Expert Judgment Network: Bridging the Gap Between Scientific Uncertainty and Evidence-Based Decision Making

Strathclyde University Business School
2 – 4 April 2014

Uncertainty erupting

Willy Aspinall

Bristol University / Aspinall & Associates

Montserrat Risk Map, showing population centres & hazard micro-management zones

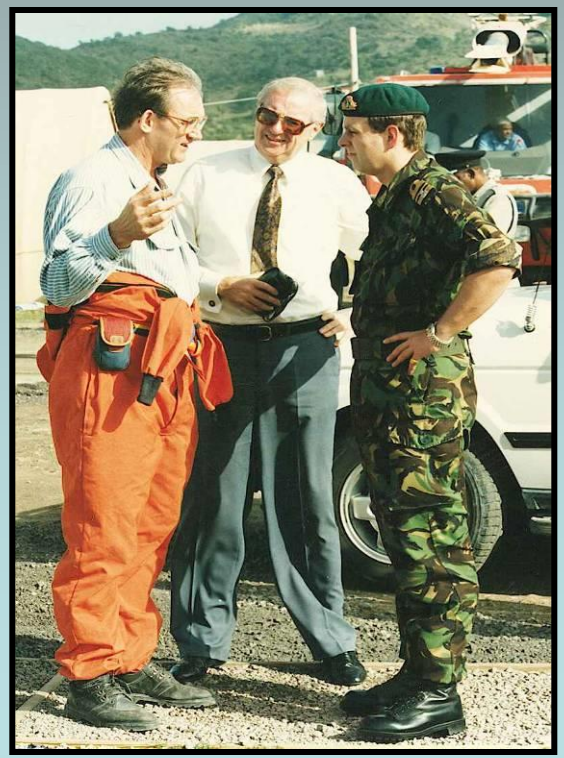
"…the island is exactly the wrong size for an eruption…"

Prompted by the Guadeloupe 1976 experience*….

….in Montserrat, we put in place a formalised procedure for providing scientific advice to the authorities

*

…..using Cooke's Classical Model and EXCALIBUR



Hincks *et al. Journal of Applied Volcanology* #CITATION
#ARTICLE_URL_DISPLAY_TEXT_FOR_STAMPED_PDF

Journal of Applied Volcanology
a SpringerOpen Journal

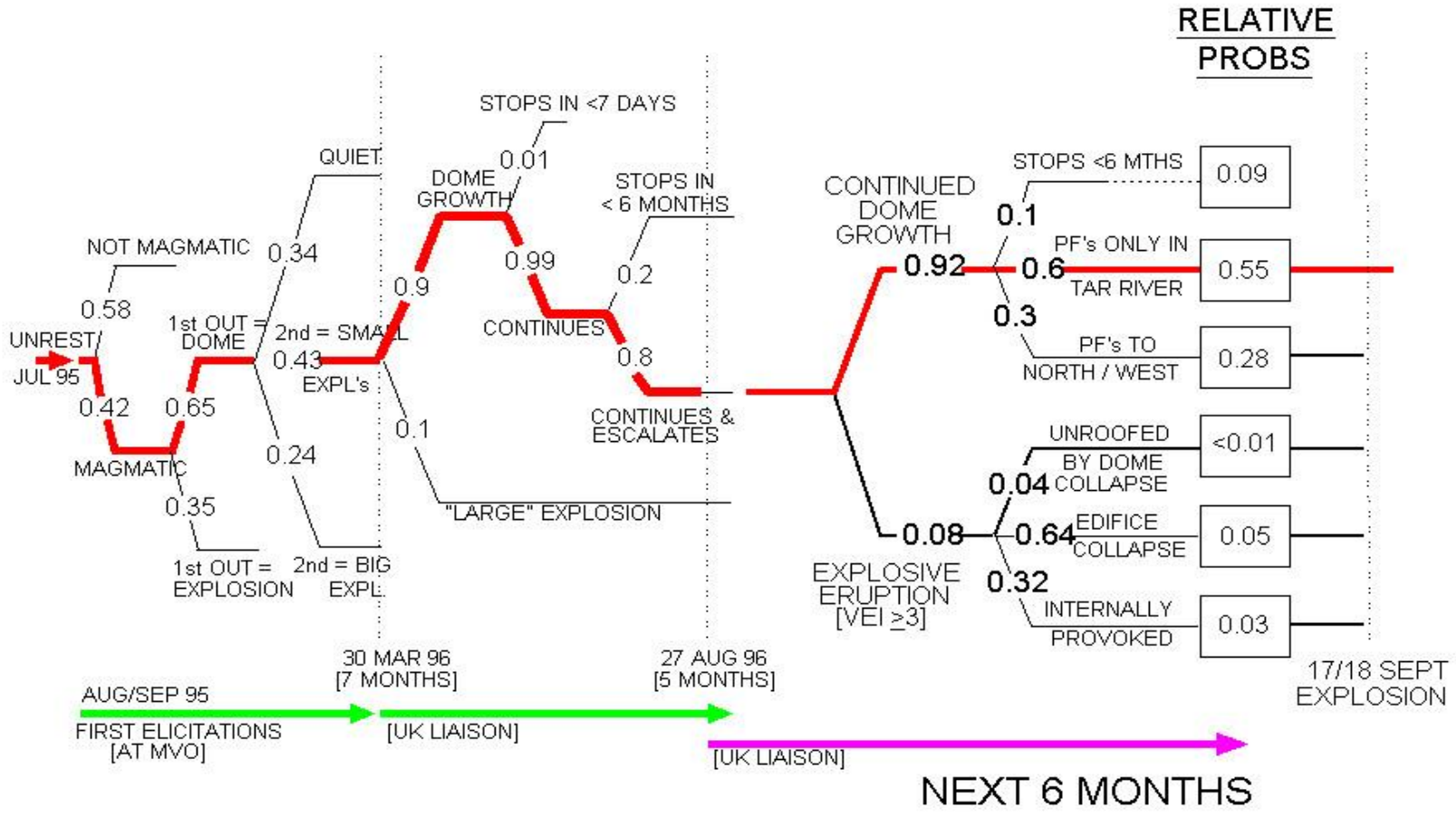
RESEARCH Open Access

Retrospective analysis of uncertain eruption precursors at La Soufrière volcano, Guadeloupe, 1975–77: volcanic hazard assessment using a Bayesian Belief Network approach

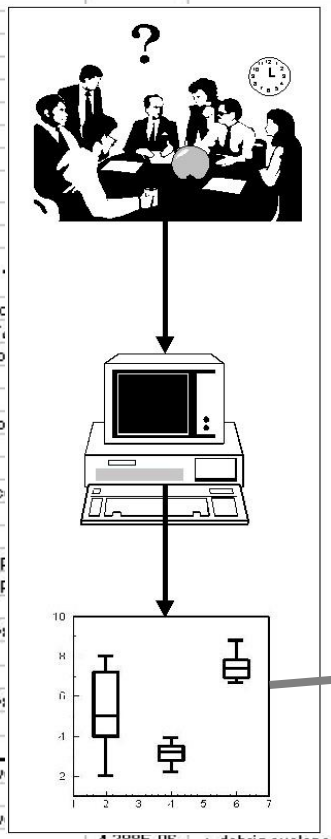
Thea K Hincks[1*], Jean-Christophe Komorowski[2], Stephen R Sparks[1] and Willy P Aspinall[1,3]

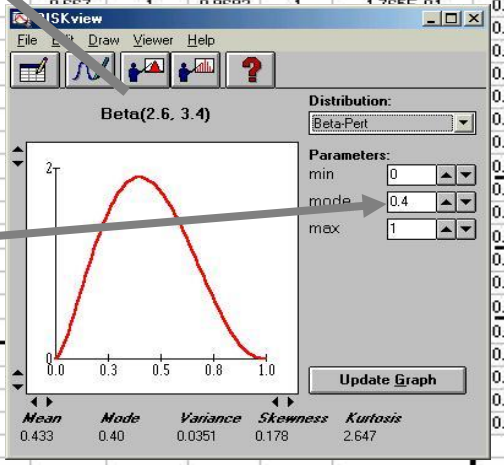
MONTSERRAT VOLCANO CRISIS

EVENT PROBABILITY TREE - UPDATE 27 AUG 96



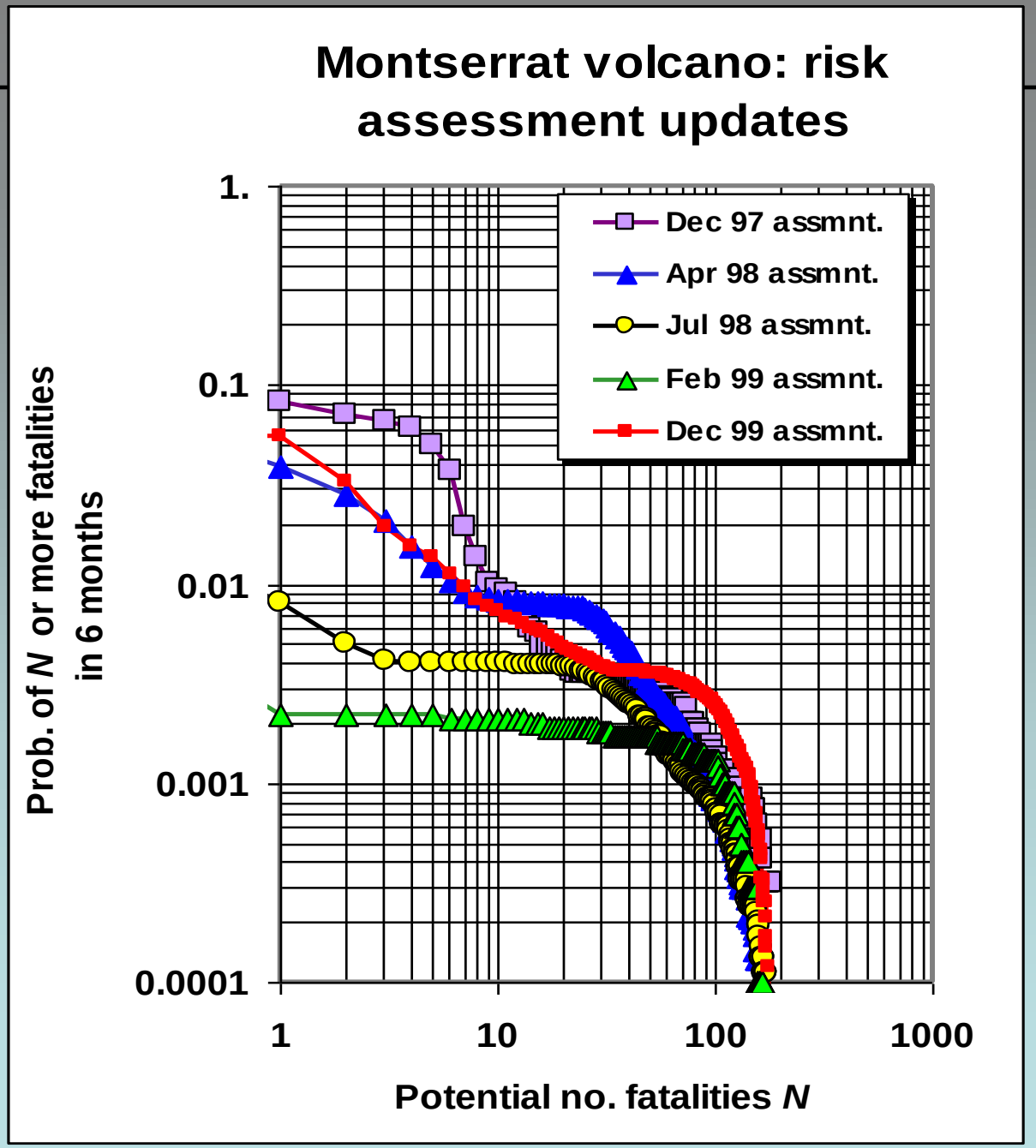
Monte Carlo simulation of potential casualty risks using parameter uncertainty distributions from probability tree



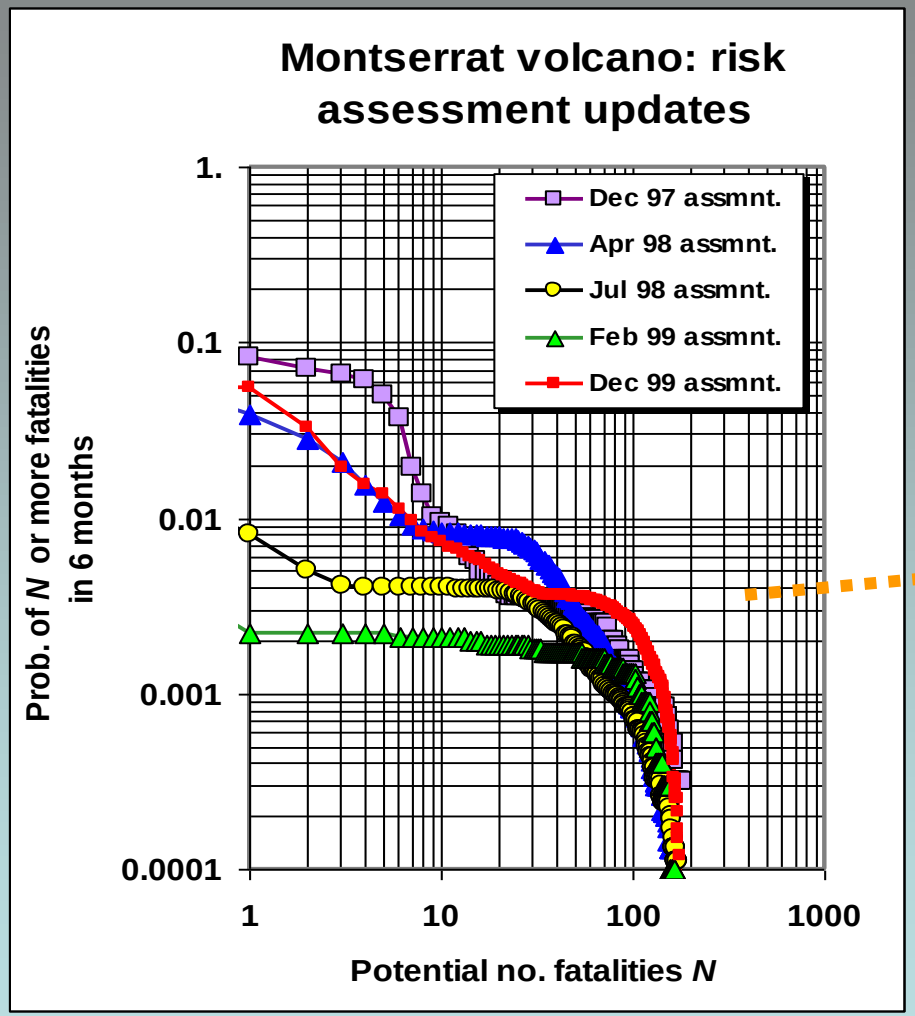


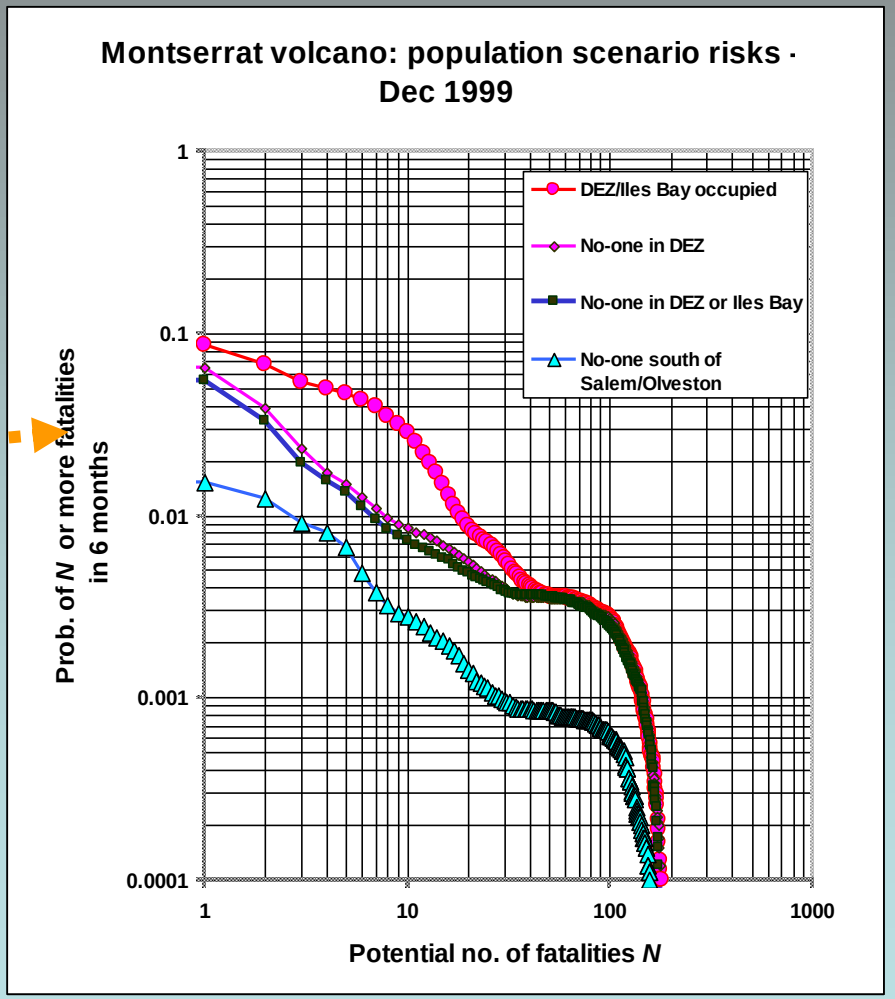
……producing so-called F-N casualty exceedance risk curves, expressing societal risk levels at different probabilities

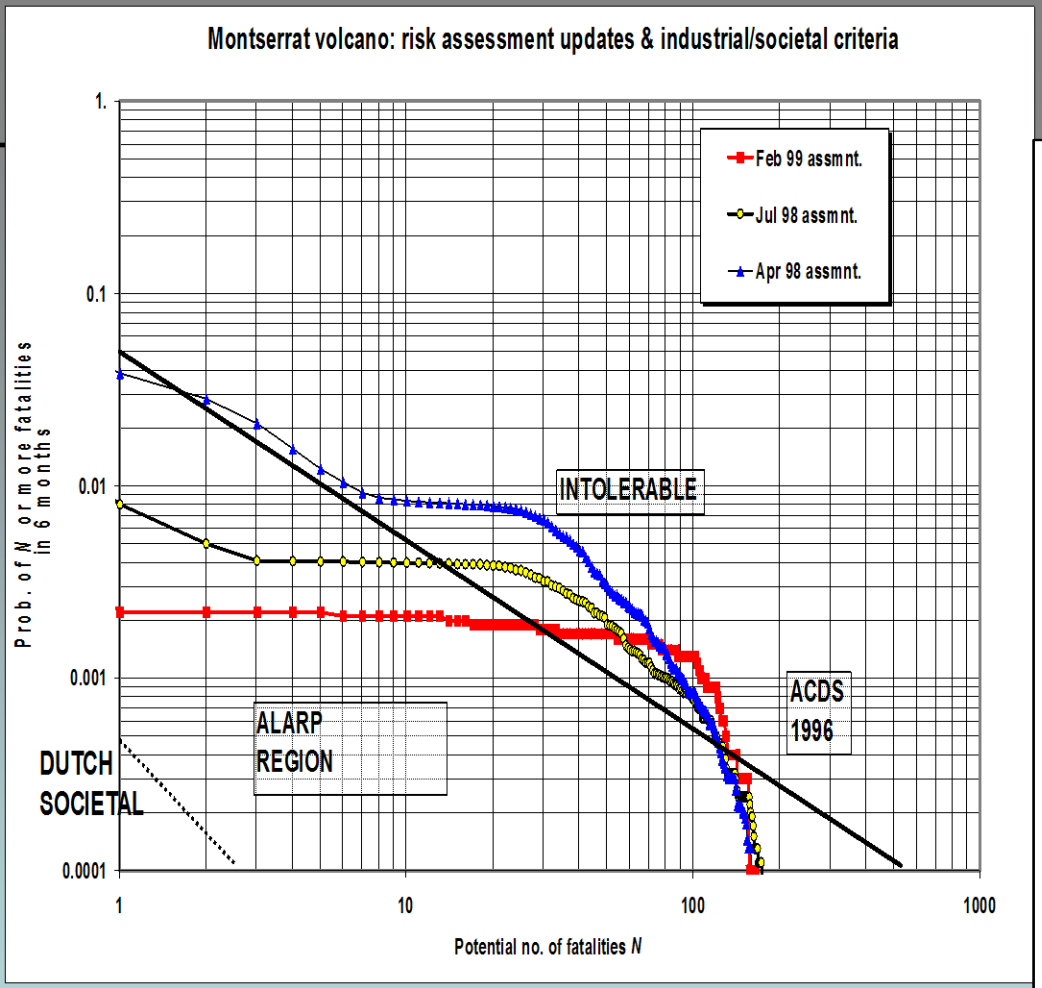
Pace Tim!



Population risk curves: regular updates,…….. and mitigation by staged evacuation



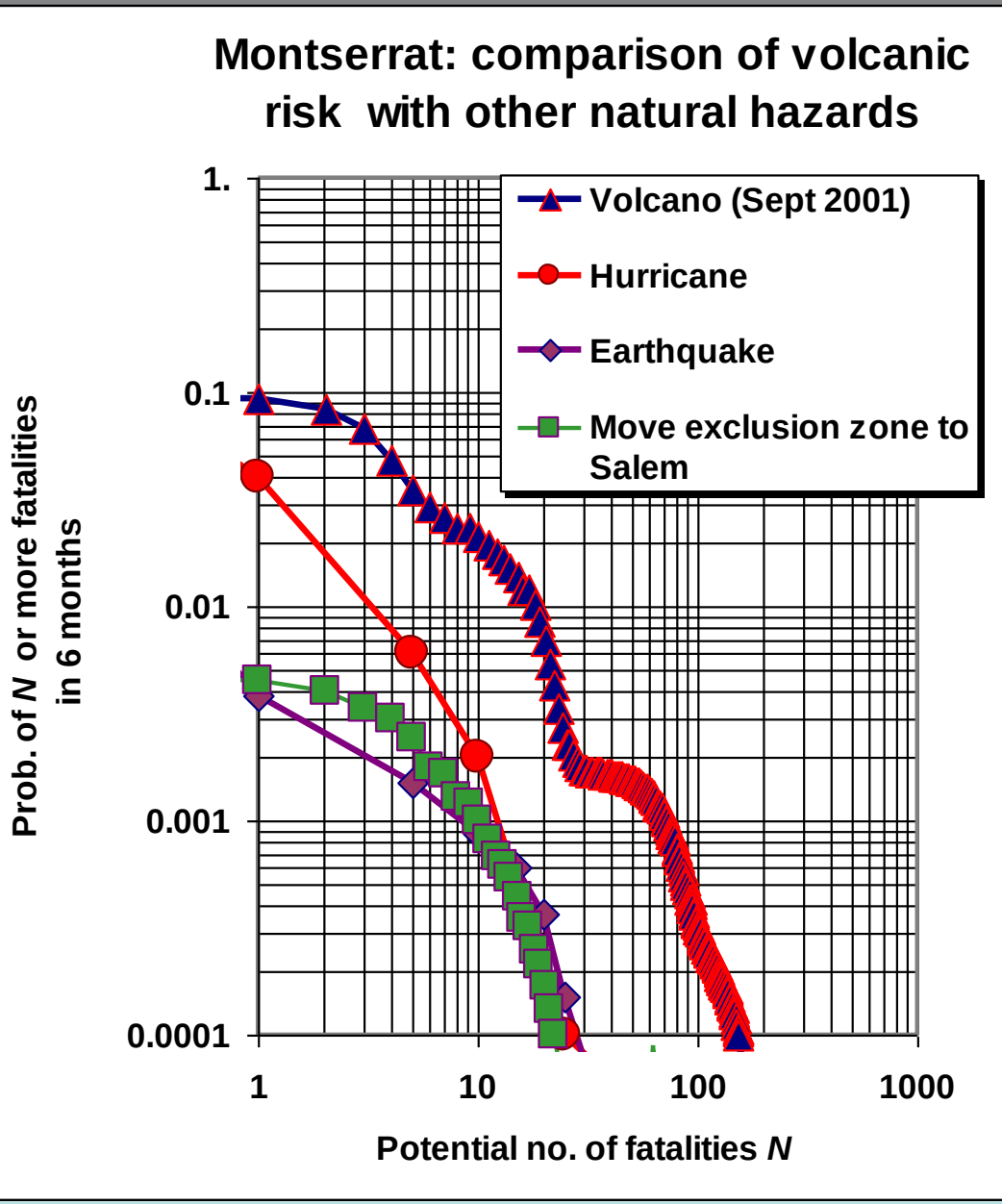




Comparative "acceptable" risk levels:

industrial criteria,…

or other natural hazards??



Probabilistic forecasting for Montserrat volcano

Typical forecast question: GIVEN current conditions, what is the probability that within the next year the first significant development will be the resumption of lava extrusion

SAC elicitation	Credible interval lower bound	Median estimate	Credible interval upper bound
Prob Odds	**6.3%** **15 - 1**	**34.1%** **2 - 1**	**66.1%** **1 - 2**

Brier Skill Score : the forecast method has predictive skill relative to some reference (e.g. climate record) if BSS is positive.

A perfectly accurate forecast method has BSS = 1;

bad forecasting leads to a negative BSS score

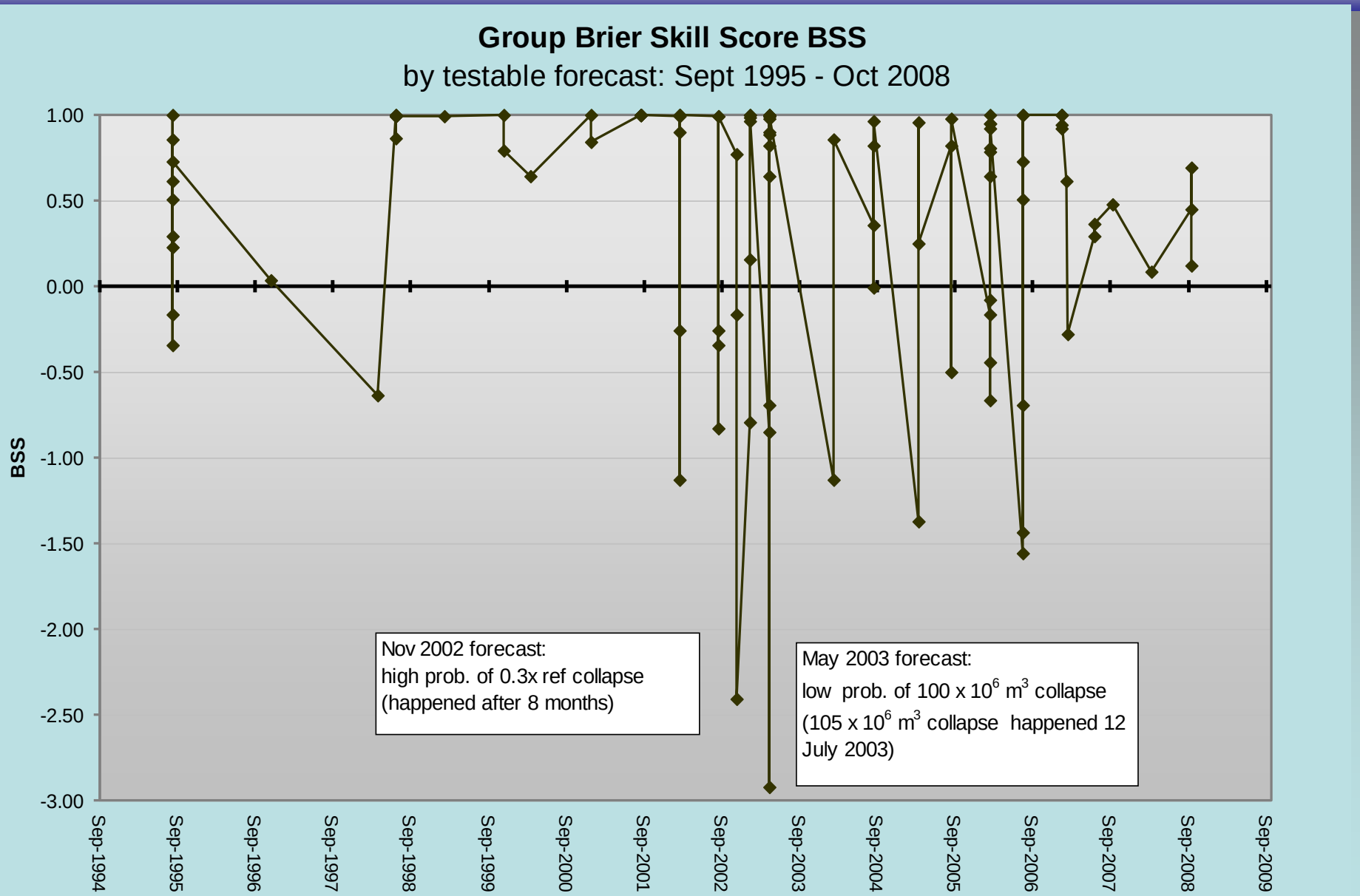
Probabilistic forecast scorecard

	+ve BSS	zero or -ve BSS
All forecasts (110 no.)	84 (76%)	26 (24%)
Life critical forecasts (75 no.)	61 (83%)	14* (17%)

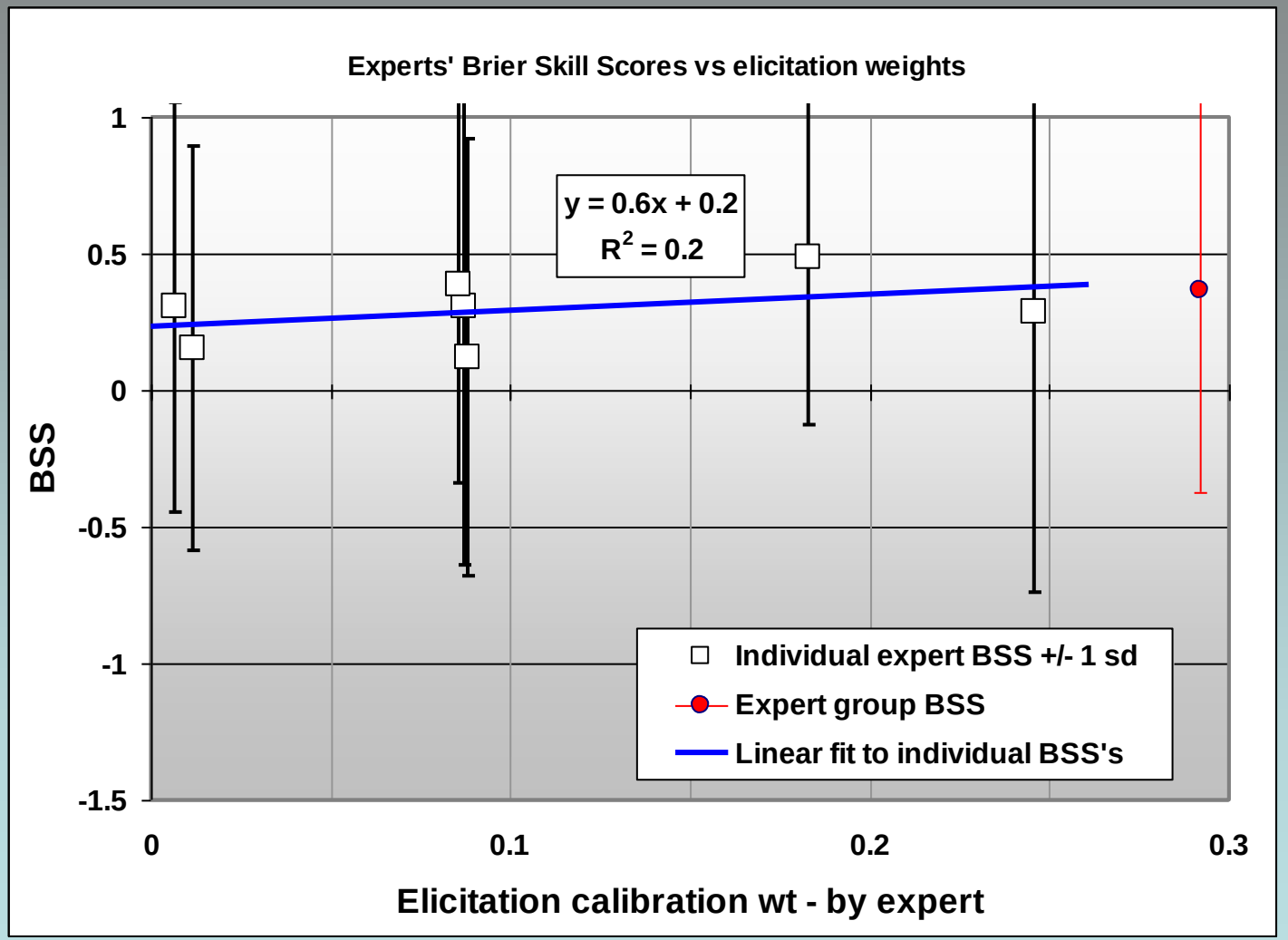
*** includes some important 'life threatening' scenarios**

∴ cautious

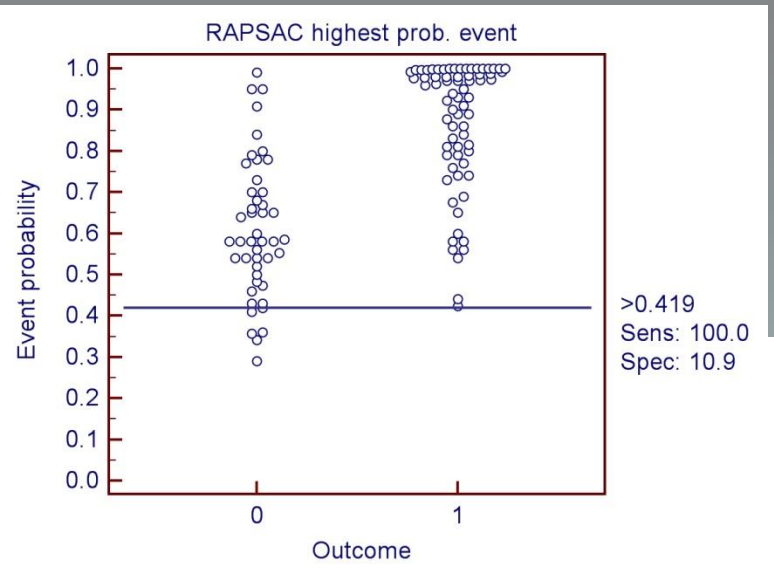
Forecast skill performance of Montserrat SAC



Brier -v- Cooke

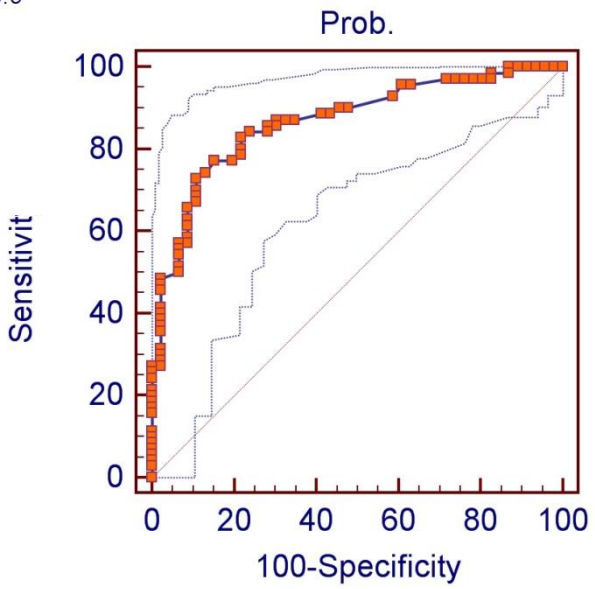


Alternative to Brier Skill Score?



*In signal detection theory, a **receiver operating characteristic (ROC) curve** illustrates the performance of a binary classifier: plots TPR = true positive rate vs. TNR = true negative rate*

FPR = false positive rate = Specificity



ROC curve

Variable	Prob.
Classification variable	Outcome

Sample size		116
Positive group :	Outcome = 1	70
Negative group :	Outcome = 0	46

Disease prevalence (%)	60.3

Area under the ROC curve (AUC)

Area under the ROC curve (AUC)	0.867
Standard Error[a]	0.0334
95% Confidence interval[b]	0.792 to 0.923
z statistic	10.996
Significance level P (Area=0.5)	<0.0001

[a] DeLong et al., 1988
[b] Binomial exact

Youden index

Youden index J	0.6199
95% Confidence interval[a]	0.4800 to 0.7261
Associated criterion	>0.8
95% Confidence interval[a]	0.7 to 0.914664123

[a] BC_a bootstrap interval (1000 iterations).

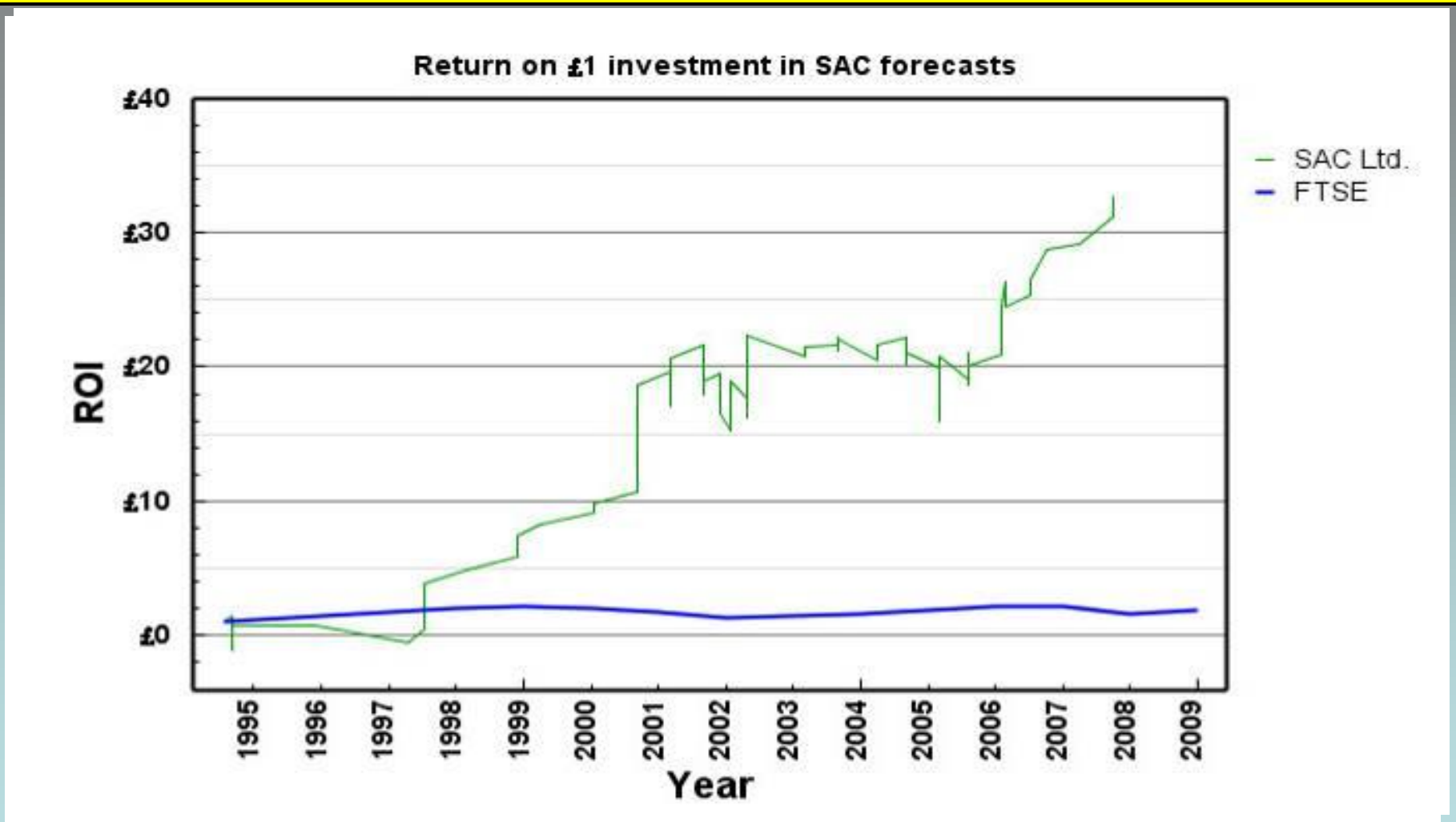
Optimal criterion

Optimal criterion[a]	>0.42
95% Confidence interval[b]	-
Sensitivity	100.00
Specificity	13.04

[a] Taking into account disease prevalence and estimated costs:
cost False Positive: 10; cost False Negative: 1000
cost True Positive: 10; cost True Negative: 0
[b] BC_a bootstrap interval (1000 iterations).

AUC area is the Mann-Whitney version of the Wilcoxon nonparametric two-sample statistic: 86% of the time, an actual event (1) has higher forecast prob than a non-event (0)

Communicating forecast skill



[Hagedorn, R., Smith, L.A. (2008) Communicating the value of probabilistic forecasts with weather roulette. Meteorol. Appl. Published online in Wiley InterScience (www.interscience.wiley.com) DOI: 10.1002/met.9.]

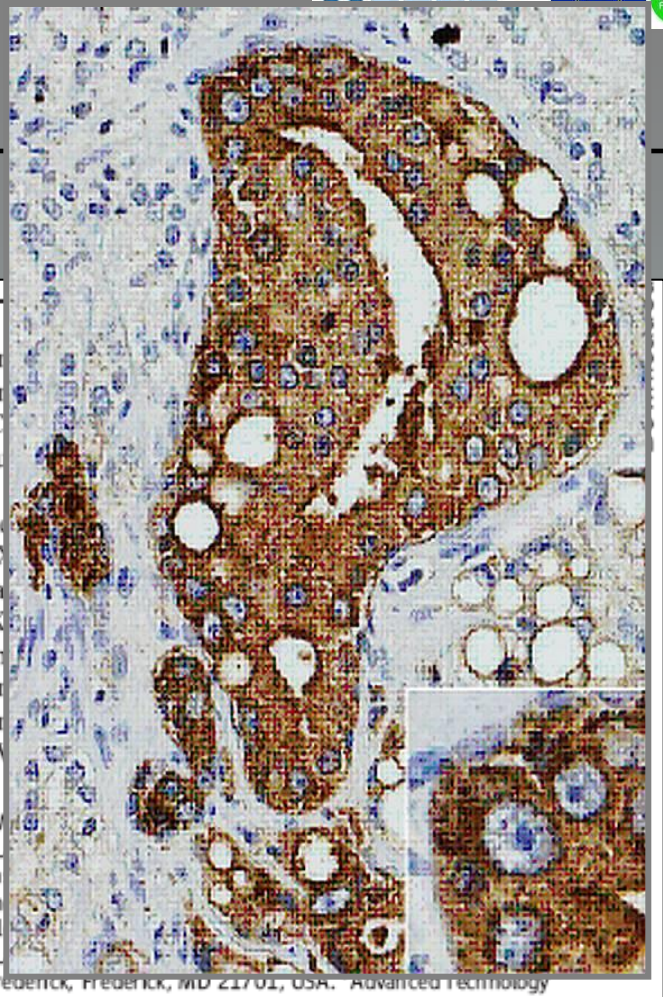
Big news!

Detection of an Infectious Retrovirus, XMRV, in Blood Cells of Patients with Chronic Fatigue Syndrome

Vincent C. Lombardi,[1*] Francis W. Ruscetti,[2*] Jaydip Das Gupta,[3] Max A. Pfost,[1] Kathryn S. Hagen,[1] Daniel L. Peterson,[1] Sandra K. Ruscetti,[4] Rachel K. Bagni,[5] Cari Petrow-Sadowski,[6] Bert Gold,[2] Michael Dean,[2] Robert H. Silverman,[3] Judy A. Mikovits[1]†

Chronic fatigue syndrome (CFS) is a debilitating disease of unknown etiology that is estimated to affect 17 million people worldwide. Studying peripheral blood mononuclear cells (PBMCs) from CFS patients, we identified DNA from a human gammaretrovirus, xenotropic murine leukemia virus–related virus (XMRV), in 68 of 101 patients (67%) as compared to 8 of 218 (3.7%) healthy controls. Cell culture experiments revealed that patient-derived XMRV is infectious and that both cell-associated and cell-free transmission of the virus are possible. Secondary viral infections were established in uninfected primary lymphocytes and indicator cell lines after their exposure to activated PBMCs, B cells, T cells, or plasma derived from CFS patients. These findings raise the possibility that XMRV may be a contributing factor in the pathogenesis of CFS.

Chronic fatigue syndrome (CFS) is a disorder of unknown etiology that affects multiple organ systems in the body. Patients with CFS display abnormalities in immune system function, often including chronic activation of the innate immune system and a deficiency in natural killer cell activity (*1*, *2*). A number of viruses, including ubiquitous herpesviruses and



Program, National Cancer Institute–Frederick, Frederick, MD 21701, USA. [6]Basic Research Program, Scientific Applications International Corporation, National Cancer Institute–Frederick, Frederick, MD 21701, USA.

*These authors contributed equally to this work.
†To whom correspondence should be addressed. E-mail: judym@wpinstitute.org

XMRV Expert Elicitation Workshop

XMRV = Xenotropic murine leukemia virus-related virus

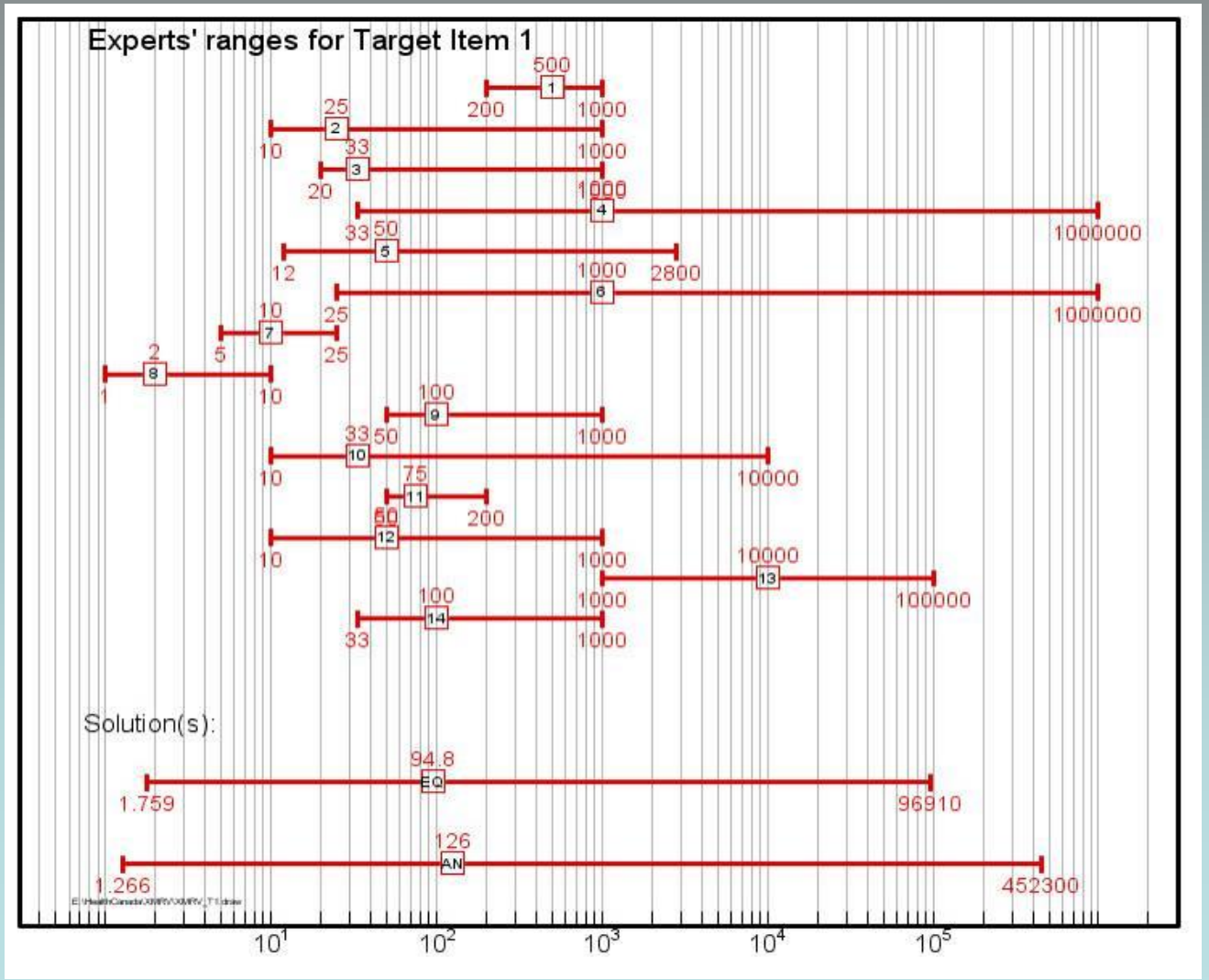
a gammaretrovirus first described in 2006

Following calibration, the experts were asked to answer a number of target questions for which answers are unknown.

Work with McLaughlin Centre for Population Health Risk Assessment, Univ. Ottawa

Target Question 1

A set of target questions that asked about the current prevalence of XMRV infection in the world (1), Canada (3), USA (4), UK (5) and France (6) in the general adult population? (1 in xxxxx)

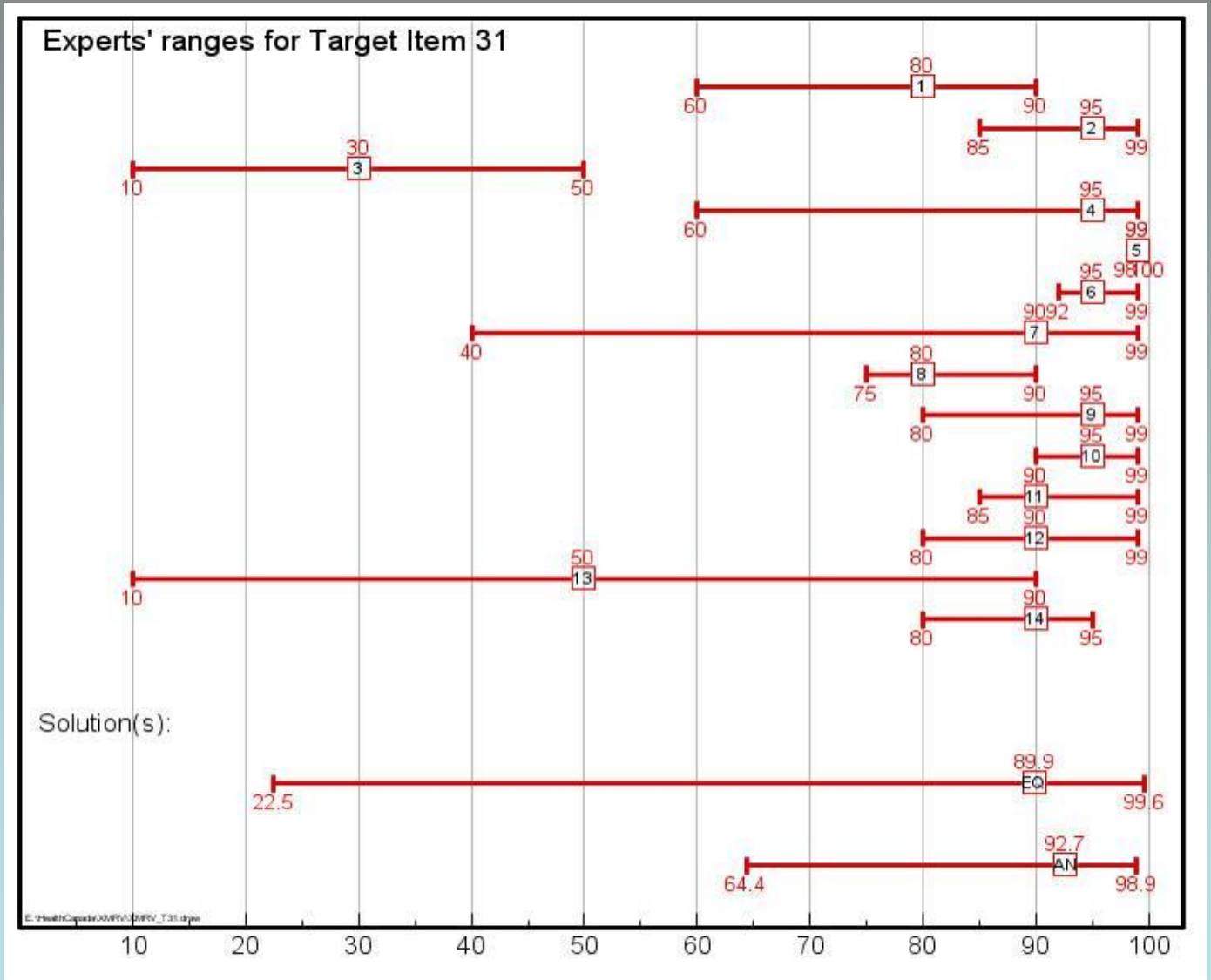


Performance Weighted Solution:

- Median: 1 in 126
- Range: 1.2-452,300

Target Question 31

What percentage of infected XMRV carriers are asymptomatic?



Expert Weighted:

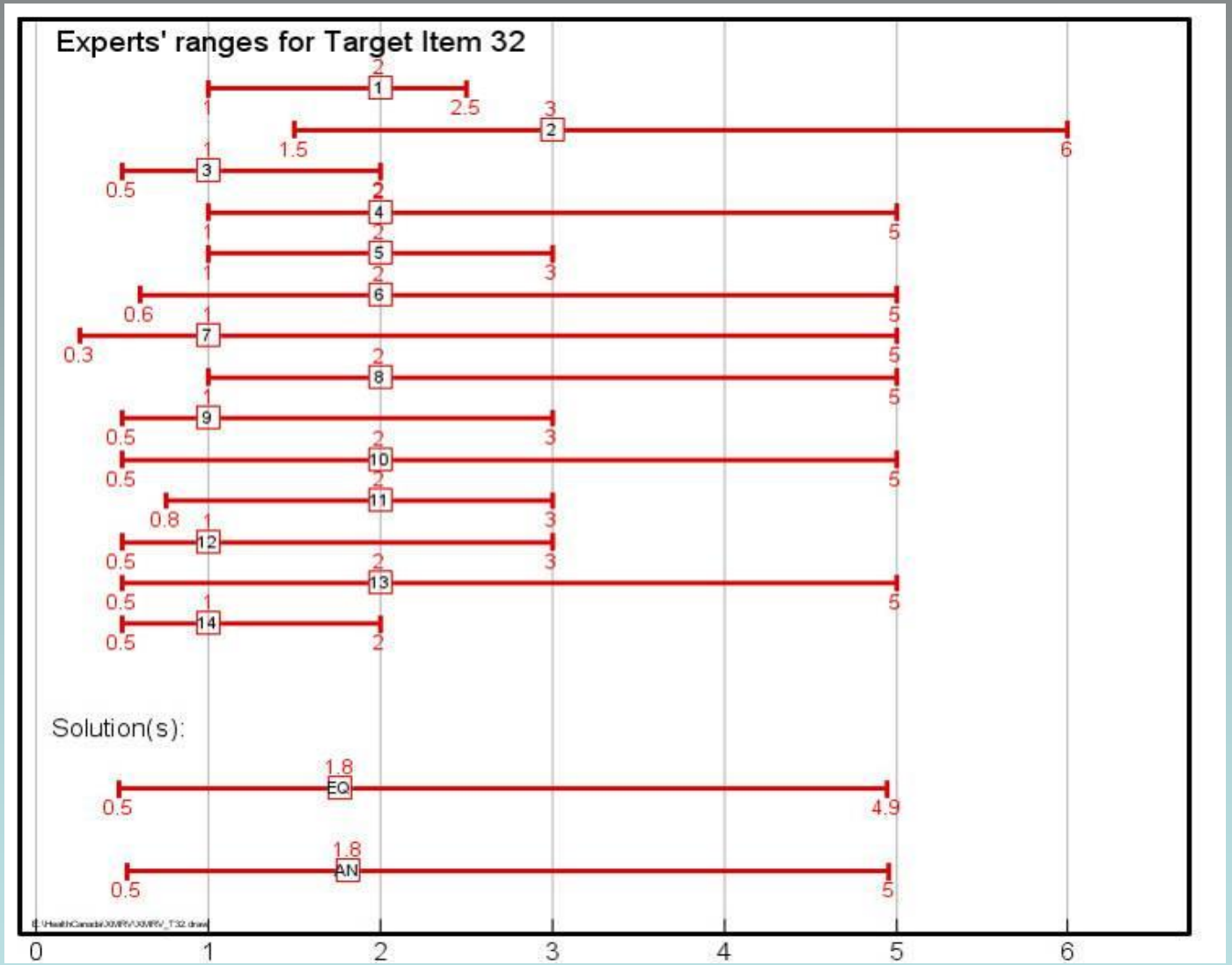
- 92.7 %
- Range: 64.4-98.9

Experts believe the majority of XMRV infections are asymptomatic.

Short right tails suggest experts are more certain that the value is higher than lower.

Target Question 32

When will the data be available to generate testing/ screening of blood donors for XMRV?



Expert Weighted:

- 1.8 years
- Range: 0.5-5.0

Experts best estimate for available data and improved techniques for XMRV testing is within 2 years (left tailed skew).

Eventual outcome …

REPORTS

No Evi Gamm Previo

Konstance Knox
John Hackett Jr.
Andreas M. Kog

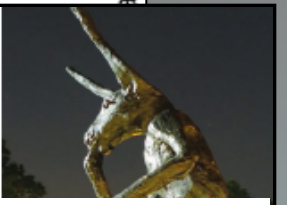
Members of the
XMRV [xenotrop
present in the b
from 61 patient
identified as XM
transcription poly
for detection of
other MLVs in t
strongly (X-MLV
and healthy con
would be unlike
laboratory reage
CFS is likely att

Xenotrop
in mice,
of being
i.e., integrated in
able to reinfect c
as the name (xe

COMMENTARY

Science sightseeing

Think again



LETT

edited by Jen

Editori

IN THE ISSUE
"Detection
patients with
al. purportin
murine leuke
67% of patie
3.7% of healt
by other inves
to detect XM
issue, we are
ing view tha
by Lombardi
research reag
gin of the ret
try of XMRV
two mouse le
imental passa
1990s. A con
analyses lead
tion with XM
early xenogra
tion of the vi
of murine-lik

PNAS PNAS PNAS

Retraction

MEDICAL SCIENCES

Retraction for "Detection of MLV-related virus gene sequences in blood of patients with chronic fatigue syndrome and healthy blood donors," by Shyh-Ching Lo, Natalia Pripuzova, Bingjie Li, Anthony L. Komaroff, Guo-Chiuan Hung, Richard Wang, and Harvey J. Alter, which appeared in issue 36, September 7, 2010, of *Proc Natl Acad Sci USA* (107:15874–15879; first published August 23, 2010; 10.1073/pnas.1006901107).

The authors wish to note the following: "Although our published findings were reproducible in our laboratory and while there has been no evidence of contamination using sensitive mouse mitochondrial DNA or IAP assays or in testing coded panels, we have the following concerns:

1. The original chronic fatigue syndrome (CFS) patient samples were of insufficient volume to distribute to other laboratories for independent confirmation.
2. Only one (1) of many laboratories has found a similar association between polytropic murine leukemia viruses (pMLV) and CFS and a careful study of 100 CFS patients (2), as well as a coded panel recently constructed by the National Heart, Lung, and Blood Institute (NHLBI) (3), have found no evidence for either xenotropic murine leukemia virus-related virus (XMRV) or pMLVs in CFS patient samples.
3. Our attempts, through collaborations, to demonstrate antibody in affected patients, to isolate the virus by culture, or to show integration sites in the human genome have failed to support the initial findings.
4. While recall of eight patients from the original cohort 15 y later showed pMLV gag sequences in seven, the copy number was very low and phylogenetic analysis showed these sequences were not direct descendents of the original dominant strains (4). Still later samples from four of these patients tested negative in the NHLBI panel. While this result could be explained by viral clearance over time, it fails to support a sustained retroviral infection in human cells.

Although a more definitive, National Institute of Allergy and Infectious Diseases (NIAID)–sponsored, coded panel of samples from 150 well-characterized and geographically diverse CFS patients and controls is being assembled for further study, in consideration of the aggregate data from our own laboratory and that of others, it is our current view that the association of murine gamma retroviruses with CFS has not withstood the test of time or of independent verification and that this association is now tenuous. Therefore, we retract the conclusions in our article."

Shyh-Ching Lo
Natalia Pripuzova
Bingjie Li
Anthony L. Komaroff
Guo-Chiuan Hung
Richard Wang
Harvey J. Alter

1. Hanson MR, et al. (2011) Detection of MLV-like gag sequences in blood samples from a New York state CFS cohort. *Retrovirology* 8(Suppl 1):A234.
2. Shin CH, et al. (2011) Absence of XMRV retrovirus and other murine leukemia virus-related viruses in patients with chronic fatigue syndrome. *J Virol* 85: 7195–7202.
3. Simmons G, et al.; Blood XMRV Scientific Research Working Group (SRWG) (2011) Failure to confirm XMRV/MLVs in the blood of patients with chronic fatigue syndrome: a multi-laboratory study. *Science* 334:814–817 10.1126/science.1213841.
4. Katzourakis A, Hué S, Kellam P, Towers GJ (2011) Phylogenetic analysis of murine leukemia virus sequences from longitudinally sampled chronic fatigue syndrome patients suggests PCR contamination rather than viral evolution. *J Virol* 85: 10909–10913.

www.pnas.org/cgi/doi/10.1073/pnas.1119641109

RETRACTION

Judgment in the face of scientific uncertainty:

the last word in rationality…

In the face of such challenges, we can extol Roger's virtues …

nature

Vol 463|21 January 2010

OPINION

A route to more tractable expert advice

There are mathematically advanced ways to weigh and pool scientific advice. They should be used more to quantify uncertainty and improve decision-making, says **Willy Aspinall**.

When a volcano became restless on the small, populated island of Montserrat, West Indies, in 1995, there was debate among scientists: did the bursts of steam and ash presage an explosive and deadly eruption, or would the outcome be more benign? Authorities on the island, a British overseas territory, needed advice to determine warning levels, and whether travel restrictions and evacuations were needed. The British government asked me, as an independent volcanologist, to help reconcile differing views within the group.

As it happened, I had experience not only with the region's volcanoes, but also with a unique way of compiling scientific advice in the face of uncertainty: the Cooke method of 'expert elicitation'. This method weighs the opinion of each expert on the basis of his or to remove it from the decision process.

Of the many ways of gathering advice from experts, the Cooke method is, in my view, the most effective when data are sparse, unreliable or unobtainable.

Rational consensus

Advice during an emergency is usually the responsibility of a chief scientist, with all the stresses that involves — including the pressure to be extremely cautious. There is a better way: pooling the opinions of a group of specialists.

There are several methods of such expert elicitation, each with flaws. The traditional committee still rules in many areas — a slow, deliberative process that gathers a wide range of opinions. This has parallels with the scientific process itself. But committees traditionally give all experts equal weight (one person, one

the Delft University of Technology in the Netherlands with his colleagues, instead produces a 'rational consensus'. To see how this works, take as an example an elicitation I conducted in 2003, to estimate the strength of the thousands of small, old earth dams in the United Kingdom. Acting as facilitator, I first organized a discussion between a group of selected experts about how water can leak into the cores of such ageing dams, leading to failure. The experts were then asked individually to give their own opinion of the time-to-failure in a specific type of dam, once such leakage starts. They answered with both a best estimate and a 'credible interval', for which they thought there was only a 10% chance that the true answer was higher or lower.

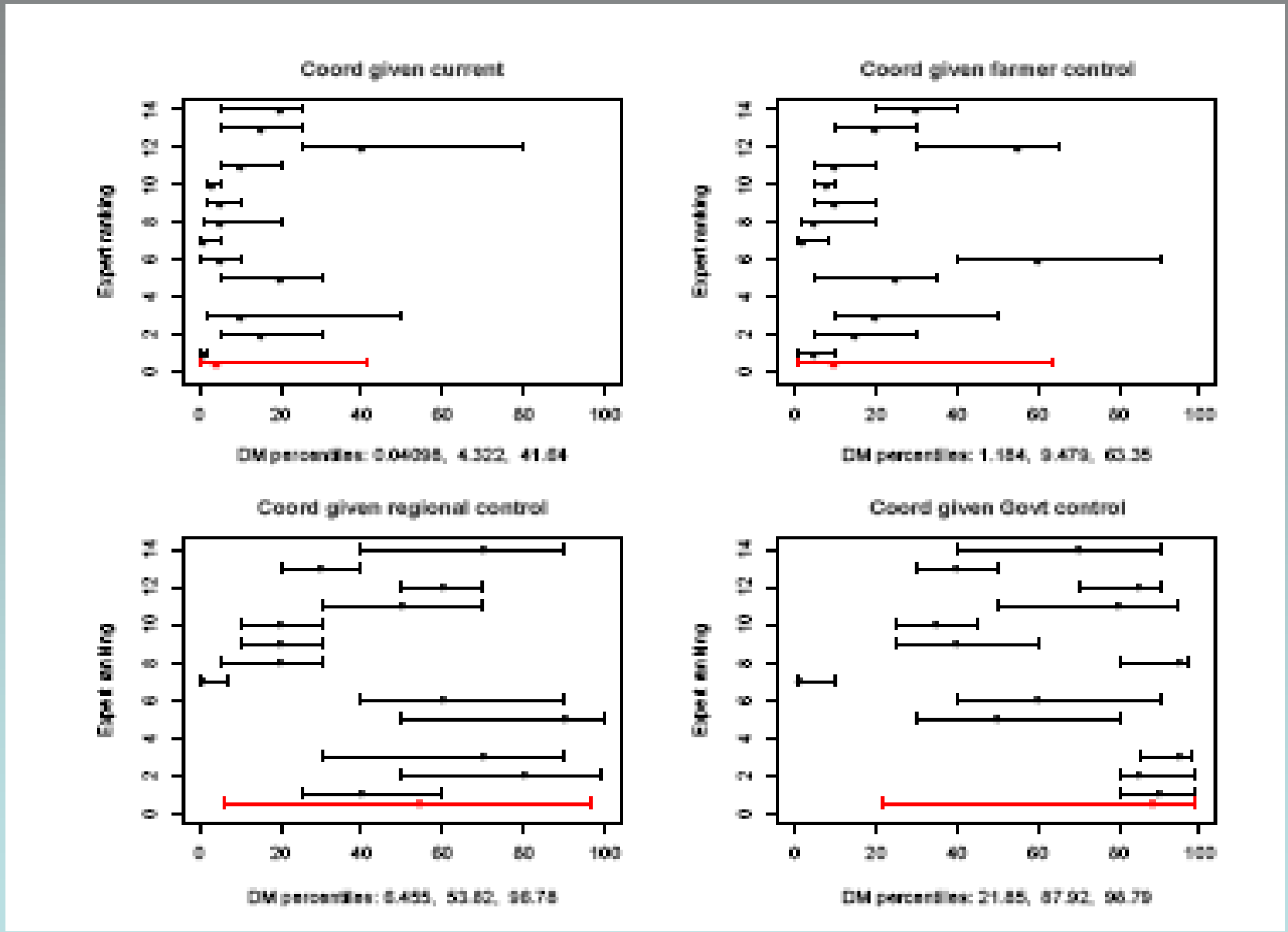
I also asked each expert a set of eleven 'seed questions', for which answers are known, so

…. and apply them to scabby sheep:

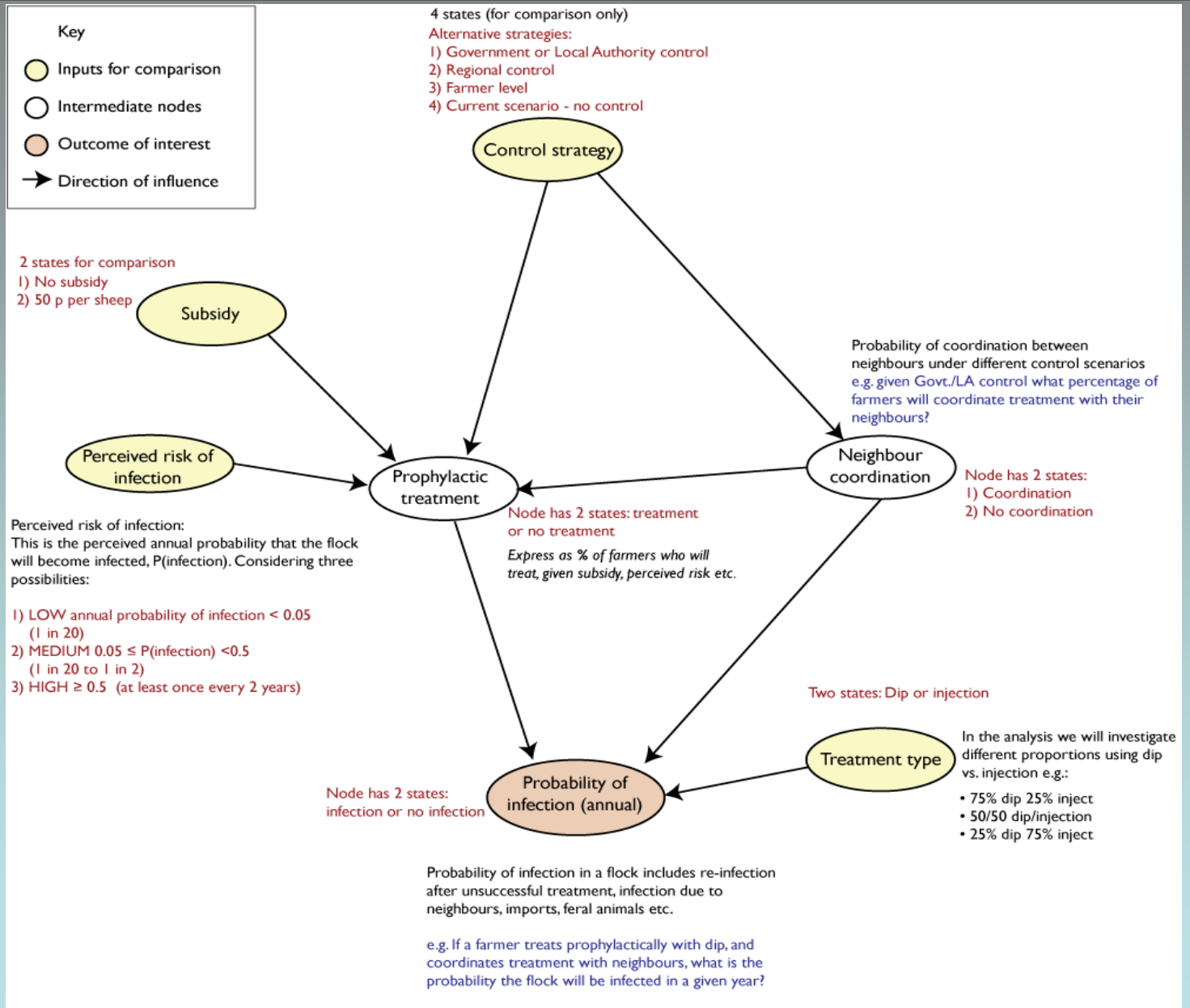
Policy options for Psoroptes ovis management

work with Thea Hincks (PhD), Jon Stone and UoBristol vets

Example range graphs for experts' judgments and weighted combination quantiles (red)



Sheep scab infection -v- policies: BBN using elicited parameters



Sheep scab elicitation BBN findings

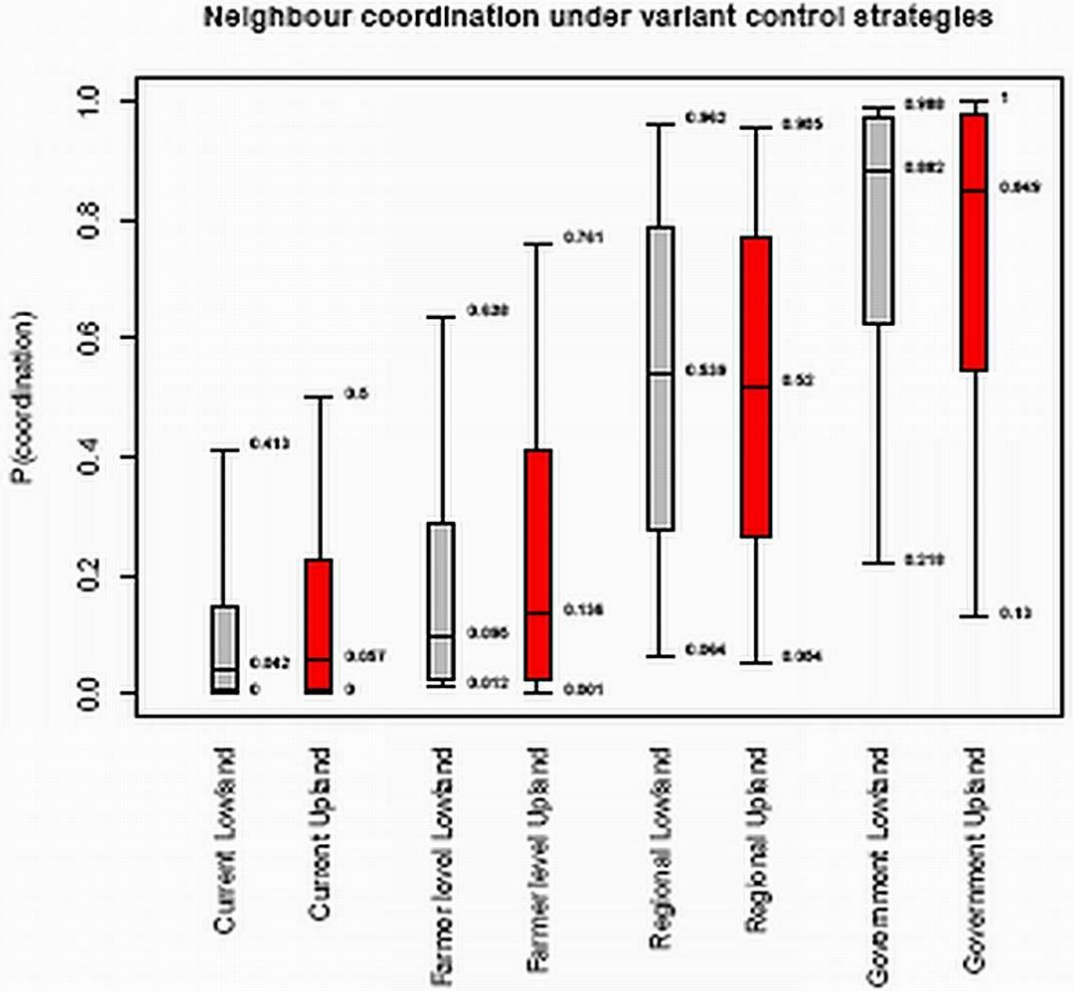


Figure 18: *Probability of neighbour coordination in upland and lowland regions, for each control scenario. The box bounds the first and third quartiles, with a line drawn to show the median value. The whiskers mark the 5 and 95 percentiles.*

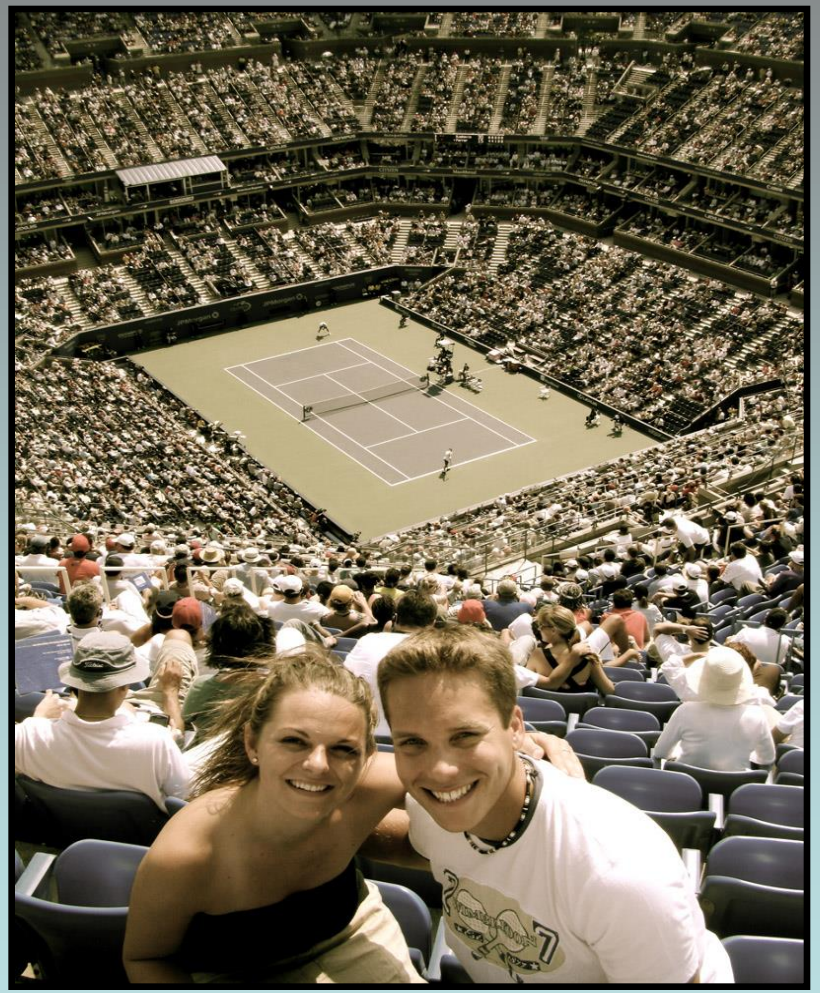
“Achieving Consensus … use in Law and Policy”

ACHIEVING CONSENSUS:
AN ANALYSIS OF METHODS TO SYNTHESIZE EPIDEMIOLOGICAL DATA FOR USE IN LAW AND POLICY

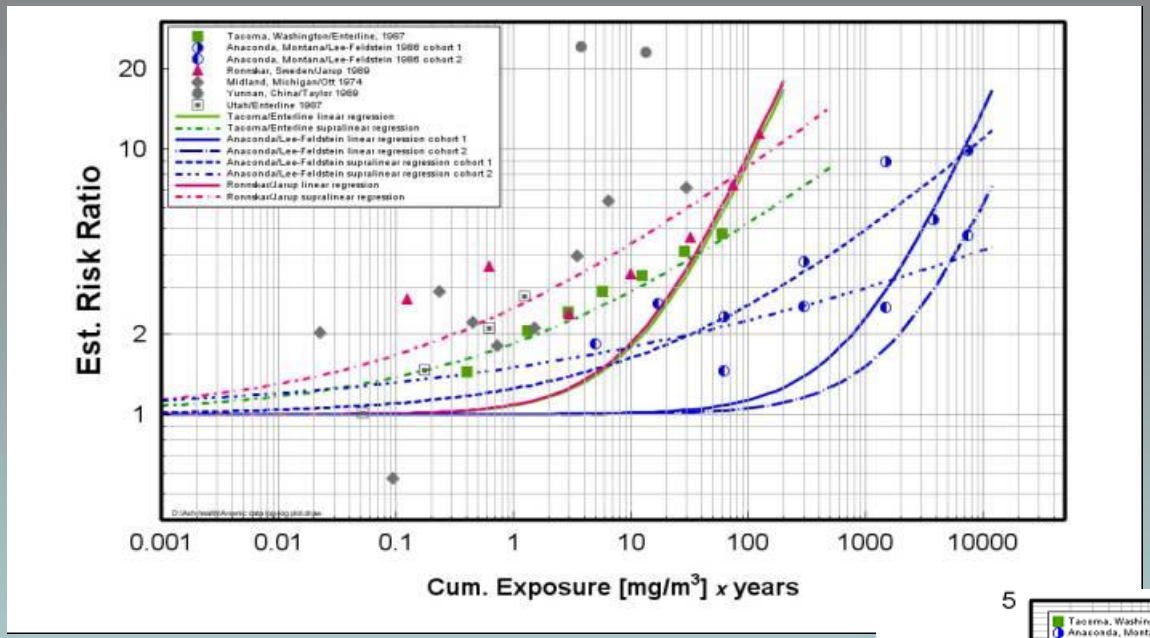
JOSEPH M. HANZICH
PEMBROKE COLLEGE

DEPARTMENT OF PUBLIC HEALTH & PRIMARY CARE
INSTITUTE OF PUBLIC HEALTH
UNIVERSITY OF CAMBRIDGE

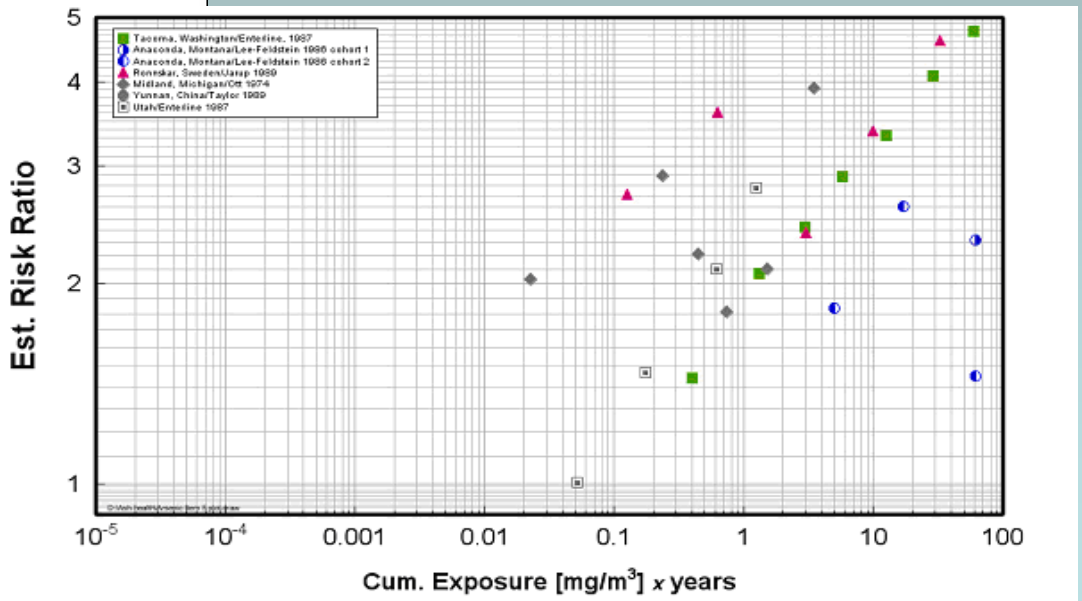
31 JULY 2007

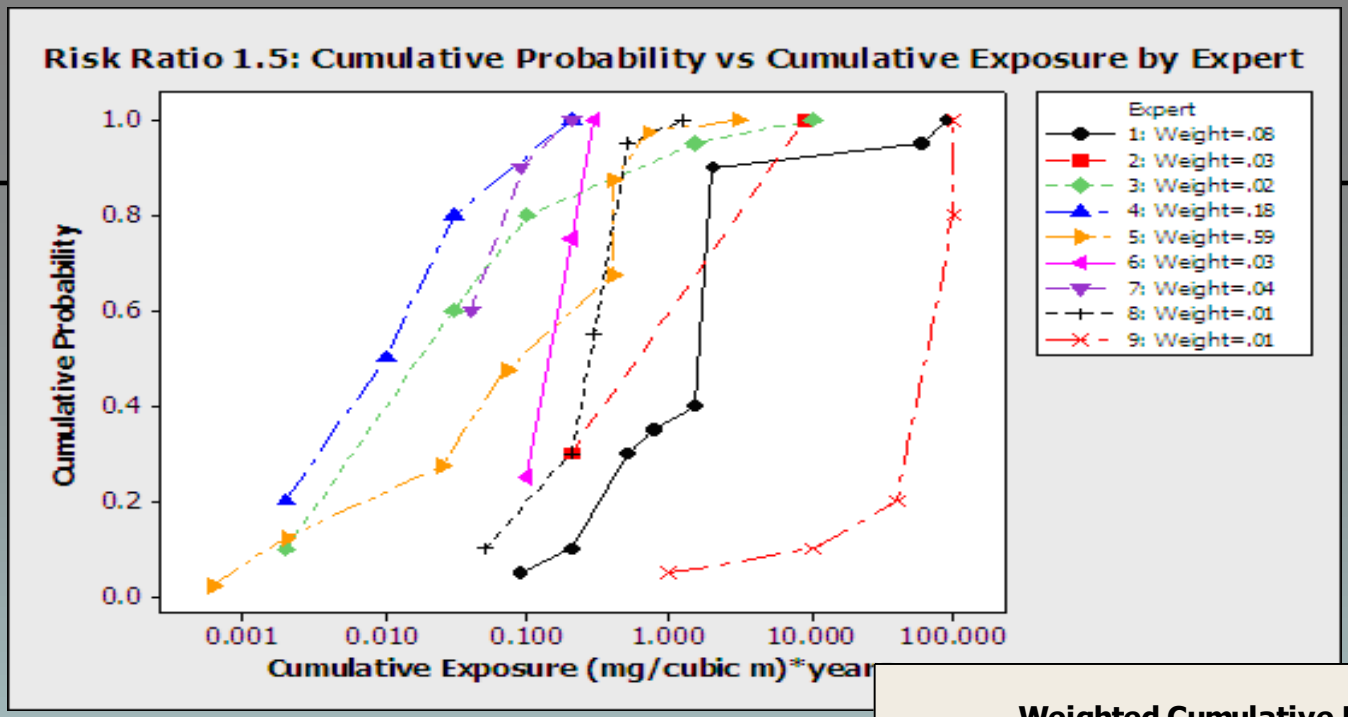


Estimating dose-response curves for cancer risk from airborne arsenic



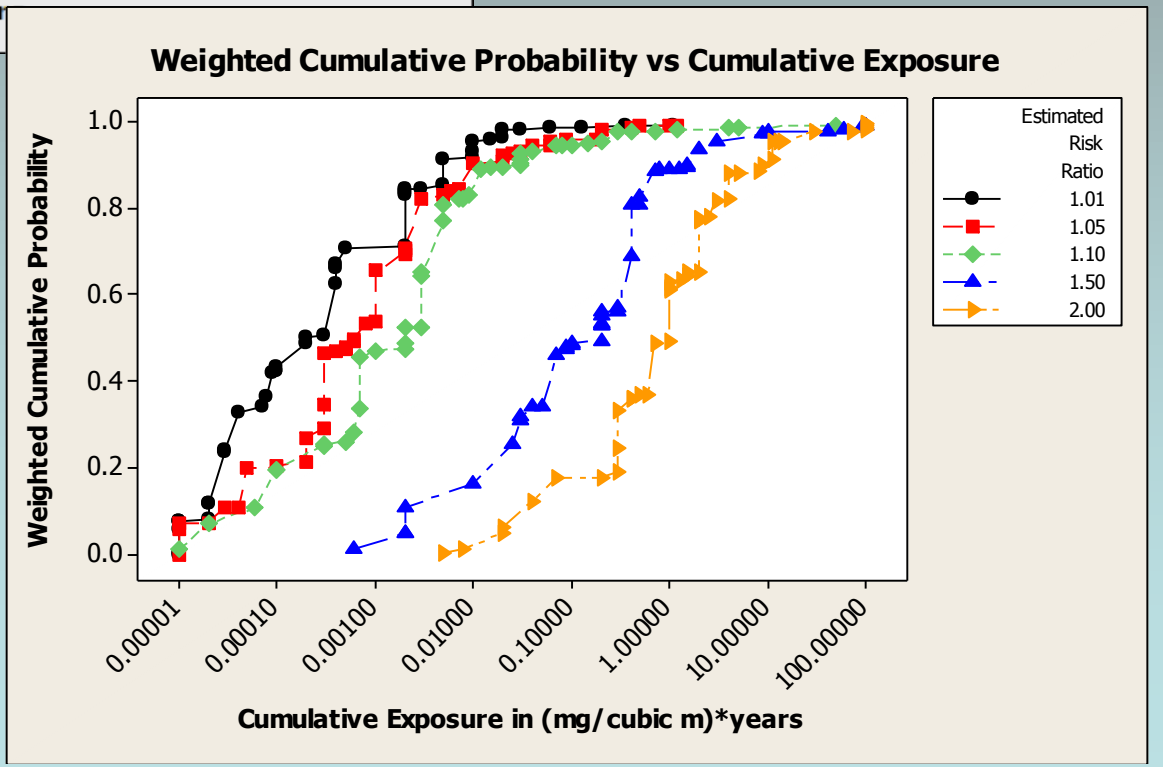
Work with the late Joey Hanzich (Cambridge University Env. Epid. MPhil 2006-07) and Dr Peter Baxter at IPH Cambridge



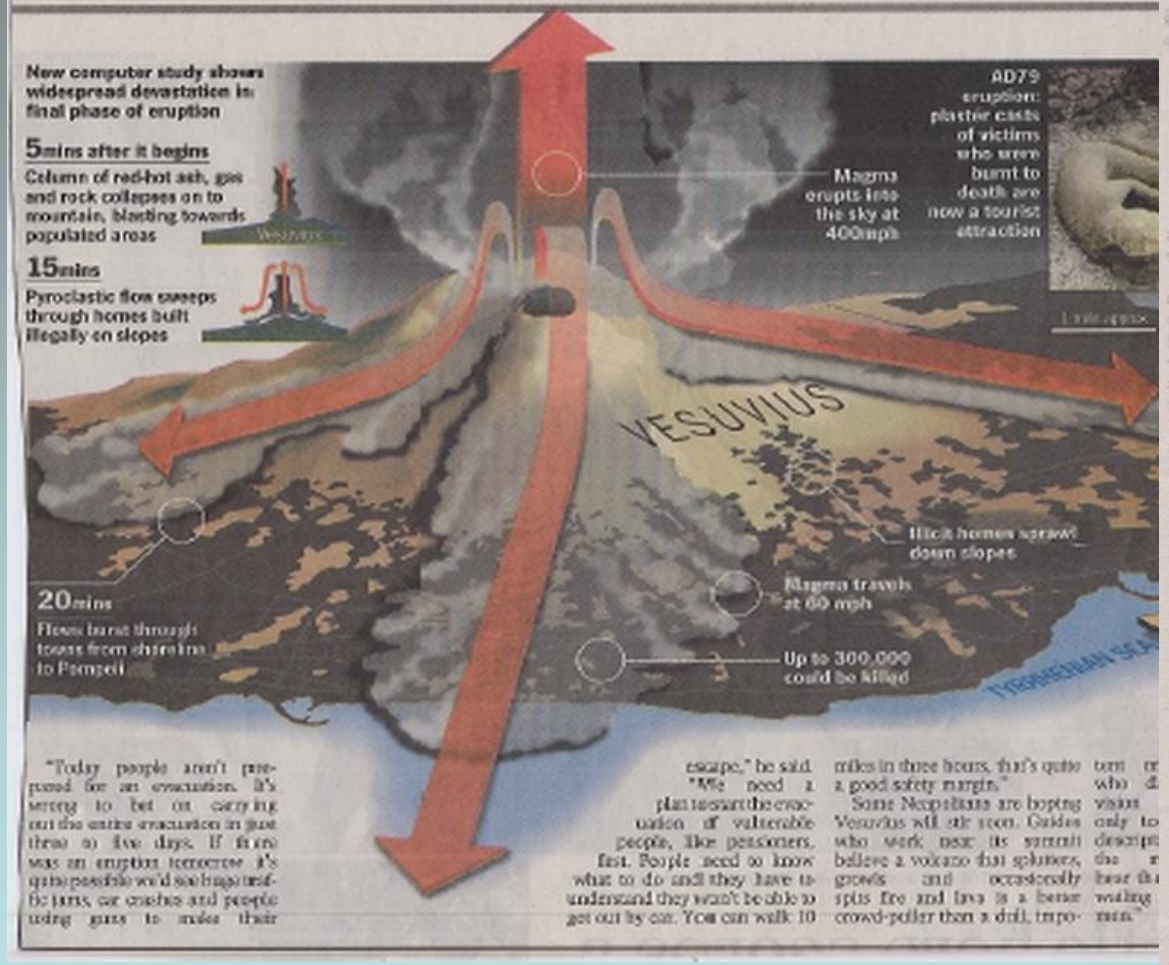


Alternative self-weighted curves from one individual expert for one risk ratio value…..

….and pooled results for group, when all combined with EXCALIBUR weights



Back to volcanoes: Vesuvius, and the future threat to Naples



"Today people aren't prepared for an evacuation. It's wrong to bet on carrying out the entire evacuation in just three to five days. If there was an eruption tomorrow it's quite possible we'd see huge traffic jams, car crashes and people using guns to make their escape," he said. "We need a plan to start the evacuation of vulnerable people, like pensioners, first. People need to know what to do and they have to understand they won't be able to get out by car. You can walk 10 miles in three hours, that's quite a good safety margin."

Some Neapolitans are hoping Vesuvius will stir soon. Guides who work near its summit believe a volcano that splutters, grows and occasionally spits fire and lava is a better crowd-puller than a dull, impo-

THE SUNDAY TIMES MARCH 4, 2007

Vesuvius blast could kill 300,000

John Follain

THE next eruption of Vesuvius could kill at least 300,000 people, nearly 20 times as many as the AD79 disaster that buried the ancient city of Pompeii, according to Italian government research.

More than half a million people live in the so-called "red zone" of 18 towns in a four-mile radius of the volcano and most would die if an evacuation could not be completed in time, the research says.

The findings are from a study by some of Europe's leading vulcanologists and public health experts, including Dr Peter Baxter of Cambridge University's Department of Public Health.

The destruction of Pompeii, the worst affected city, has inspired many books and films, including Robert Harris's 2003 bestseller, which features Pliny the Elder and which is to be adapted by the director Roman Polanski in a £100m movie.

Some 2.5m tourists visited Pompeii last year, where the people by bus from each of the 18 towns.

Professor Giuseppe Luongo of the University of Naples, a former director of the Vesuvius Observatory which monitors the volcano, believes plans are inadequate and local people are ill-informed about them.

Expert elicitations

Vesuvius last awoke with a small blast in 1944. A large eruption could unleash incendiary avalanches and ash that would threaten millions of people.

EUROPE'S TICKING TIME BOMB

Vesuvius is one of the most dangerous volcanoes in the world — but scientists and the civil authorities can't agree on how to prepare for a future eruption.

BY KATHERINE BARNES

It starts with a blast so strong that a column of ash and stone rockets 40 kilometres up into the stratosphere. The debris then drops to Earth, pelting the surface with boiling hot fragments of pumice and covering the ground with a thick layer of ash. Roofs crumble

small eruption in 1944, but recent studies suggest that Vesuvius could be more dangerous than previously assumed, which has prompted a vigorous debate about the risk and scale of

interpret this layer as an active magma reservoir[1], which could produce large-scale 'plinian'-style explosions — named after Pliny the Younger, who described the AD 79 eruption.

The first rumblings of activity at Vesuvius could come weeks to years before an eruption,

BETTMANN/CORBIS

….. Nature, 12 May 2011

Cooke's Classical Model used to characterize hazards and risks for various possible future eruption scenarios at Vesuvius

Neri, A. et al. (Editors) (2008). Evaluating explosive eruption risk at European volcanoes. J. Volcanol.Geotherm. Res. Spec. Vol. 178.

Aspinall WP, Woo G, Voight B, Baxter PJ. (2003). Evidence-based volcanology: an application to volcanic crises. J. Volcanol.Geotherm. Res. 128: 273-285.

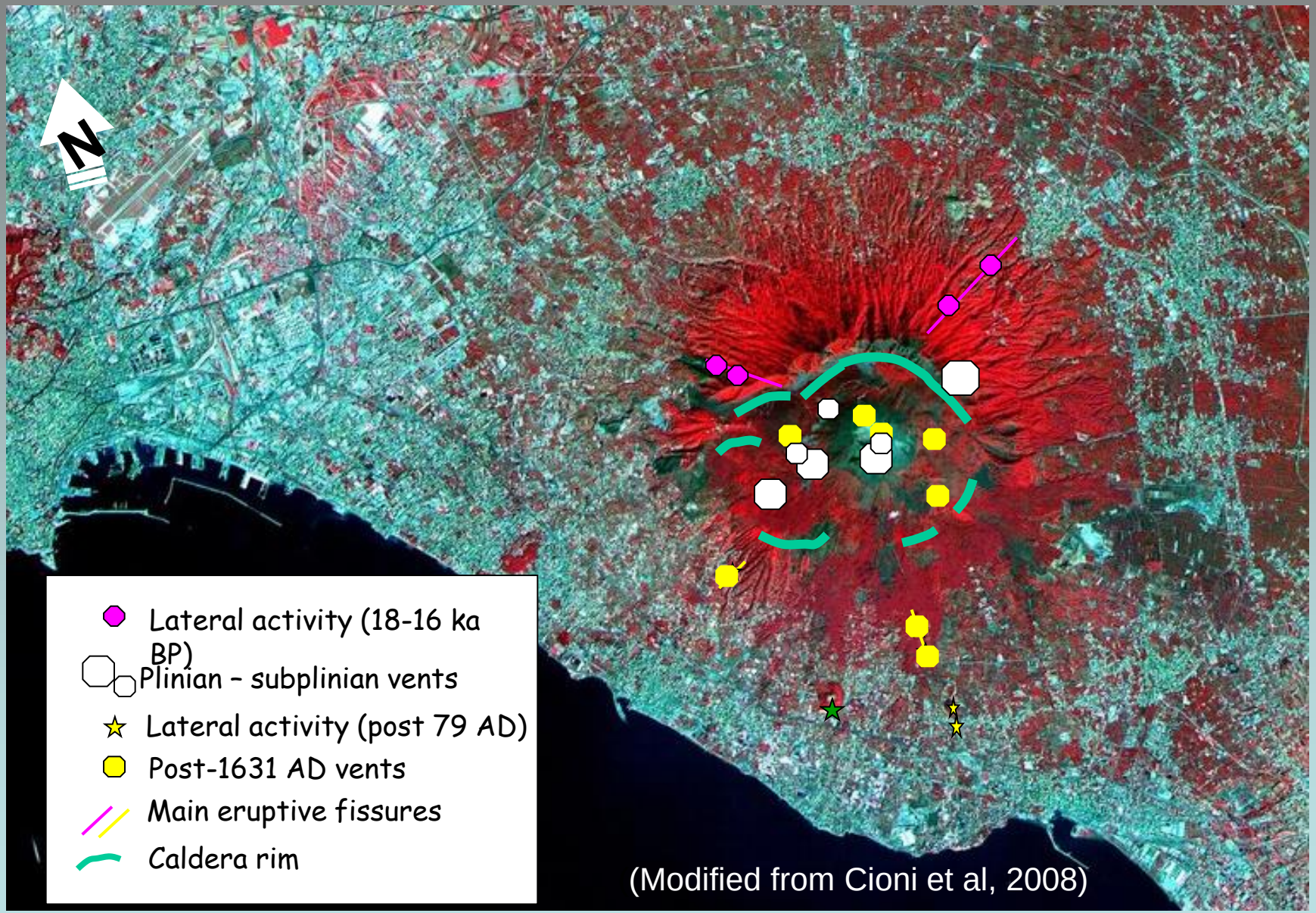
… and many more

Main uncertainty sources in modelling PDC dynamics

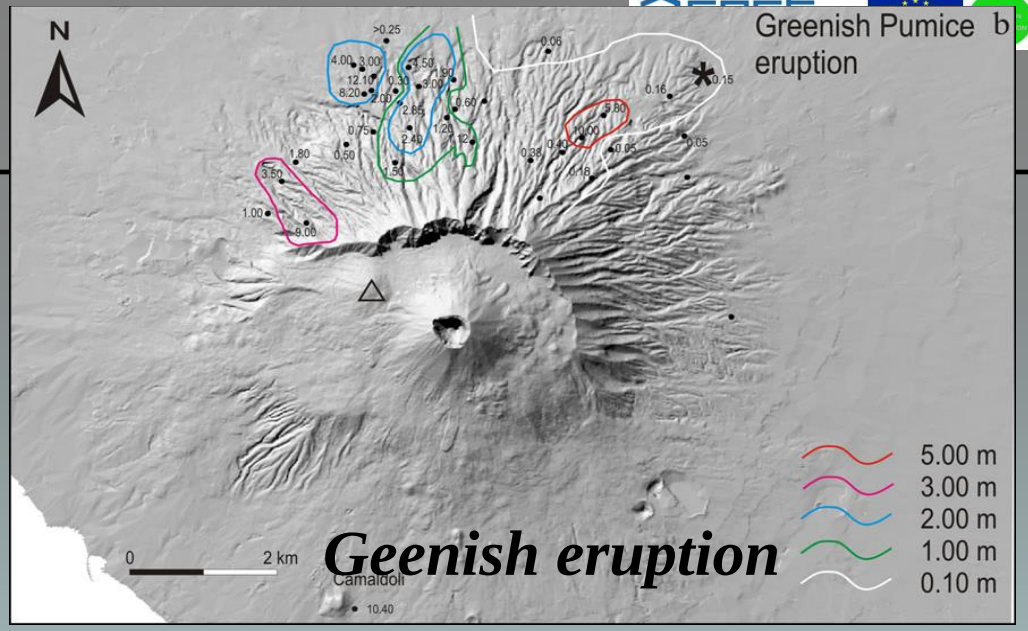
One eruption ‘magnitude’ scenario (Sub-Plinian):

- Variability of the mass flow rate (2 - 8 x 10^7 kg/s)
- Variability of collapse mechanism (column/caldera collapse, partial/total column collapse)
- Variability of flow properties and emplacement (dilute vs dense PDC)
- Variability of volcano topography (past, present topography, and syn-event changes)
- Variability of vent location

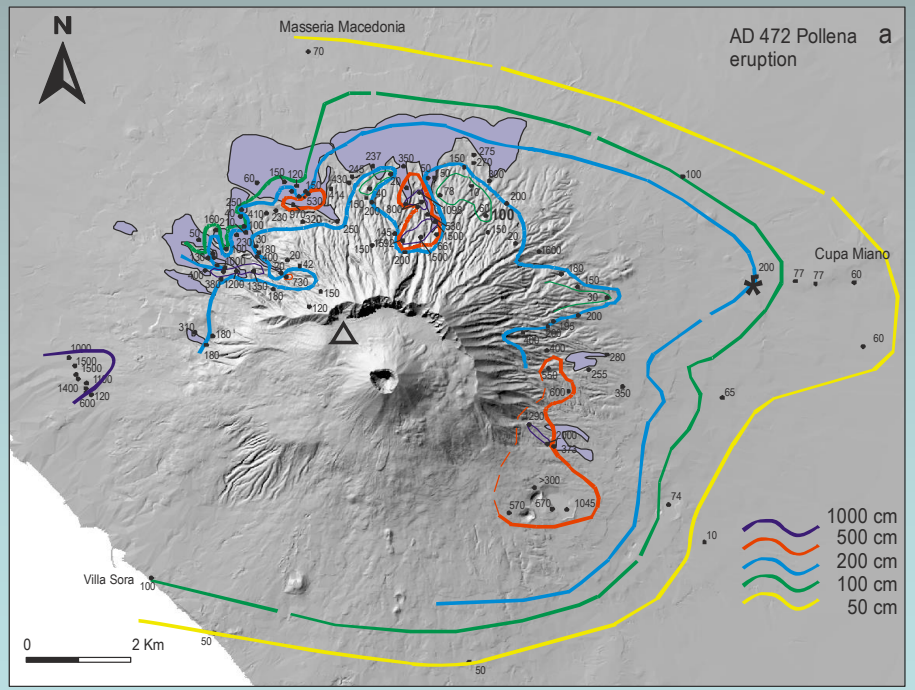
Distribution of past vent locations

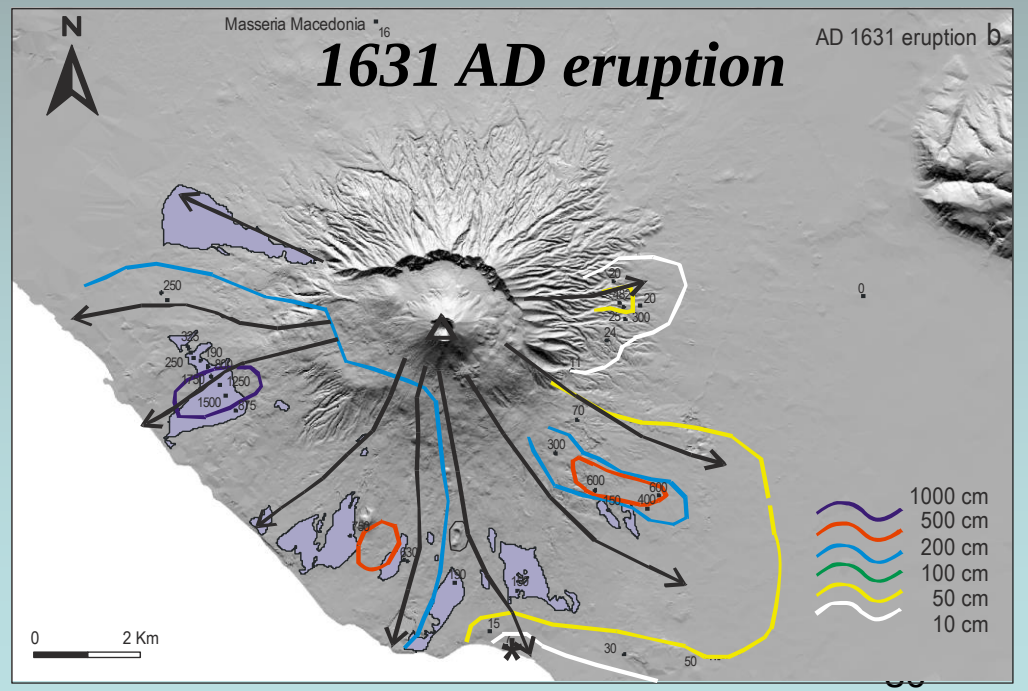


Sub-Plinian I PDC distributions by field reconstruction



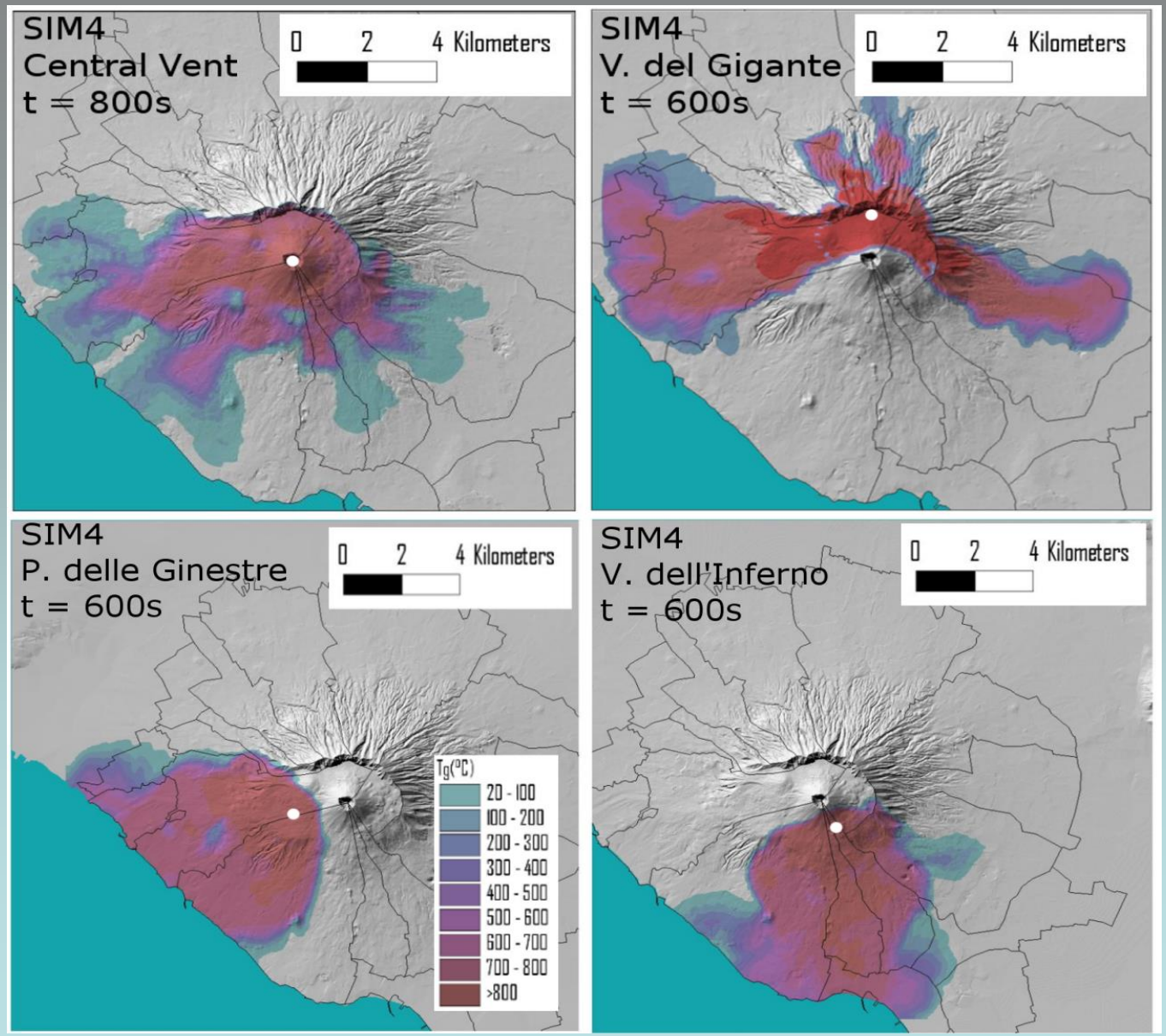
472 AD Pollena eruption





(Gurioli et al., 2010)

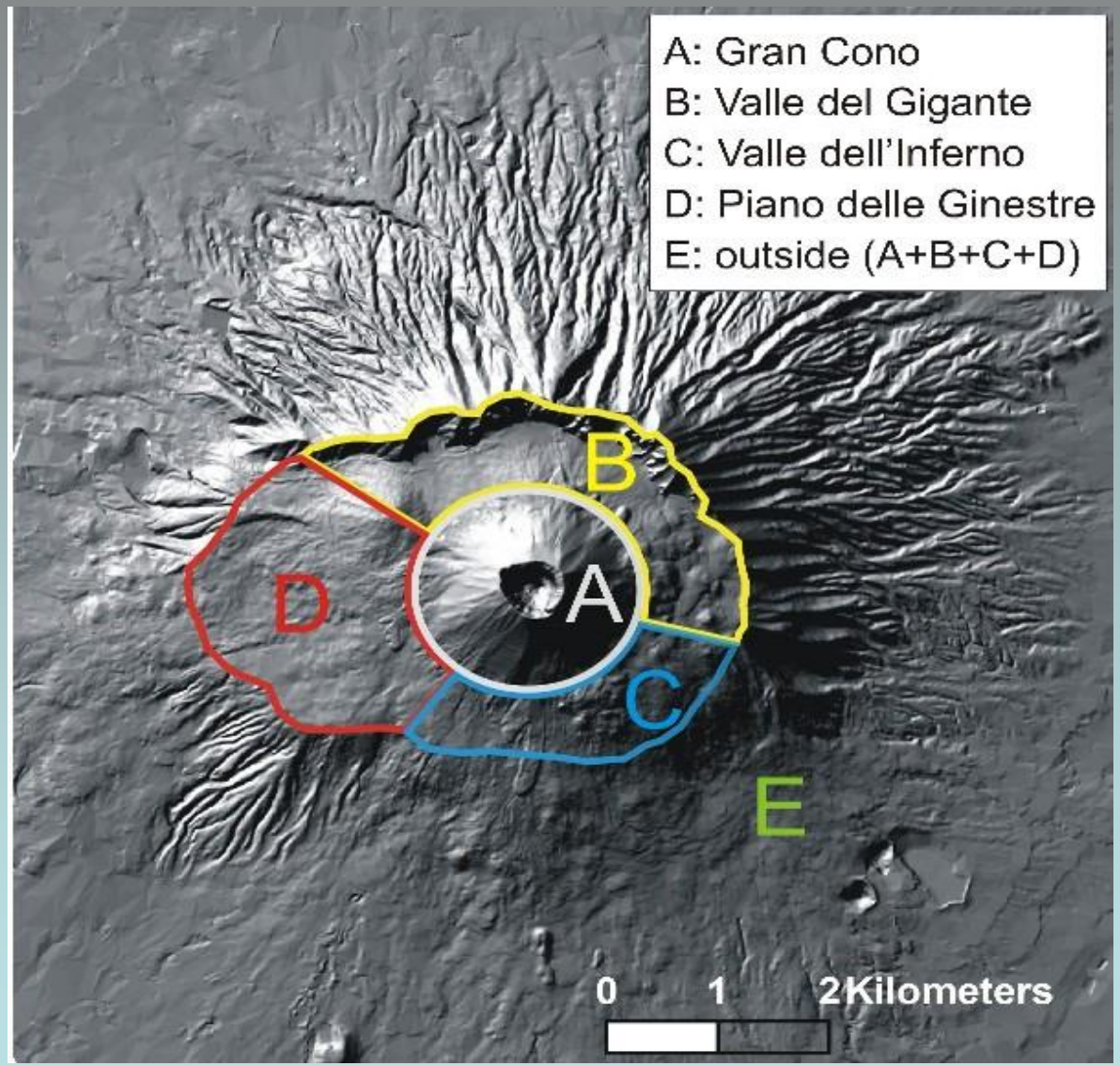
Numerical simulation of near-total collapse scenarios
Temperature distribution



(Esposti Ongaro et al.

In preparation)

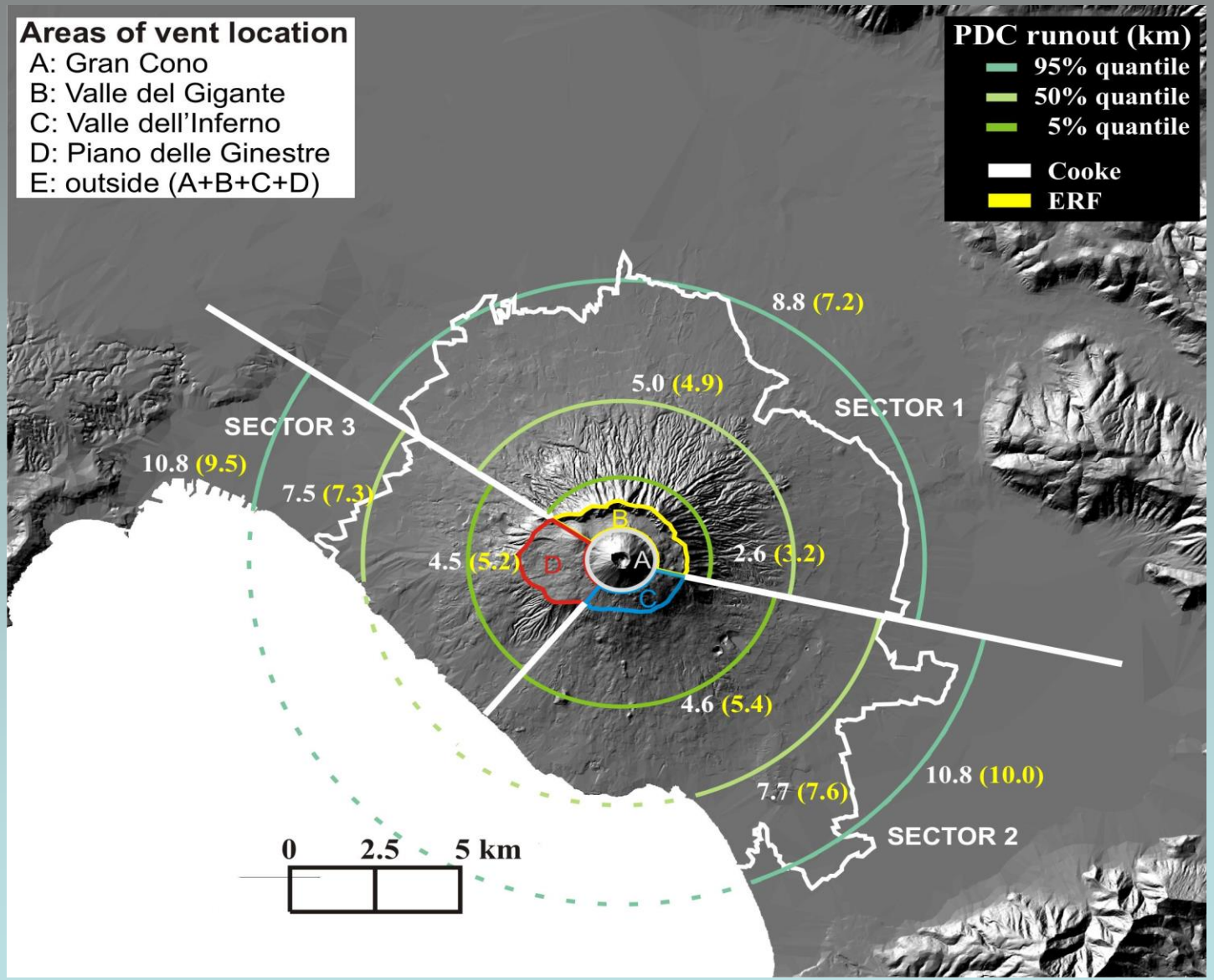
Zones defining potential areas for new vent opening



Elicited ranges of probabilities of vent opening [5%ile, 50%ile, 95%ile]

Area A (Gran Cono):	**[22, 41, 94]**
Area B (Valle del Gigante):	**[2.4, 20, 62]**
Area C (Valle dell'Inferno):	**[2.3, 20, 62]**
Area D (Piano delle Ginestre):	**[2.3, 18, 62]**
Area E (Outside A+B+C+D):	**[0.01, 1, 17]**

Vesuvius: probability map for pyroclastic flow total runout distance



Famous last words….

of a volcanologist:

Experts, expert judgment, elicitation, and the law?

HOME NEWS SPORT FINANCE COMMENT BLOGS CULTURE TRAVEL LIFESTYLE

UK World Politics Obituaries Education Earth **Science** Defence Health News

Science News Space Roger Highfield Dinosaurs Evolution Steve Jones Science Pictur

HOME » SCIENCE

The legal aftershocks of the earthquake in L'Aquila

Science is in the dock in Italy as local witnesses finally confront the earthquak
on trial for manslaughter who, it is alleged, failed to warn them of the risks.

Nuns walk past the ruins of a building after the earthquake on April 6, 2009 in L'Aquila, Italy Photo: AFP/GETTY IMAGES

Share:

Recomm

Recomm

Share

0

Science
Italy »
Science Ne
Environm
Geology »

IN SCIENCE

By Michael Day
7:30AM GMT 22 Nov 2011

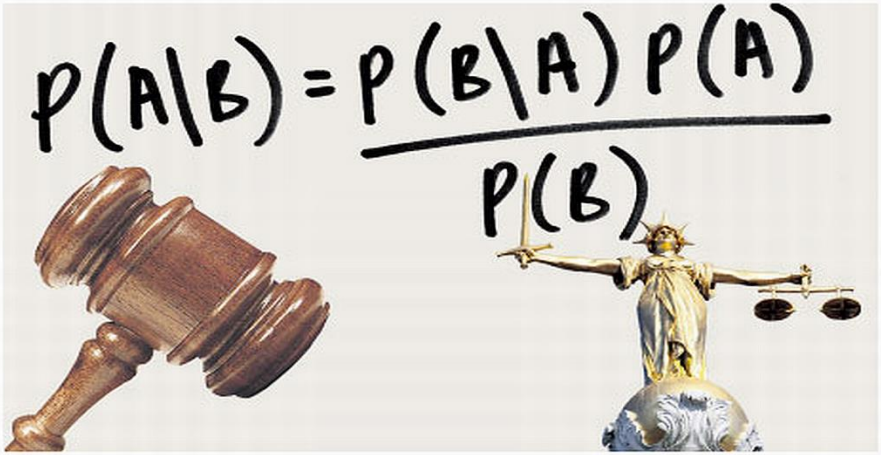
Telegraph 22 Nov 2011

News › Law

A formula for justice

Bayes' theorem is a mathematical equation used in court cases to analyse statistical evidence. But a judge has ruled it can no longer be used. Will it result in more miscarriages of justice?

Angela Saini
guardian.co.uk, Sunday 2 October 2011 21.30 BST
Article history



Bayes' theorem. Photograph: guardian.co.uk

Challenges to expert judgment elicitation

The Harvard study on Kuwait's First Gulf War reparations claim

- **Health effects claim based on expert elicitation: ~ 35 deaths**

Individual experts' best mortality estimates:

13, 32, 54, 110, 164, 2874

Equal Weights (82 deaths;

90% conf.: 18 to 400)

Performance Weights (35 deaths;

90% conf.: 16 to 54)

The judicial decision of the UN Commission eventually rejected the admissibility of this form of evidence: "…not actual data….."

Challenges to expert judgment elicitation

WORLD VIEW *A personal take on events*

N. TUCKER

Check your legal position before advising others

Next week's trial of seismologists in Italy highlights the risks to scientists who offer public advice. **Willy Aspinall** *considers what can be done.*

SCIENTISTS IN **SENSITIVE** SITUATIONS SHOULD THINK CAREFULLY ABOUT THEIR USE OF **SOCIAL MEDIA.**

Nature (2011) Vol 477, page 251

What next on my elicitations agenda?

- Railway bridge scour
- Structural fragility curves for quake and fire impacts
- WHO burden of food-borne disease – pathogen attributions
- Japanese radwaste siting
- Climate influence on extreme storms in Europe

First ever probabilistic expert elicitation in Japan: tectonic and volcanic hazard factors for radwaste repository siting

Ice sheet melting – projected contributions to future sea-level rise

nature climate change

ARTICLES

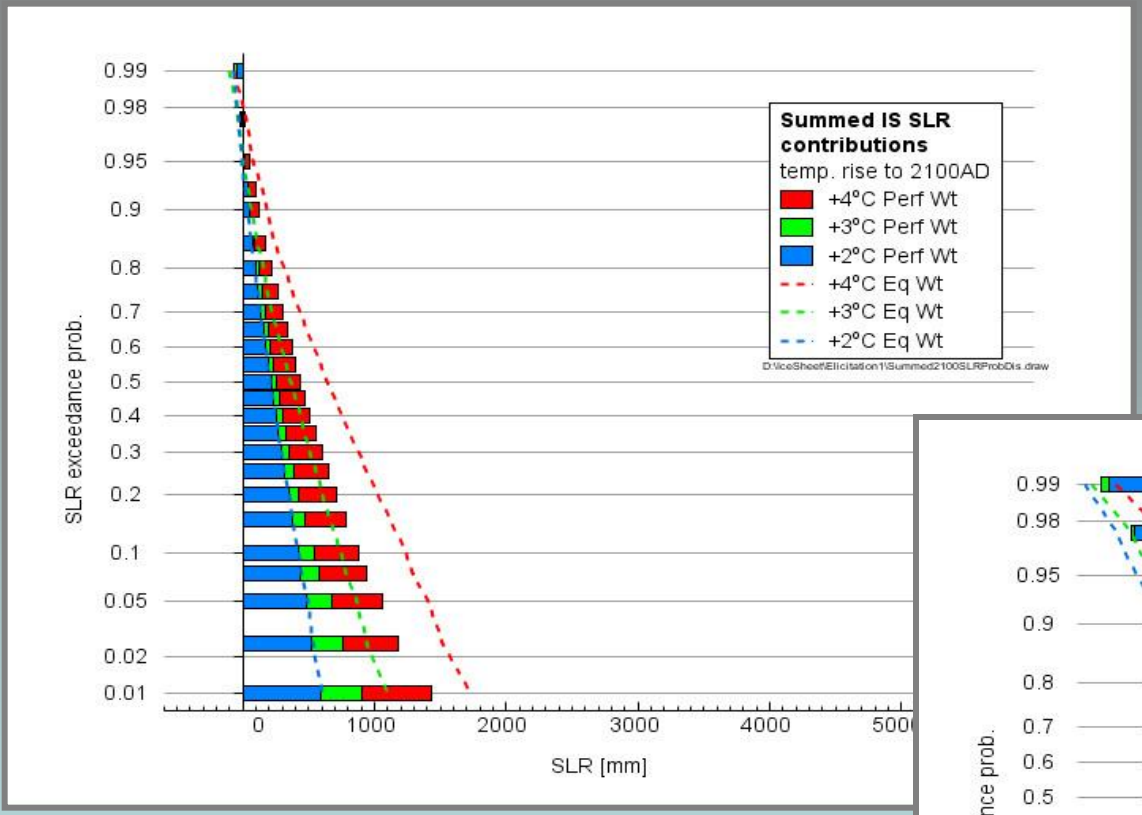
PUBLISHED ONLINE: 23 DECEMBER 2012 | DOI: 10.1038/NCLIMATE1778

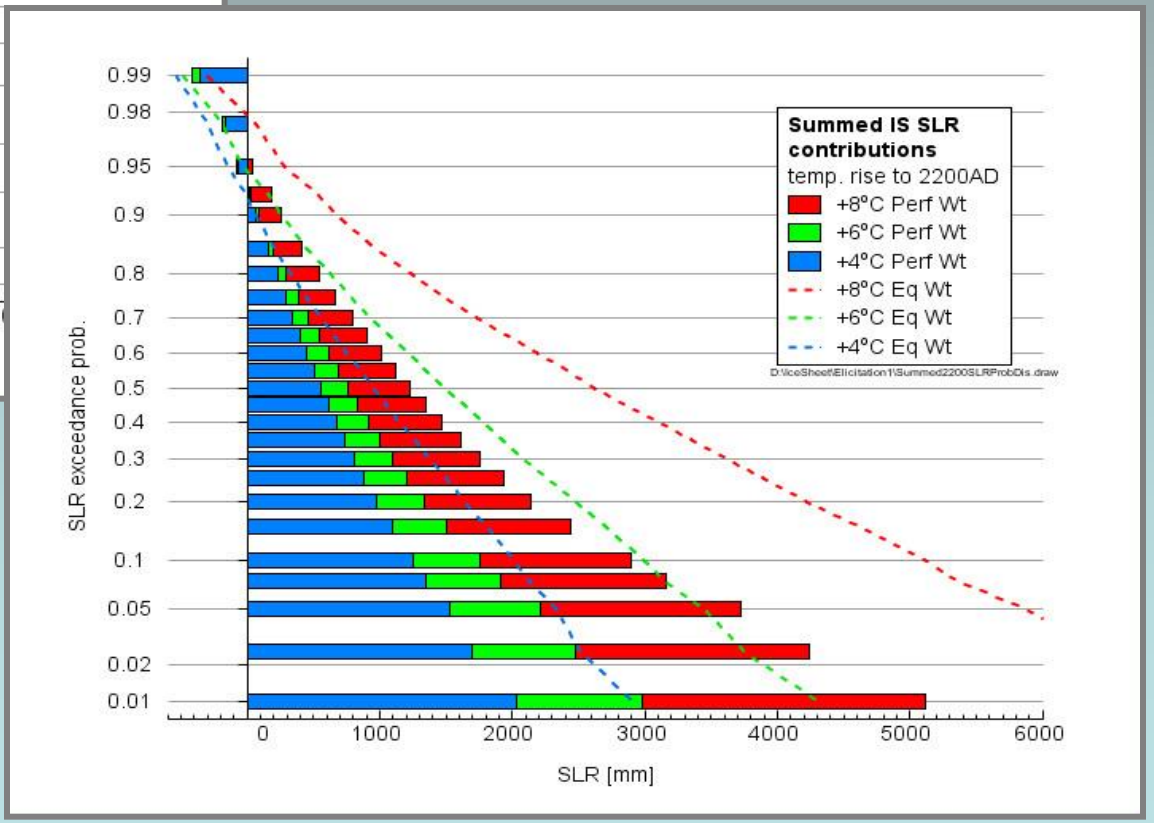
An expert judgement assessment of future sea level rise from the ice sheets

J. L. Bamber[1]* and W. P. Aspinall[2]

A major gap in predictive capability concerning the future evolution of the ice sheets was identified in the Fourth Assessment Report (AR4) of the Intergovernmental Panel on Climate Change. As a consequence, it has been suggested that the AR4 estimates of future sea-level rise from this source may have been underestimated. Various approaches for addressing this problem have been tried, including semi-empirical models and conceptual studies. Here, we report a formalized pooling of expert views on uncertainties in future ice-sheet contributions using a structured elicitation approach. We find that the median estimate of such contributions is 29 cm—substantially larger than in the AR4—while the upper 95th percentile value is 84 cm, implying a conceivable risk of a sea-level rise of greater than a metre by 2100. On the critical question of whether recent ice-sheet behaviour is due to variability in the ice sheet–climate system or reflects a long-term trend, expert opinion is shown to be both very uncertain and undecided.

Pooled expert judgements on combined ice-sheet contributions to sea-level rise: 2100AD; 2200AD





Latest news from our US correspondent on evidence for climate change …….

What happens when experts differ …..