Structured expert judgment with the Classical Model – experiences of an NBD facilitator

Willy Aspinall
Harry Enfield’s TV character
• Summary of this talk

Personal experiences of challenges to the application of the Classical Model, rather than successes ……

…… starting with a big missed opportunity.
The Icelandic volcanic ash crisis, 2010
An Invitation from the President of Iceland
Olafur Ragnar Grimsson

In the middle of the Atlantic, at the crossroads of European weather systems, the youngest country on Earth has recently reminded the world how, through a combination of a volcanic eruption and enormous quantities of glacial ice, the forces of nature can produce huge ash clouds, impacting on nations far away, even disrupting for a while the travel plans of millions in different continents. For the first time since modern aviation became the pillar of international transport, a volcanic eruption has reminded all of us that our technological and economic systems must take into account the will of nature and how the Earth is capable of ultimate surprises.

Aviation stakeholders must now understand and adapt to this kind of potential disruption. Eyjafjallajökull has become a symbol of this situation, and thus it is appropriate to invite aviation leaders, experts and policy-makers to assemble at the recently established Keilir Aviation Academy at Keflavík International Airport, to discuss the lessons learned and how technology, rules, regulations and aviation training must be improved.

For us in Iceland it is important to make a contribution of this type to the constructive international dialogue that is now needed. Therefore we invite you all to participate in the Atlantic Conference on Eyjafjallajökull and Aviation. It will certainly be a worthwhile experience.

[Signature]
“Safe to Fly – Chart” presentation by Rolls-Royce at Atlantic Conference.

Note implied quantitative uncertainties on incidents and test data.
Volcanic Ash and Aircraft Engines

BATA Volcanic Ash Workshop – 15 October 2013

Rory Clarkson
Engine Environmental Protection
Rolls-Royce

Follow-on presentation in 2013 ........
Apart from addition of photos, essentials of graph are unchanged and, as far as is known, no formal development of quantifications of uncertainty.
My view

- *De facto* (if not *de jure*) flight operational risk ‘standard’ for loss of airworthiness safety is no higher than 10⁻⁸ per sector.

- One can conjecture plausible scenarios with consequence factors:
  - Plane passes inadvertently through ash for more than a few minutes (e.g. night-time)
  - Actual ash concentration in plume is higher than forecast/monitored
  - Dispersion model fails to forecast wind bringing two plumes together
  - Modern engine technology vulnerability is higher than last generation engines

The probabilities and hence joint scenario probabilities are ill-defined – even unknown. But if jointly they >> 10⁻⁸, then the “normal” flight safety margin is eroded.

No-one could, or would, say how the decision to fly would be defended *post hoc*, in the absence of any uncertainty quantification……
As far as I know (and I was involved), no effort was made to quantify all related uncertainties with structured expert judgment, some parties did not even want to discuss this option. Thus I believe the basis of the “deal” was flawed.
Meanwhile, others were ready to make money on the uncertainties!
• Current modelling capabilities

The Barcelona Supercomputing Center (BSC) FALL3D volcanic ash dispersal model. Aim is to merge model forecasts with ATM data (airports, routes, FIRs and flights) at Volcanic Ash Advisory Centres to evaluate impacts based on user-defined criteria: e.g. concentration threshold and maximum engine dose.

BUT, as far as is known, the latter and their uncertainties have not been established (at least publicly).
However, as a consequence of the Eyjafjallajökull episode, work was commissioned on potential hazards from a Laki-type eruption for the UK National Risk Register, with extensive SEJ on volcanological uncertainties:

**Executive Summary**

In response to the recent introduction of large, long-lasting gas-rich volcanic eruptions to the UK National Risk Register (risk H55) a modelling project has been conducted to improve our understanding of potential hazards to the UK from such an eruption on Iceland. A precautionary “reasonable worst case” eruption scenario based on the 1783-1784 CE Laki eruption has been determined using the results of an expert elicitation of scientists. This scenario has been simulated 80 times using two different atmospheric chemistry and transport models (NAME and EMEP4UK) over 10 years of meteorology (2003-2012).

Authors: Claire Wilham1, Willy Aspinall2, Christine Braban3, Jane Hall1, Sue Loughin1, Anja Schmidt1, Massimo Vismara2, Bill Beasley2, Matthew Holt3, Evgenia Ilyinskaya3, John Kentisbeear3, Elin Roberts3, Ed Rowe4.
Other failures (by me) to get the Classical Model engaged:

Major London insurer of geo-political risks in Russia

Big pharma company – drug development selection criteria

Lesson: need “buy in” from senior management
An unfinished elicitation: risks from arsenic in the air

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

JOSEPH M. HANZICH
PENBROKE 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
Example self-weighted curves from one individual expert for one risk ratio value.....

.... these look a noisy mess, but when combined with Classical Model weights a rational pattern emerges.
A salutary case

Big news!

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


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...
XMRV Expert Elicitation Workshop

International panel in Ottawa, Canada
# Target Question Grouping

<table>
<thead>
<tr>
<th>Questions</th>
<th>Subject Area</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-7</td>
<td>Prevalence</td>
</tr>
<tr>
<td>8-11</td>
<td>Risk Parameters</td>
</tr>
<tr>
<td>12-15</td>
<td>Latency</td>
</tr>
<tr>
<td>16-22</td>
<td>Routes of Transmission</td>
</tr>
<tr>
<td>23-25</td>
<td>Risk Mitigation</td>
</tr>
<tr>
<td>26-30</td>
<td>Disease Relationships (causal and non-causal)</td>
</tr>
</tbody>
</table>
Target Questions 1, 3-6

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)

Weighted DM:

- 1 in 126
- Range: 1.2-452,300
Prevalence: Target Questions 1, 3-6

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)

<table>
<thead>
<tr>
<th>Country</th>
<th>Weighted DM median</th>
<th>DM 90% Credible Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>Canada</td>
<td>1 in 335</td>
<td>1 in 12 → 1 in 305,500</td>
</tr>
<tr>
<td>USA</td>
<td>1 in 280</td>
<td>1 in 12 → 1 in 305,500</td>
</tr>
<tr>
<td>UK</td>
<td>1 in 450</td>
<td>1 in 12 → 1 in 305,500</td>
</tr>
<tr>
<td>France</td>
<td>1 in 450</td>
<td>1 in 12 → 1 in 305,500</td>
</tr>
<tr>
<td>First author, country</td>
<td>Journal, date</td>
<td>Patients positive for XMRV?</td>
</tr>
<tr>
<td>-------------------------------</td>
<td>---------------------------------------------------</td>
<td>----------------------------</td>
</tr>
<tr>
<td>Lombardi, USA</td>
<td>Science, October 2009</td>
<td>Yes (67%)</td>
</tr>
<tr>
<td>Erlwein, UK</td>
<td>PLoS One, January 2010 &amp; March 2011 (re-analysis)</td>
<td>No</td>
</tr>
<tr>
<td>van Kupplweid, Netherlands</td>
<td>British Medical Journal, February 2010</td>
<td>No</td>
</tr>
<tr>
<td>Groom, UK</td>
<td>Retrovirology, February 2010</td>
<td>No</td>
</tr>
<tr>
<td>Swizer, USA</td>
<td>Retrovirology, July 2010</td>
<td>No</td>
</tr>
<tr>
<td>Lo, USA</td>
<td>Proc Natl Acad Sci, August 2010</td>
<td>No (but 86.5% MLV)</td>
</tr>
<tr>
<td>Hong, China</td>
<td>Virology Journal, September 2010</td>
<td>No</td>
</tr>
<tr>
<td>Henrich, USA</td>
<td>J Infect Dis, November 2010</td>
<td>No</td>
</tr>
<tr>
<td>Hohn, Germany</td>
<td>PLoS One, December 2010</td>
<td>No</td>
</tr>
<tr>
<td>Satterfield, USA</td>
<td>Retrovirology, February 2011</td>
<td>No</td>
</tr>
<tr>
<td>Furuta, Japan</td>
<td>Retrovirology, March 2011</td>
<td>No</td>
</tr>
<tr>
<td>Schutzer, USA</td>
<td>Ann Neurol, April 2011</td>
<td>No</td>
</tr>
<tr>
<td>Shin, USA</td>
<td>Journal of Virology, May 2011</td>
<td>No</td>
</tr>
<tr>
<td>Knox, USA</td>
<td>Science, May 2011</td>
<td>No</td>
</tr>
</tbody>
</table>
Detection of an Infectious Retrovirus, XMRV, in Blood Cells of Patients with Chronic Fatigue Syndrome

Vincent C. Lombardi, Francis W. Busceti, Javdie Das Gupta, Max A. Most

No Evidence of Murine-Like Gammaretroviruses in CFS Patients Previously Identified as XMRV-Infected

Konstanze Kneif, Donald Carrigan, Graham Simmons, Fernando Teigue, Yanren Zhou

Editors' Expression of Concern

In the issue of 23 October 2009, Science published the report "Detection of an infectious retrovirus, XMRV, in blood cells of patients with chronic fatigue syndrome," a study by Lombardi et al., purporting to show that a retrovirus called XMRV (xenotropic murine leukemia virus-related virus) was present in the blood of 67% of patients with chronic fatigue syndrome (CFS) compared with 3.7% of healthy controls (1). Since then, at least 10 studies conducted by other investigators and published elsewhere have reported a failure to detect XMRV in independent populations of CFS patients. In this issue, we are publishing two Reports that strongly support the growing view that the association between XMRV and CFS described by Lombardi et al. likely reflects contamination of laboratories and reagents with the virus. In one Report, "Recombinant origin of the xcm2 virus XMRV" (2), T. Paprotka et al. trace the ancestry of XMRV and provide evidence that the virus originated from two mouse leukemia viruses. In a second Report, "Evidence for murine-like gammaretroviruses in CFS patients previously identified as XMRV-infected" (3), A. E. R. MacGillivray et al. show that the sequence of XMRV is highly similar to that of MLV and that the XMRV sample used by Lombardi et al. was contaminated with MLV.
Judgment in the face of scientific uncertainty: 
the last word in rationality…
Diagnosing elicitation issues - UK reservoir risks

Last UK dam failure 1970 – no casualties

Experts, in former times!
Example of group inconsistency

Experts’ opinions on the time-to-failure (in days from first detection) of the 10%ile slowest cases, and two alternative ways of pooling weighted opinions – Equal weights and Performance-based DMs

Note the “two schools of thought” effect… and the strong ‘opinionation’ of many experts
Reservoir engineers: performance-based scores, and peer assessment rankings

Note: peer weighting is poor predictor of performance-based weights!
Matters arising ..... 

• Who wants to know their own calibration score?

• Voight effect .... proximity of expert’s medians to realizations for experts with identical statistical accuracy scores?

• Paper, pencil, eraser –v– spreadsheet?
• **How erratic are elicited scientists??**

Despite built-in check formula and Warning Flags in spreadsheet, multiple non-increasing quantiles (NIQ) appeared in CDC elicitation files:

Study: 7188 expert-target responses over 15 panels;
- no panel had zero NIQ errors;
- 73 NIQs in total on 444 targets;
- 20 of 48 experts had at least one NIQ;
- highest rate of NIQs was 4.9% on Panel1 (from 225 expert\-targets);
- average 1.4% NIQs over all panels;
- biggest recidivist had 19 NIQs on 143 targets

• Revising responses – one expert changed calibration item quantiles; their P-value went from $10^{-7}$ to 1.0!
NIQs not correlated with Classical Model calibration p-value

CDC elicitation: number of Non-Increasing Quantiles, by expert ranking

Calibration p-value rank: low [1] to high [48]
CDC study

- Another error encountered was the "impossible event" or the "certain event". For some items, some experts assigned 0% to all quantiles, and others 100% to at least the 50% and 95% quantiles.

- A frequency-based approach can help, by guiding experts to think in terms of "1 in 1,000,000" or "9999 in 10000", etc.

- The message: probabilistic thinking is not easy for everyone, and improbable is often mistaken for impossible. Cromwell’s advice, writing to the General Assembly of the Church of Scotland on 5 August 1650: “I beseech you, in the bowels of Christ, think it possible that you may be mistaken.”
A reflection on the Classical Model in the Montserrat volcano case:

“Those are not the exact numbers I would use, but I can’t argue with them“

Senior USGS volcanologist, Montserrat Volcano Observatory: 1995.
• Despite my remarks, not all things are negative

“...”

“A.A. Milne: “The House at Pooh Corner” 1928

Thank you!