculable probability whatsoever. We simply do not know."*


'Ignorance'
5. Unknown inadequacies
Limited knowledge
3. Model structure
'Indeterminacy'
Limited information
2. Model parameters
4. Acknowledged inadequacies
Unavoidable unpredictability
1. Future events

Expressing possible effects of factors left out of quantitative model

Source of uncertainty	*Direction and magnitude of effect*
Moderate under reporting of consumption is known to occur	--
Misreporting: some subjects will have reported the food that they ate in a wrong food category	+/-
Use of broad food categories causes over-estimation of exposure	+++
etc	*etc*
Qualitative evaluation of overall effect of identified uncertainties Estimates for high consumers are likely to over-estimate adult exposure by a moderate amount, but might be under-estimates for regional populations consuming locally-produced food and are probably under-estimates for children	**++ adults +/- local populations - - children**

European Food Standards Agency:

Qualitative evaluation of influence of uncertainties on an assessment of ochratoxin A exposure for high consumers

How can we communicate deeper uncertainties due to the quality of the evidence?

High quality	Further research is very unlikely to change our confidence in the estimate of effect
Moderate quality	Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate
Low quality	Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate
Very low quality	Any estimate of effect is very uncertain

Part of GRADE scale used in Cochrane Collaboration and 25 other organisations to assess confidence in estimates of medical treatment effects

Guadalajara, May 2009

The risk of swine flu? I haven't a clue . . .

. . . writes a professor of risk. But I'm still sending my daughter in Mexico some Tamiflu

It could have been designed to make me feel inadequate. I am a professor of risk, and when my daughter Rosie wanted to spend part of her gap year working on a newspaper, she chose, with a true nose for a story, to go to Mexico.

So it is assumed that I know the chances of her, and everyone else, getting or even dying of, swine flu. But I just don't know; risk is such an odd thing — no instrument can measure it but it constantly changes as we find out more information, just as the odds on Barack Obama being President oscillated wildly in the year before the election. What do we really mean by chance and risk anyway?

In some circumstances we can comfortably put a number on risk: if I spend £1 on a lottery ticket, I can calculate from the number of ways the balls can be drawn that there is a 1 in 14 million chance of winning the jackpot. Doing the sums for swine flu is a different matter: a heavenly compere doesn't pull balls with our names on out of a large bag, so epidemiologists resort to computer models of how epidemics work.

But instead of just having pure unavoidable chance, ignorance of the mechanics of the epidemic starts to dominate the calculations. It's a bit like trying to work out the odds of winning the lottery when you don't know how many balls there are.

The shape of the epidemic would be a lot more predictable if we knew all about this virus, and in particular something called the "reproductive number", which is how many people each case is expected to infect in an unaffected and unprotected population. For example, each case of measles would be expected to infect twenty people, which is why the fall in MMR vaccinations is viewed so anxiously; for smallpox it's about five and Sars about three.

Epidemiologists and insurers are rushing to estimate this quantity from the limited data: for this virus, it seems to be less than two, so a bit of effort might push it below the magic threshold of one, when the epidemic should disappear.

Meanwhile, my girl in Guadalajara reports that nobody there seems to care much about the reproductive number, and the lack of any clear information has brought a mixture of blind terror and indifference. For every few people not wearing masks someone is wearing four at once, just in case. And it's never long before the wearer's intrinsic Mexican-ness overrides instructions and face masks are yanked down to kiss a cheek or smoke a cigarette. The masks sold out completely on the second day of the scare, leading many people to fashion their own from dishclothes and bits of string.

The health minister in Guadalajara has only just admitted that there may possibly be some local cases, whereas in the UK the papers are providing full histories of every

Rosie reports that for every Mexican without a mask, another has four

contact — invaluable information for the epidemic model. But our ignorance goes beyond not knowing how infectious the virus is and the proportion of cases that die — the virus could mutate or, the feared outcome, join with avian flu to create a new strain: despite the opportunities for flying-pig jokes, this would be no laughing matter.

At least we can think of these possibilities and weigh them up, inevitably using a lot of judgment stirred in with the science. But our journey through ignorance can lead into the pitch-black of deep uncertainty — Donald Rumsfeld's unknown unknowns. It can be disastrous to believe that you have thought of everything — it seems clear that a big reason for the financial crisis was a belief that risk models were somehow "true" and that the world really worked according to the rules, and there was no preparation for when events did not fit the model.

So we need some humility and to admit that we may be wrong. Pundits may mock the level of uncertainty that says the eventual UK body count could be none or could be a million, but that is simply an expression of honest ignorance. The need is to have emergency systems that are precautionary at first, and then rapidly adapt to new knowledge obtained from good data. Deciding which vaccines to prepare for the winter flu season will require a delicate balance of risks and benefits — a real gamble in the face of uncertainty.

And even if a judgment is inevitable, the reasoning should at least have some science behind it, unlike Egypt's slaughter of the innocent pigs. Perhaps even that is better than the conspiracy theories circulating in Mexico, inviting us to believe that the virus was introduced by the Americans, the pharmaceutical industry or to distract attention from the drug cartels.

Anyway, my gut feeling is that the chances we will see the girl again are looking quite good. But we've sent out Tamiflu just in case.

David Spiegelhalter is Winton Professor of the Public Understanding of Risk at the University of Cambridge. Rosie Spiegelhalter is sticking it out in Mexico

Government response to scientific uncertainty?

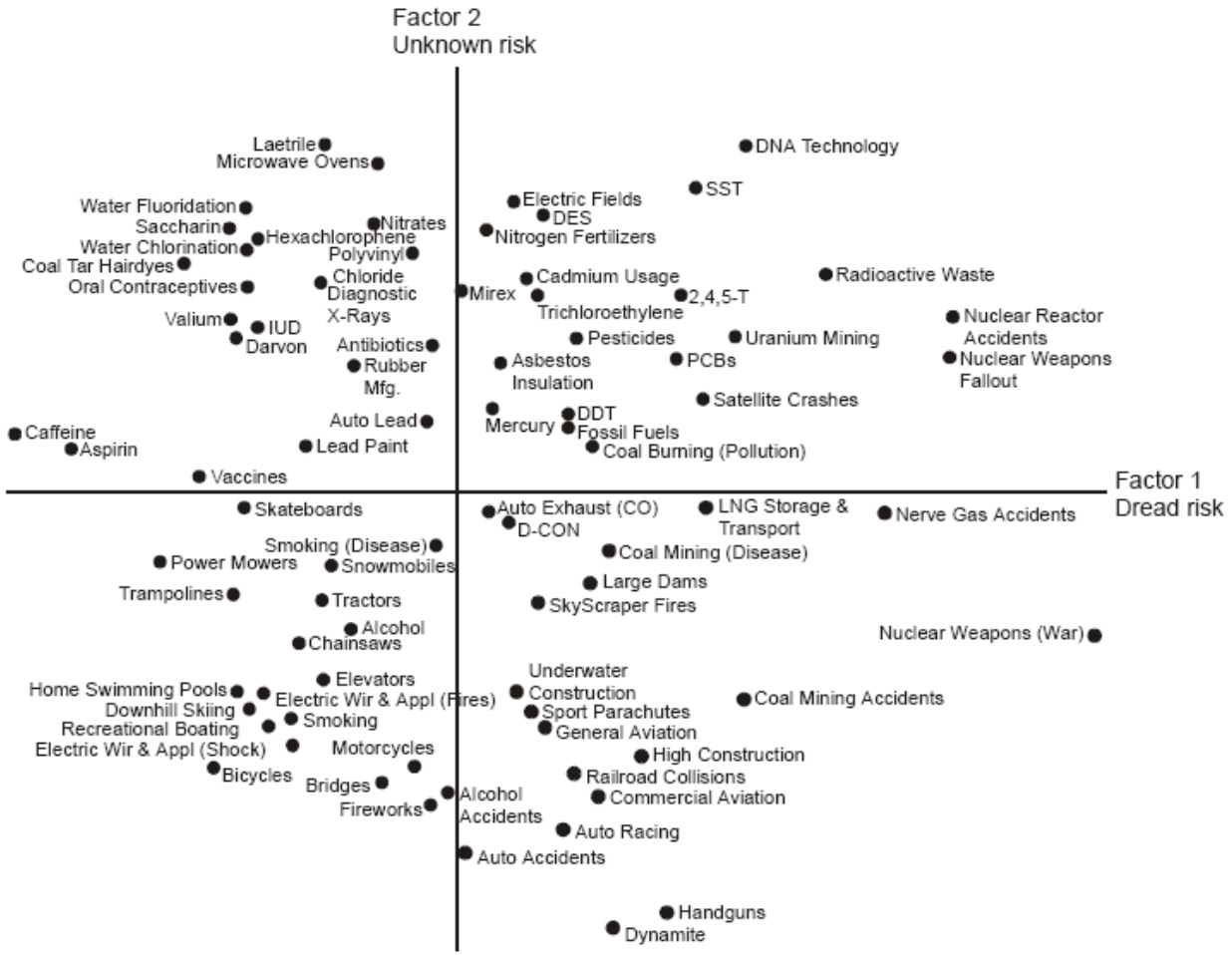
- 'Worst case scenarios' of 30% clinical cases, of which
  - July 2009 : 1/300 die - 65,000 deaths
  - Sept 2009: 1/1000 die - 19,000
- Ultra-precautionary planning – these were implausible combinations even given knowledge at the time
- In fact around 450 deaths
- Can we afford this level of caution?

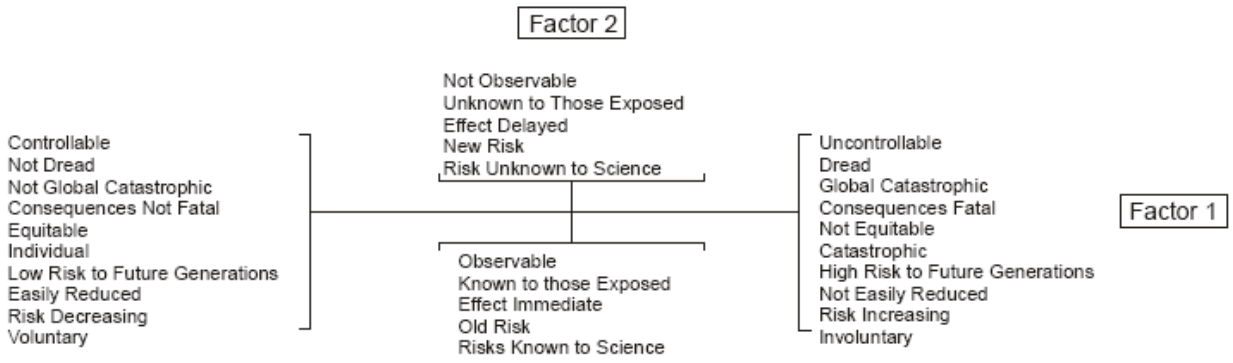
How do people respond to risk?

Personal responses dominated by

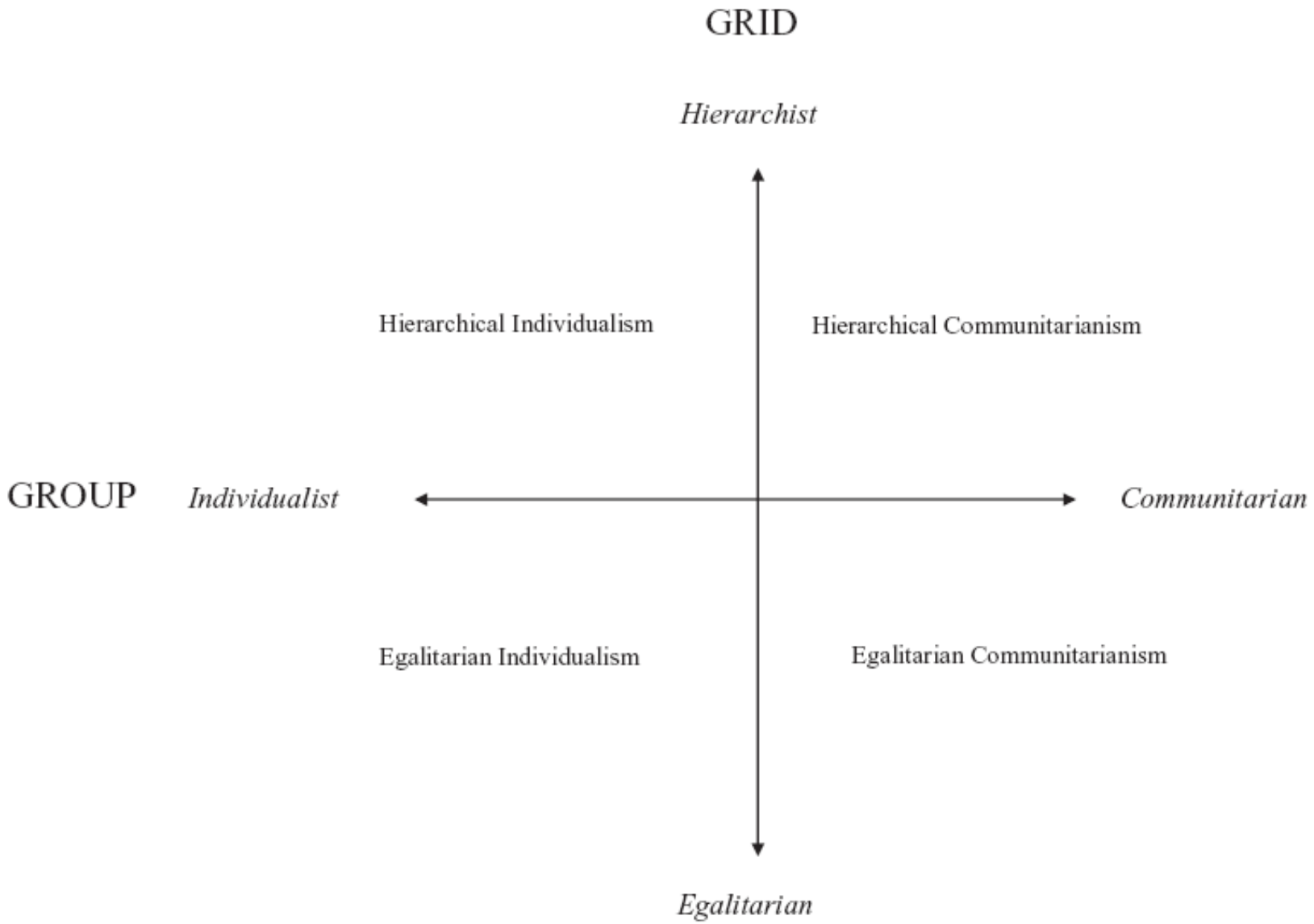
- Emotion and personality
- Personal experiences
- Feelings of control / imposition
- Cultural beliefs about how society should be organised: individualist/communitarian, hierarchical/egalitarian
- Trust (or lack of it) of authority
- Familiarity / 'dreadness' of hazard
- 'Innocence' of victims
- Social norms

'Probabilities' are largely ignored

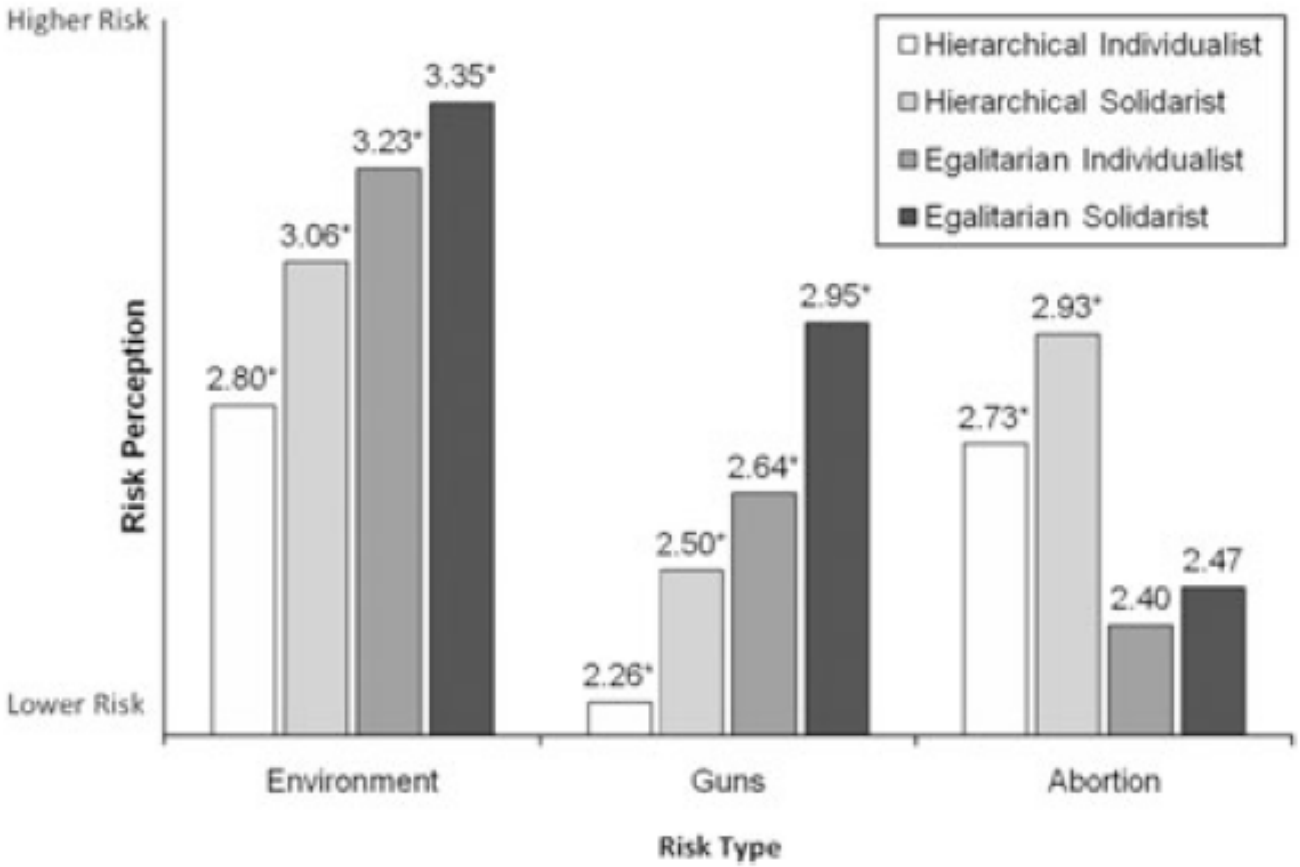




- Risk perception Slovic (1987) - ‘psycho-metric paradigm’



- Cultural theory (Douglas and Wildavsky, 1982



Journal of Empirical Legal Studies
Volume 4, Issue 3, 465–505, November 2007

Culture and Identity-Protective Cognition: Explaining the White-Male Effect in Risk Perception

*Dan M. Kahan, Donald Braman, John Gastil, Paul Slovic, and C. K. Mertz**

Synthesis: "Cultural cognition"

OPINION

NATURE|Vol 463|21 January 2010

Fixing the communications failure

People's grasp of scientific debates can improve if communicators build on the fact that cultural values influence what and whom we believe, says **Dan Kahan**.

"People endorse whichever position reinforces their connection to others with whom they share important commitments."

TIMESONLINE

NEWS | COMMENT | BUSINESS | MONEY | SPORT | LIFE & STYLE | TRAVEL | DRIVING

UK NEWS | WORLD NEWS | POLITICS | SCIENCE | ENVIRONMENT | WEATHER | TECH & WEB | VIDEO

Where am I? › Home › News › Environment

From The Times

February 2, 2010

Beat that: double yolks defy one in a trillion odds

All the eggs in the box had double yolks

TIMES RECOMMENDS

- How the Saints preserved New Orleans
- Please God, can I have a daughter next?
- We need a new capitalism to take on China

EUREKA ZONE ››

Egg Council said 1/1000 eggs double-yoked

So chance of 6 eggs = 1/1000 x 1/1000 …

= 1 in 1,000,000,000,000,000,000

What's wrong with this?

Acknowledging uncertainty/error at different levels:

1. *Event probability wrong*: 2,000,000,000 half-dozen eggs in UK every year, and so would expect to wait 500,000,000 years for an event this rare to happen.

2. '*Parameters' are wrong*: double-yokes more common in extra-large

3. '*Model*' is wrong: eggs in a box are not independent, likely to come from similar batch

I had 6 double-yolks in the next box of eggs I bought!!

£2.49 from my local Waitrose

Models are like guide books

They can be

- Out-of-date
- Too simple
- Too complicated
- Wrong

But they can still be useful if used with caution!

And we have to acknowledge that disputes are not only because of ignorance but because of different cultural world-views

The Treatment of Uncertainties in the Fourth IPCC Assessment Report

Clearly a long and arduous struggle for consistency

'Likelihood' used by WG1

Terminology	*Degree of confidence in being correct*
Virtually certain	> 99% probability of occurrence
Very likely	> 90%
Likely	> 66%
About as likely as not	33% to 66% probability
Unlikely	< 33% probability
Very unlikely	< 10% probability
Exceptionally unlikely	< 1% probability

Table 2: Likelihood scale recommended for use of Working Groups of the IPCC (9)

*"Most of the observed increase in global average temperatures since the mid-20th century is **very likely** due to the observed increase in anthropogenic greenhouse gas concentrations"*

‘Confidence’ used by WG2

Terminology	*Degree of confidence in being correct*
Very high confidence	At least 9 out of 10 chance of being correct
High confidence	About 8 out of 10 chance
Medium confidence	About 5 out of 10 chance
Low confidence	About 2 out of 10 chance
Very low confidence	Less than 1 out of 10 chance

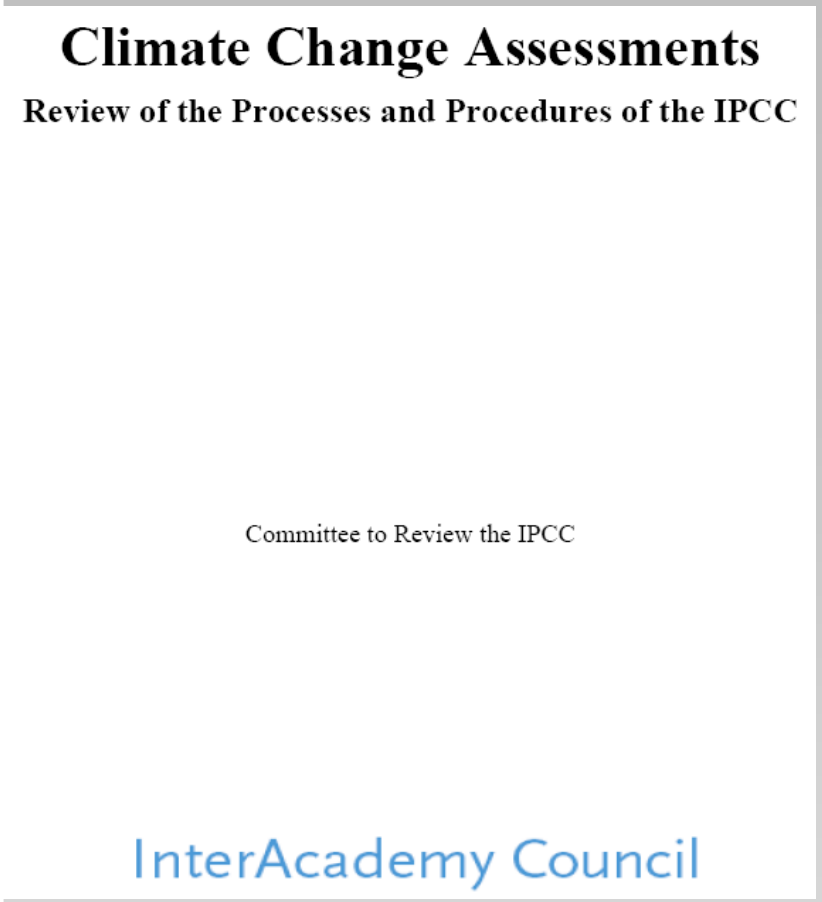
Table 3: Quantitatively calibrated levels of confidence recommended for use of Working Groups of the IPCC (9)

“In some countries, yields from rain-fed agriculture could be reduced by up to 50% by 2020.
(High confidence; IPCC, 2007b, p. 13)”

WG3 used a qualitative scale

Level of agreement or consensus	High agreement, limited evidence	…	High agreement, much evidence
	…	…	…
	Low agreement, limited evidence	…	Low agreement, much evidence
	Amount of evidence (theory, observations, models)		

Table 1. Qualitatively defined levels of understanding recommended for use of Working Groups of the IPCC



August 2010

On uncertainty –

- Pointed out anomalies
- Guidance not followed
- Recommended using level-of-understanding scale
- Drop numerical confidence scale

Department of Statistics public lecture

Trying to Quantify Uncertainty

Professor David Spiegelhalter
Winton Professor of the Public Understanding of Risk, University of Cambridge and senior scientist, MRC Biostatistics Unit

Professor Anthony Atkinson
Chair, LSE