Aviation and its Contributions to Healthcare

Professor Martin Elliott

25 November 2015

I do not want to scare you, but a few weeks ago an untrained pilot manually landed an Airbus 320 at Gatwick Airport.

That pilot was me. But, fortunately for the people of South East England, it was in a simulator at CTC Aviation near Southampton. It was my third attempt at such a landing, despite the outstanding teaching of Captain Guy Adams of CTC Aviation, to whom I am very grateful.

I made lots of mistakes. It is not surprising that a doctor like me makes errors. We know, from a seminal US Institute of Medicine report (To Err is Human, 2000[i]) that medical errors and health-care-associated accidents lead to 200,000 preventable deaths per year in the USA alone. That is the equivalent of 20 large jet airliners crashing every week; with no survivors. In the UK, estimates of the mortality associated with medical error vary from 250 to 1000 per month: that's about 1 to 4 jumbo jets. This remains shocking, especially remembering the words of Hippocrates; "First Do No Harm".

Imagine the public outcry there would be if so many plane crashes did happen and consider for a moment whether you would get on a plane or not. The death of 250 healthy people in one catastrophic and spectacular event always triggers headlines and comment. But deaths in hospital do not occur in groups like that. Usually it is an individual who is already ill that dies, and the tragedy primarily affects the victim's family and friends. If similar events take place in other wards or other hospitals, in other towns, it is only when someone spots a pattern that the scale of the problem emerges.

Despite the obvious risks, the airline industry, along with the nuclear and railroad industries are defined as high-reliability organisations. There are few fatalities per encounter with the system. Five characteristics of high-reliability organisations have been described by Weick and Sutcliffe (2001)[ii], these are:

1. pre-occupation with failure,
2. reluctance to simplify interpretations,
3. sensitivity to operations,
4. commitment to resilience and
5. deference to expertise.

As Professor Rhona Flin[iii] has pointed out, significant levels of protection and redundancy are built in to modern technology-based systems, but as hardware and software get ever more reliable, the human contribution to accidents becomes ever more apparent. The estimated contribution of human factors to accidents in hazardous technologies varies between 30% and 90%. Against that background, remember that most hospitals spend about 60-70% of their budget on people, and most of these are at the ‘sharp end’, dealing with patients.

Patient care, and especially cardiac surgical care, involves multiple interactions between people, technological hardware, software and the environment, as has been beautifully described by Ken Catchpole (http://csmc.academia.edu/KenCatchpole), and based on the SHELL model from aviation[iv]. Human error is inevitable[v], and these interfaces between people and systems are also prone to failure. The more interfaces there are, the more opportunity there is for failure. The rapid scaling up of the number of potential interactions between increasing numbers of people can be demonstrated by imagining each point on a polygon as a person, and then drawing lines between them. By the time you reach eight people, the 'map' looks like a confusing spiders web. And errors can be amplified like Chinese whispers as the number of interfaces increases.

Thus it may not be surprising that, when plotted on a similar scale, medicine cannot be defined as highly reliable.

The question that has fascinated those of us in medicine concerned with improving safety is exactly how has the airline industry become so reliable? Certainly, technological advances in material science, software and communications have played their part; planes are self-evidently much better than they were. But it became clear during the 1970’s, that many major fatal airline accidents did not have a primarily technical cause.

I want to consider three infamous air-crashes which led to major change in the airline industry and which have subsequently influenced ours. More detailed information is easily available on the Web, with detailed reports being searchable at the National Transportation Safety Board site http://www.ntsb.gov/Pages/default.aspx. Excellent summaries of these individual cases can also be found in the work of Gordon, Mendenhall and O’Connor[vi].


This Lockheed L-1011 bound for Miami crashed as a result of its experienced air crew becoming distracted by a
burned-out landing light, which had a financial value of only 50 cents. Whilst they tried to sort out this technical problem (which involved them checking that the landing gear was down), they put the plane on autopilot, set at 2000ft. Without them realizing it, they had actually put the plane into a very slow, shallow descent and the plane crashed into the Florida Everglades, killing 101 of the 176 people on board. A warning from air traffic control (ATC) was vague and non-specific "How's it going Eastern?" despite its monitors clearly showing dangerously low, and decreasing, altitude. The crew had lost awareness of the severity of the situation and there was inadequate challenge from junior officers.

**KLM Flight 4805 and PanAm Flight 4805. March 27th 1977**[viii]

This remains aviation's most deadly accident. The two flights involved were both Boeing 747 aircraft, diverted to the tiny airport of Los Rodeos on Tenerife in the Canary Islands because of a small terrorist bomb explosion at nearby Las Palmas. It was a Sunday, and the Los Rodeos ATC had only two people on duty, and no ground radar. The airport had limited facilities, including aircraft parking slots, and only one runway, which had to double as a taxiway because of overcrowding. On the day of the accident, the weather gradually worsened, and fog rolled in from the sea. Visibility became minimal. The KLM flight had been told to taxi to the far end of the runway, turn 180 degrees and prepare for takeoff. The Pan Am flight was preparing to taxi along the runway from the opposite end with a plan to leave the runway beyond the congestion of parked planes.

Communication between ATC and the aircraft was difficult because of the poor English of the ATC staff on duty, and the very non-standard language of the KLM Captain (Captain Jacob von Zanten, who was the poster-boy for KLM at the time. Indeed his photograph was beaming out of the in-flight magazine the doomed passengers were reading). Capt. Von Zanten was very confident in his own abilities, but used non-standard phraseology in his communications with others, both his team and ATC. He misinterpreted the ATC instruction of "stand by for Takeoff" as clearance, and said "Let's go". The co-pilot transmitted the rather meaningless "We are now at takeoff‖ further confounding matters between ATC and the Pan Am flight. And remember there was no ground radar, so the ATC were effectively 'blind'.

The KLM flight accelerated down the runway, itself blind to the presence of the approaching Pan Am jumbo. Von Zanten saw the Pan Am flight at the very last minute, pulled back on the joystick, but not soon enough to clear the plane and a horrific collision occurred. Things got worse because the ATC could not initially see what had happened through the fog and when finally the fire crews arrived, they spent 20 minutes at the KLM flight, on which everyone died, unaware that a hundred meters away, people in the Pan Am flight, potentially rescuable, were being incinerated.

In addition to the obvious communication issues between ATC and the aircraft, the investigation report believed it was possible that the KLM 1st Officer was intimidated by 'the Captain's legendary status' and was not assertive enough in preventing the captain from making such a huge error, despite the 1st Officer clearly understanding that they had not been issued takeoff clearance.

Five Hundred and Thirty Eight people died.

**United Airlines Flight 173, December 28th 1978**[ix]

Like the two previous examples, this accident was not associated with major technical failure. Rather it highlighted other important human factor issues, and proved a 'tipping point' in aviation safety[x]. Here is a section from the summary of the official NTSB Report: "About 1815 Pacific standard time on December 28, 1978, United Airlines, Inc., Flight 173 crashed into a wooded, populated area of suburban Portland, Oregon, during an approach to the Portland International Airport. The aircraft had delayed southeast of the airport at a low altitude for about 1 hour while the flight crew coped with a minor landing gear malfunction and prepared the passengers for the possibility of a landing gear failure upon landing. The plane crashed about 6 nautical miles southeast of the airport. The aircraft was destroyed; there was no fire. Of the 181 passengers and 8 crewmembers aboard, 8 passengers, the flight engineer, and a flight attendant were killed and 21 passengers and 2 crewmembers were injured seriously.

The National Transportation Safety Board determined that the probable cause of the accident was the failure of the captain to monitor properly the aircraft's fuel state and to properly respond to the low fuel state and the crewmember's advisories regarding fuel state. This resulted in fuel exhaustion to all engines. His inattention resulted from preoccupation with a landing gear malfunction and preparations for a possible landing emergency.

Contributing to the accident was the failure of the other two flight crewmembers either to fully comprehend the criticality of the fuel state or to successfully communicate their concern to the captain."

The captain "had a management style that precluded eliciting or accepting feedback". The first officer and flight engineer (who died) failed "to monitor the captain" and give effective feedback and provide sufficient redundancy. It was only when it was too late that the first officer expressed a direct view, "Get this **** on the ground". The crisis was neither prevented, nor managed or contained. The NTSB believed that the accident exemplified a recurring problem - a breakdown in cockpit management and teamwork during a situation involving malfunctions of aircraft systems in flight.
Culture

Prior to that time, aviation culture was centred around the pilot and his (mainly his) or her flying skills. As has been pointed out, early aviators flew alone with no radio contact. Those who took the risks took the consequences. All pilots learnt to fly solo early in their training, embedding some independence in their thinking. The culture in the 1970's was characterised by a steep hierarchical arrangement, with the captain at its apex. The Captain's word was the law, and he was not to be challenged. Capt. Chesley (Sully) Sullenberger (the captain who safely brought a plane to land on the Hudson river in New York, has said that "in the bad old days, when the captain was a god with a small 'g' and a Cowboy with a capital 'C', first officers carried little notebooks that listed the idiosyncrasies and personal preferences of different captains". There was not really a concept of a team at all, first officers and engineers being thought of like fire extinguishers "break glass if they're needed" (Robert Helmreich, quoted in[xii]). This has sometimes been called the 'trans-cockpit authority gradient', a term attributed to the late Elwyn Edwards[xii] in 1972.

What these accidents highlighted was the importance of the prevailing culture in which the aircrew operated and the overarching importance of Human Factors in influencing safety. Indeed, the NTSB Recommended after the United 173 investigation that the NTSB should "Issue an operations bulletin to all air carrier operations inspectors directing them to urge their assigned operators to ensure that their flightcrews are indoctrinated in principles of flightdeck resource management, with particular emphasis on the merits of participative management for captains and assertiveness training for other cockpit crewmembers. (Class II, Priority diction) (X-79-17)". Training was to be radically reformed to include aspects of culture and behaviour.

It is hard to pinpoint exactly when it started, but the recommendations from the United 173 investigation certainly helped the development of non-technical skills by Crew Resource Management. Several key people and organisations began to work simultaneously to develop a better understanding of the way aircrew worked together and to develop more effective training. These were; United Airlines, NASA and several academic psychologists, notably John K. Lauber who worked at the Ames Research Center for NASA and Robert Helmreich and Clay Foushee in Austin Texas. Pan Am and United Airlines had also been working on flight simulation and direct pilot observation in the 1970's. The academic and professional observations enable researchers to both define behaviours and to develop rational training programs, with the specific aim of improving teamwork and driving safety. An excellent detailed description of all that CRM entails is beyond the scope of this lecture, but is available in text book format[xiii], and there is a good 'do it yourself' guide at this website http://www.crewresourcemanagement.net/. CRM training improves the non-technical skills of aircrew, and comprises:-

1. communication,
2. leadership skills,
3. decision-making,
4. situational awareness,
5. team-working,
6. managing stress and fatigue and
7. understanding ones limitations.

When such training was first introduced by United, it was far from popular with the pilots who called it 'charm school'. But that training (beautifully described in detail by Rhona Flin and colleagues[xiv]) has evolved considerably over time, and is now used in a wide variety of high-reliability organisations.

The impact of CRM training on the safety of commercial airliners has been immense. There has been a marked decline in the number and severity of accidents (as judged by the number of hull losses and deaths), despite a huge increase in the number of flight departures. The accident rate per million departures has fallen exponentially since the 1960's and is now around 2.7 per million departures[xv], a 1:400,000 chance. Even so, being a pilot or flight engineer is still the 3rd most dangerous occupation in the USA after logging and fishing, and much more so than being a soldier, as the US Bureau of Labor statistics from 2014 testify.

Surgery, and especially cardiac surgery, was populated in the 70s and 80s by surgeons, usually male, who demonstrated many of the behaviour patterns seen in commercial pilots of the time. Sir Lancelot Spratt typified these behaviours, as seen in the Doctor in the House movies. Dominant, aggressive, confident, secure in their beliefs and teaching by humiliation. My early training was very much like that. As a junior, it was incredibly difficult to challenge the judgement, decisions or authority of a consultant. You just had to do it their way (whatever 'it' was), and all their ways were different. There is an old joke asking how many cardiac surgeons it takes to change a light bulb. The answer is one; he just holds onto the light bulb and the whole world revolves around him. Most of those setting out to train as cardiac surgeons are high achievers with high self-confidence. They have been described as 'goal-orientated, with a strong sense of their ability to control their actions and environment'[xvi]. Historically, they also sacrificed their personal needs on the altar of their career. And the role, as Winlaw describes, does bring with it a degree of positional power which some may find attractive.

A cardiac surgeon, like a pilot, must be situationally aware, be able to marshal available resources, be able to initiate rapid changes in management and do so in a way that sensitively uses adaptive and command and
control skills. It is perhaps not surprising that there are some surgeons in whom the boundaries between such appropriate behaviour and narcissism become blurred. These surgeons will be arrogant, have an inflated sense of their own importance, expect special treatment from others, but are quick to blame rather than taking personal responsibility. Humility in these people is rare. In its worst form, this behaviour is associated with throwing instruments, shouting and belittling colleagues. I have seen all of these in my career, and in every continent.

Sometimes, the boundaries between certain individual characteristics which might be regarded as good and those which have more negative consequences are rather blurred, but I think this comparison table (first brought to my attention by Tony Giddings) between a strong leader and a 'derailing' leader emphasises some to the features which were once accepted as normal but are now sensibly being questioned.

Such inappropriate behaviour, whilst not always aggressive, can result in real harm to patients, under-reporting of incidents, self-justifying explanations or even cover-ups, and an environment where others feel unable to speak up. Exactly like in the air crashes described earlier.

Things were not much better in the early 2000s, as pointed out in a survey of operating room staff by Rosenstein and O'Daniel[xvii]. They discovered that abusive behaviour was common, especially amongst senior staff, and this bad practice was passed on to junior staff. Most observers felt that adverse events were more likely to happen in this environment.

Technical performance was seen to be the most important skill, and 'a good pair of hands' was the most important requirement for success. The idea that other skills might be important has been slow to develop, much to the amusement of pilots who had already adopted the methods. For example, Pilot John Nance suggested the following interaction between surgeon and patient when he opened the US National Patient Safety Foundation, "Sorry, Ms. Wilson, that we cut off the wrong hand; but how was the food?" The wrong focus and the wrong interpretation.

Just as the disasters in the 1970's focused the collective mind of the airline industry, so in the 1990s and early 21st century, so there came about the realization that medicine was not 'safe' and that human error, whilst predictable, was an un-mitigated risk in healthcare, and that something should be done about it. As ever, ideas have their time, but the vision and energy of individuals are required to develop them.

The first of these individuals is Professor Don Berwick. Berwick is and always has been a paediatrician in Boston. He also had a Masters in Public Policy, and in 1983 he became Vice president of Quality of Care Measurement to the Harvard Community Health Plan. In 1989, he co-founded the Institute for Health Improvement which has been seminal in influencing a generation of healthcare workers around the world, and whose methods of Quality Improvement (QI) have been widely copied. The IHI has grown in stature and influence under his tutelage and the NHS (of which he is in favour) has benefited directly from both his opinions and the work of those he trained. At the start, though, he was responsible for bringing the 'safety' of medicine into public and political focus. He received an honorary knighthood in 2005, a reward for the work he did for safety and QI in the NHS.

Equally important was Professor Lucian Leape, also of Harvard. A surgeon, Leape, became a founder member of the Medical Practice Study (MPS) of medical injury, and conducted some of the first studies into the over and underuse of medical procedures. The observation of the extent of potentially preventable harm led him to study the causes of errors, and in 1994 he published a landmark paper[xviii], Error in Medicine which called for the application of systems theory, similar to that used in the airline industry, to prevent medical errors. This led ultimately to the establishment of the National Patient Safety Foundation and the Institute of Medicine's most influential publication; To Err is Human[xix]. More recently, through the Lucian Leape Institute, he has been highlighting problems of the culture in medicine, notably that of disrespect, and it is striking how much of the behaviours he highlights in his publications reflect those behaviours seen in pilots in the 1970s and as pointed out by Rosenstein's survey.

What is important about these Berwick and Leape is that they did not just spot the problem, but had the training, experience and desire to do something systematic about it. The organisations they both established and work in have been profoundly influential, and their methods are in widespread use.

The open, transparent, self-critical and non-punitive approach they fostered stimulated workers in my specialty to consider the human factors in our work. Pre-eminent amongst these was my predecessor, Marc de Leval. Marc was a highly talented, creative, driven surgeon who expected the highest standards from himself and those around him. At the height of his career, when he already had a spectacular international reputation, he published an article analysing a cluster of surgical failures (a high mortality rate) in the neonatal arterial switch, for which previously he had excellent results.

He asked the following questions: - (1) Could the cluster of failures be due to chance alone? (2) Could procedural risk factors and their variation across time explain the mortality? (3) Could human error account for the cluster of failures? (4) Could appropriate monitoring techniques allow early detection of trends in surgical outcomes? (5) Could outcome measures other than death provide a refined way of monitoring surgical performance? (6) If the surgeon's performance varies across time how is it best expressed? (7) How can the failure rate be reset to its
At the time, this was considered extraordinarily brave. Surgeons published good results, and neither they nor the journals had much interest in negative results. Marc introduced several ideas to our specialty that had previously not been widely considered.

These were, presentation of results in the form of a CUSUM chart and the use of alert lines, the idea that human factors might be important and the concept of a ‘near-miss’. In his work, Marc referred specifically to lessons from aviation, pointing out that Near Misses were (and are) routinely reported to the Civil Aviation Authority and analysed to see if anything could be improved to make flying safer. The degree of risk inherent in each incident is assessed, trends are analysed, and recommendations for remedial action are made.

Marc and his colleagues defined as a near miss in cardiac surgery as ‘the need to go back onto cardiopulmonary bypass to correct something after the operation was completed’, thus creating a more sensitive ‘failure’ indicator than death alone. Not only did Marc introduce the concept of ‘near-miss’ into cardiac surgery, but with this paper he made people think about the key human factors, including the potential for failing performance with age. The fact that, after Marc ‘retrained’, mortality fell almost to zero suggests that was not an issue for him. Such self-referral for retraining with another surgeon, based on data, was very unusual. His openness and honesty made him, quite appropriately in my view, something of a legend in the field, and this paper is widely cited, and its methods routinely applied. Here, for example is a CUSUM chart showing the continuously improving operative mortality for surgery for congenital heart disease at GOSH.

Marc and his colleagues defined as a near miss in cardiac surgery as ‘the need to go back onto cardiopulmonary bypass to correct something after the operation was completed', thus creating a more sensitive 'failure' indicator than death alone. Not only did Marc introduce the concept of 'near-miss' into cardiac surgery, but with this paper he made people think about the key human factors, including the potential for failing performance with age. The fact that, after Marc 'retrained', mortality fell almost to zero suggests that was not an issue for him. Such self-referral for retraining with another surgeon, based on data, was very unusual. His openness and honesty made him, quite appropriately in my view, something of a legend in the field, and this paper is widely cited, and its methods routinely applied. Here, for example is a CUSUM chart showing the continuously improving operative mortality for surgery for congenital heart disease at GOSH.

Marc de Leval wrote an essay in the Lancet in 1997, in which he compared errors in surgery with conventional theories of error. He wrote:-

"Human beings can contribute to the breakdown of a system through errors or violations, which can be divided into skill-based, rule-based, and knowledge-based.

Skill-based errors are associated with some forms of attentional capture, either distraction or preoccupation, and are unavoidable. Even complex tasks such as open-heart surgery can become largely automatic. Failure to administer heparin before initiating cardiopulmonary bypass, for example, is a skill-based error. Skill-based violations typically involve corner-cutting: they are customary among surgeons for whom speed is the main criterion of surgical excellence.

At rule-based level, human performance is governed by memory-stored rules, solutions acquired as a result of experience, training, or both. Such mistakes arise either from the misapplication of good rules or the application of bad rules. Rule-based violations are deliberate acts of deviation from safety procedures. To turn off an alarm system to use equipment beyond its range of safety is a common example.

The knowledge-based level of performance is relevant in unfamiliar situations for which action must be improvised. Knowledge-based errors are the attribute of the learner. Knowledge-based violations have their origins in particular situations in which non-compliance seems to be justified to accomplish the task. To allow the blood pressure to fall temporarily to control massive bleeding is an example."

De Leval argued that medicine should preoccupy itself with error (as Weike and Sutcliffe suggested [xxiv]), and
search for the latent as well as the active failures in systems, and use human factors science in both analysis and training. This essay proved very perceptive, and he went on to study the impact of human factors on the outcome of the same type of operation done in many centres in the UK[xxv]. The study was difficult because of the physical method of observation; a researcher was needed to be present at each operation and to record events on paper. Audio and video recording would have made the study much more effective. Nonetheless, once again this work had a big impact in helping surgical teams begin to understand the importance not just of what they did technically, but how they worked together; their non-technical skills.

Nothing, though, has the impact of a personal story. And I want to show you two.

This is Martin Bromiley. Martin Bromiley is a pilot whose wife Elaine died during anaesthesia for a simple elective sinus procedure. He tells the story much more effectively in an online video (from www.risky-business.com) than I can, but the anaesthetists looking after her lost all situational awareness and persisted in trying to intubate Elaine, even though they should have been performing a tracheotomy, as the nurses knew but were not assertive enough to make it happen. Martin has subsequently formed the Clinical Human Factors Group to foster better safety management within the healthcare using CRM techniques. He often makes the point that safety is not just a product of data analysis; if you achieve safe outcomes, it doesn't mean that you will do it every time. It is the process that is important in achieving high reliability. Key issues in healthcare were listed in a recent letter to the BMJ[xxvi]. These are:

1. Analysis of accidents should include an examination of "human factors issues," especially workplace behaviours
2. The findings from these analyses must be linked to ongoing training of the behaviours that constitute non-technical skills in healthcare
3. Humans will always be prone to fail in systems that have not been designed using ergonomics/human factors principles.

Here is another important personal story, this time from Clare Bowen. Again the video is self-explanatory, and very harrowing. First, though, let me explain what a morcellator is. It is a powered device, shaped rather like a gun, which has rotating blades within a tube and which can shred large pieces of tissue when used laparoscopically, theoretically reducing the need for large incisions.

Clare makes a passionate plea for the use of human factors in medical school training, and points out that in aviation the pilot's life is also at risk. In surgery, that is definitely not the case.

In paediatric cardiac surgery the expectations of families, administrators and other clinicians are uniformly high.[xxvii] The specialty has become a focus of attention because it is has been both one of the earliest disciplines to engage multiple different old-style specialties (e.g., surgery, anaesthesia, intensive care, radiology) into a single team. It has also been the centre of several internationally known cases of team failure, notably in Bristol, Oxford, Winnipeg and most recently St. Mary's Hospital in West Palm Beach, Florida. As Bromiley has suggested, whilst initial analysis of mortality data may highlight the problem, and whilst the common kneejerk response is to name and shame the surgeon, subsequent investigations usually highlight systems issues, as they have done in the airline industry.

I want also mention Professor Atul Gawande, the 2014 Reith Lecturer for the BBC. Gawande is a surgeon and also a professor of Public Health. Politically active on the democratic side of American politics, he is also an accomplished and prolific writer, famously for The New Yorker. Later to the table than the first two names I mentioned, he nonetheless has made significant contributions to the debate about quality, safety and waste. He wrote a seminal article in 2009[xxviii] looking at the staggeringly expensive healthcare in the Texas town of McAllen, and an equally important and best-selling book, directly relevant to the subject of today's lecture, The Checklist Manifesto[xxix]. Combining academic skills, a wide international network and personal charisma, he both challenged the lack of use of checklists in medicine, compared with other high reliability organisations, and carried out studies to emphasise their potential benefits.

This is the core checklist you need to get a 747 off the ground. You are basically not going to fly if any of these things are not ticked off. It is relatively simple, and thus cannot cover all the issues on a plane, but does deal with core elements, and must be crosschecked by the co-pilot. There was no culture of checklists in most surgical departments, although many individual surgeons and anaesthetists had mental lists. We relied on memory, the patient's notes and the experience of those around us.

Such reliance on memory was not a good idea. Significant mistakes continued to be made. Wrong site surgery, for example; taking out the left rather than the right kidney; performing the wrong operation, but on the correct organ; and even performing the correct operation, but on the wrong patient. Guwande and like-minded colleagues realised that if a checklist including core details (e.g. the identity of the patient, the site of the surgery, the diagnosis and planned operation) was always used, such risks would obviously be mitigated. They published a paper in 2009[xxx] that led to WHO standards being introduced for the use of checklists in surgery. They demonstrated that mortality and morbidity could be significantly reduced if checklists were used effectively and routinely. More specialty specific checklists emerged, and here is one on for cardiac surgery designed by the US Society of Thoracic Surgeons, currently in regular use in the USA.
Evidence was clearly accumulating that medicine remained dangerous, that mistakes Imperial College London were being made by healthcare workers and that in such a labour intensive field with a wide variety of individual ways of doing things, human error was inevitable. The application of CRM methods to healthcare and the use of aviation-based checklists would surely be appropriate. As Kevin Fong (an anaesthetist at UCLH and Horizon presenter) put it, "Standardise until you absolutely have to improvise". Individual, safety conscious and quality-driven hospitals and departments throughout the world started to adopt these techniques. And several institutions with good leadership established quality improvement programs based on the IHI principles. Academic departments grew up, notably in the UK those of Rhona Flin in Aberdeen and Charles Vincent at, and Peter McCulloch in Oxford. My own hospital established a Zero Harm program in 2007 under the leadership of a previous CEO, Dr Jane Collins, now CEO of Marie Curie.

The combination of the metrics delivered by Quality Improvement programs and the philosophies inherent in CRM methods produced powerful stimuli for change, and safety standards improved. Those units that have successfully implemented these techniques have seen similar improvements in quality and reduction in error.

But, this is very different from the massive improvement in safety in the airline industry that followed the introduction of CRM. The introduction of CRM in aviation was industry wide. When a pilot joins an airline he or she has to complete a three-day structured course on CRM, even if they have completed similar courses with previous employers. After that there is an annual refresher of these skills as well as getting critiqued during four days of annual simulator check/refreshers. Flight crews are taught and have to demonstrate that they are skilled in the areas of leadership, crew cooperation, situation awareness and decision-making. Assertion and communication skills are taught and assessed. You have to pass these courses.

In healthcare, the implementation of CRM has been haphazard and inconsistently led. Such training only happens via an interested Society or via a Royal College course or perhaps by a forward thinking Hospital Trust. But even then it seems attendance is voluntary, with those most in need of such training finding excuses for non-attendance. Very few medical schools include it in any part of the curriculum. It is staggering that something which has been shown so clearly to be of benefit, and accurately relates to every aspect of the teamwork modern healthcare workers need to espouse, is neither compulsory nor routinely assessed.

In truth, the NHS is a very complex organisation, with its component parts being able to function with considerable autonomy. Implementing strategies across the whole organisation becomes incredibly hard as the system leaves so much to local 'preference' and investment choices. These words of Captain Guy Hirst, former senior training pilot for BA and later with Attrainability, a company devoted to spreading the benefits of CRM training to healthcare, accurately reflect the state of play in the NHS today:-

"When I was still involved with Attrainability we had some great successes when proper initiatives were put in place and the outcomes were most impressive. That however was usually the exception rather than the rule. A good example is the introduction of the Surgical Safety Checklist. I spoke at the NPSA’s launch conference and was incredulous at the way it was introduced or should I say not introduced. Like so many of the initiatives in healthcare it was poorly thought out - it was ridiculous to expect clinical teams to understand the rationale and the nuances of checklists without proper explanation and training. Indeed it was and is used as a crude auditing tool that does no more than audit that someone has ticked a box!!! Harsh but true."

I can confirm this observation. I have had the privilege of operating in many hospitals in many countries, and in each one there is a different application of the WHO checklist, and the quality of use, and level of understanding of the significance is equally variable. I have seen brilliant junior medical and nursing staff, who clearly understand the rationale for the checklist, struggling hard to get any serious engagement out of senior surgeons, who still seem to see themselves in the same 'above it all' 'this is a waste of my time' position as the captains and Lancelot Spratts of old.

Guy went on to add that some teams, including many at my own institution, are doing it well and the checklist clearly helps, but he continues to wonder whether this is due to luck rather than judgement. Guy's view is that healthcare, especially high-end surgery is much more complex than aviation, and that our patients are 'much more idiosyncratic' than a 747.

"You never totally know what you will find until you start operating, even in spite of the amazing imaging available these days. However that to me is another reason why it is so important to try and standardise. Human beings are exactly that - HUMAN - and thus error is ubiquitous. For that reason it is essential to have a robust error management strategy. Yes we try and avoid making errors but that is not always possible hence we need to either trap our own errors or more likely hope our team members trap them and if that layer doesn't work then we need to mitigate the potential consequences."

Airlines have some tools at their disposal that are rarely available in healthcare. These are black box data recording, including audio and video records of cockpit activity, and simulators, which allow aircrews' behaviour and responses to various scenarios to be assessed and used for regular performance review. To protect the public, aircrew can be removed from active service if they fail appropriate safety standards. Assessments are frequent; at least annually and for the very safety conscious airlines perhaps 6 monthly. They are observed and assessed by senior training pilots on real flights too, part of a process called Line Operations Safety Audit.
In my world, I have never had any kind of formal assessment of my technical performance, other than direct observation by a close colleague, or as part of a specific scientific study. We get very little feedback on our technical skills, often because that feedback usually has to be given by the same close colleague or friend who will have had no training in such a ‘difficult’ conversation. Whilst I have had some CRM training, it has never been formal, has never been repeated, and remains voluntary. I have never been assessed with my team in a simulator (although there is a good program of such work at UCLH). And as I said in my last lecture, there is little culture of rehearsal in surgery, rather one of analysing after the event.

Appraisal of consultants in the NHS, although improving slowly, remains largely supportive of the development of the individual and is far from the performance review one might expect in the harsh environment of a private sector company or even a university. It is often conducted by a friend and against variable criteria. A program of revalidation of doctors has been introduced by the GMC, hopefully to weed out those that are poorly performing, but it does not include CRM training nor make any assessment of technical skills, other than reviewing gross outcomes, and in my view sets a very low bar. Only a handful of doctors (0.7%) have failed to be revalidated. If we want excellence and to deliver consistent safety to the patient, we need to become tougher. Since we know inappropriate behaviour is dangerous, it should not be tolerated, as was the case with the two Fly Be pilots who were sacked in 2012, after an abusive mid-air row in which the pilot called the co-pilot ‘his bitch’ and was told ‘f*** off’ in return. Fortunately, no harm came to the passengers.

Sadly, no system of training is ever going to be perfect, as these two slides given to me by Manfred Mueller of Lufthansa demonstrate. The airline is one of the safest in the world, with the highest standards of training assessment. They have adopted very scientific approaches to their training methods and have demonstrated a clear relationship between excellent performance in psychometric tests at recruitment and subsequent complex flying skill tests in the simulator; these with poor PM results not getting through to become full pilots.

Five days later a GermanWings (Lufthansa subsidiary) co-pilot, Andreas Lubitz, (in the green zone on the testing charts) locked his pilot out of the cockpit and locked his Airbus 320 into a fixed fatal descent into the Alps killing all 150 people on board. We know all this from the black box flight recorder and ATC records.

Over the last decade, several groups, including our own, have studied the potential of black box technology in the operating room. Initially this proved technically challenging, because of the large teams involved, the varying sound and light levels, the range of technology to be monitored and the lack of a uniform time code between pieces of equipment. The time shown on one screen may be seconds different from that shown on another making tracking of an event to its root cause a considerable challenge. Despite that, enough early evidence exists to say that similar threats to those occurring in the cockpit can result in unwanted states or even errors. Examples include bad behaviour, interruptions or distractions.

Recently, video, audio and data acquisition have all improved, and new potentially viable black box solutions are becoming available, and early reports are very encouraging, notably Toronto carried out by Teodor Grantcharov in laparoscopic surgery (in which it is relatively easy to adapt the technology). I know from my own experience that people's behaviour changes greatly when they see how they behave and appear to others on playback of video. It is very revealing. I think that such equipment should be everywhere and form part of the electronic patient record, and become the basis for simulator datasets for scenario testing. Unfortunately, given the parlous state of the NHS finances, unless there is a massive change in priorities, it is unlikely to happen unless funded via research projects. However, in my view this should be core in-service training and not subject to the fickle nature of the research grant world.

I want to bring my talk to a close by telling you about some research from Toronto Sick Kids Hospital which to my mind brings together many of the themes which link aviation and healthcare. Many people have been skeptical of that link, applying what anaesthetist Kevin Fong has described in a Horizon program for the BBC as "The Law of False Equivalence"; just because it worked in aviation doesn't mean it will work elsewhere.

Toronto Sick Kids has always been a top children's cardiac centre, with great tradition of both high quality and innovation. Over the last decade, they have made a series of appointments of people dedicated to data collection and created a culture of analytical self-criticism. They were collegial, introspective and supportive of each other. In the late 'noughties' they had introduced weekly 'performance rounds' in which each case which had had surgery was presented and discussed in front of a congregation of the whole multi-disciplinary Heart Centre. In 2010, Ed Hickey, an excellent British surgeon (whom we have sadly lost to Canada) was presenting cases at this meeting and realised that the journey of sick children through their centre was really rather like the 'flight path' of an aircraft, and developed a graphical way of simplifying the way in which cases were reviewed. He built an Access database to feed this graphic, but soon found it 'unanalyzable'. But it did trigger a detailed review by him and others of airline models of preparing and reporting flight plans and paths, and so recoded his data using similar terminology, based on the threat and error management model. I am very grateful to Ed for allowing me to use some of his slides and some of his data.

A pilot, preparing for a flight, has to submit a flight plan to ATC and to his organisation, and this will form the basis of briefings with the crew on the plane itself. You need to know where you are taking off from, where you are going to and what special things you need to plan for. Pilots are well aware that threats exist which may
impact on their journey, for example weather, terrain, traffic, etc, and that if you fail to mitigate these risks, an **unintended state** can be entered, in which it is relatively easy to have some slip in solving the problem, creating **errors** can occur which threaten survival. Errors can, of course, be corrected, but under the pressure that can occur at such times, sequential unintended states and errors can develop, creating **error cycles** which are much more dangerous and difficult to resolve. Thus the planning of a flight requires that the threats are identified and both they and potential mitigations discussed and communicated. There may be a thunderstorm nearby which it might be better to avoid, of terrible weather at the destination airport that might prompt a diversion.

The **threats** that threaten safe flying have equivalents in cardiac surgery. These range from the disease itself, co-morbid conditions such as diabetes or lung problems, equipment issues, various stressors and distractions and the underlying culture of the organisation.

The **errors** can be classified too. A **violation error** is a conscious deviation from a standard procedure or care pathway. A **procedural error** is a mistake such as giving the wrong dose or leaving a swab behind. A **communication error** is self-evidently failing to get the right information across to someone else. A **proficiency error** implies sub-optimal execution of a task, and has been the dominant form of mutual criticism of surgeons for years! A **judgment error** is when the wrong course of action is chosen; a poor decision.

If we plot the relative risk to a patient against their time course in hospital, we see considerable variation over that time. There is a little risk when they are admitted, depending on how sick they are, but risk increases dramatically over the course of the operation, and gradually declines over the post op care period until the patient is ready to go home. It was a graph like this that made the Toronto team realise the similarities with flight planning, and that if they got together before surgery to discuss the patient they could draw up a formal personalized flight plan which would consider all the identified threats and discuss and agree appropriate mitigations, hopefully to prevent the development of unintended states and errors. They then observed the patients during their course and recorded what happened to them in a revised database, before (as a whole team) reviewing each case at the end of their care.

So they had to accept the idea, develop the software, engage the team, and embed it into the working practices of the unit and then begin to collect baseline data. This is not an easy task, and one must emphasise the importance of the leadership of their unit, namely Professors Glen van Arsdell and Andrew Redington, another brilliant Englishman we have lost. They were truly supportive. They now have hired one person to maintain the database and prepare presentations for their meeting.

The early data from their experience seems to support the threat and error model as being relevant. To demonstrate that, I will quickly take you through the story of a patient undergoing theoretically simple surgery. This child aged 46 days had a simple hole between the two pumping chambers of the heart, known as a VSD. He first had his carotid artery punctured instead of the adjacent vein during preparation for anaesthesia. During the course of his surgery, it was found that the VSD was not completely closed so he had to have a 2\textsuperscript{nd} episode of bypass to fix it. His heart showed a higher than usual pressure at the end of the operation and later he could not have his breathing tube removed at the predicted time and later developed a wound infection. A complicated, but ultimately successful, outcome. If we superimpose on his charts what were threats, errors and unintended states, you get some idea of how this works.

Out of the first 519 patients they studied with this method, errors occurred 260 patients (50%), and in 173 patients (33%) these had clinical consequences. One hundred and nine patients (21%) actually entered cycles of unintended states and error, which, just as in aviation, are associated with a progressive loss of safety margins. There is an associated increased risk of very adverse outcomes from residual problems in the heart, through brain injury and death.

The group has extended and automated a lot of this work, from simply automatically creating PowerPoint slides to more detailed observation based studies with video and audio recording, similar to LOSA (Line Operations Safety Audit) assessments of aircrew. There have been technical difficulties, but overall it has proved easy to implement after initial skepticism by some. They feel it has improved objectivity, reduced the tendency to blame and allowed them to concentrate on things other than just mortality, which like at GOSH, is now too low to be used as a performance metric.

It has improved accountability; nothing can be swept under the carpet, and no case is excluded for discussion and analysis. And by following patients through to the time of discharge they have been able to identify errors that amplify as the hospital stay extends. They have begun to remove some bits of data from collection which proved not to be useful and have successfully implemented the process in another hospital in Toronto. It has required continued leadership, as all safety initiatives do, but a key learning point emerged quickly, “**The minute you are in a spot you did not expect to be in, you are now in a much more dangerous spot than you appreciate**.”

There are many physicians out there who think that comparing medicine to aviation, or indeed any other industry, is to denigrate the human side of medicine and cannot reflect the wide variation in individual patients and their needs. The methods used in aviation may indeed not be applicable throughout medicine (I am not sure what difference they would make in a dermatology clinic), but they are self-evidently relevant to the complex
technology and team heavy disciplines such as cardiac surgery, neurosurgery, emergency medicine and intensive care. The American Heart Association recently published\[xxx\] a detailed review of the scientific background to the study of patient safety in the cardiac operating room and, whilst asking for more detailed and prospective research to be done, nonetheless felt that there was sufficient evidence available that many of the aviation based approaches I have described should be implemented across the board.

We have yet to see the massive improvement in safety in medicine that has been seen in aviation since 1980. There seems little doubt that if those benefits are to be fully appreciated in medicine; we need to continue to make significant changes.

**Firstly**, and most importantly, safety has to be seen as a top priority for the organisation, from Secretary of State down, and not just something to which lip service is paid. A *climate* of safety can only exist if it is both believed in and proselytized by the leadership and maintained throughout the organisation. This is really difficult if your CEO changes every two to three years, as is the average in the NHS.

Safety and quality improvement should not be optional or dependent on charity or research funding as they so frequently are in the NHS. Safety should not be compromised for financial reasons. Whilst investment may be needed to incorporate some of the changes into our system, we must remember that both complications and litigation are expensive, and both can be reduced by improving safety. The current squeeze being put on NHS finances, combined with increased demand threatens safety anyway, and we have to be hyper-vigilant to make sure that safety programs are not marginalized or even closed down.

When short cuts are taken in safety, risk increases and it is a brave manager who puts profit ahead of safety. Such actions were taken by Jeff Smisek, formerly CEO of Continental, when in an attempt to reduce the fuel costs of his business he reduced the amount of fuel carried to cover for emergencies. He was quoted (Manfred Mueller) as saying "Safety IS our top priority; flights can stop for extra fuel if necessary". During his tenure, the number of continental fuel emergencies at Newark Airport, NJ (less than 30 minutes of fuel left at landing) rose from 19 in 2005 to 96 in 2007. This was extremely unpopular with the pilots who had to be on board the planes. What if there had been bad weather, or a terrorist incident? As they put it, they'd be breathing fumes.

**Secondly**, the lessons of CRM must be incorporated, from the outset, into the training of all staff in the NHS, after all we work in teams most of our lives. Understanding the importance of successful interactions, behaviour and communication is critical to patient safety, and to effective teamwork. Making it core-business will reduce the risk of it being perceived as 'charm-school'.

**Thirdly**, and especially for those working in complex specialties, regular formal assessment both by observation and regular CRM training should either be added to or replace the current rather soft appraisal. Technical skills assessment may also be done this way. This needs to compulsory and not managed in the same way as current NHS mandatory training in such things as moving and handling, blood transfusion and so on, which is either didactic in groups or by e-learning as individuals. Only resuscitation training really approaches the personal assessment that CRM would require. Simulation would be ideal, but it is likely still to be restricted to very few sites. However, there is no reason that we can't have simulation and training centres to which surgeons and their teams must go for assessment, rather as aircrew training is outsourced now to companies like CTC.

**Finally**, we must continue to research this area and develop methods and metrics which allows us to improve the way in which we deliver the new treatments we discover elsewhere.

The organisations in which we work have a duty to support that, and the NHS must put safety first and send out the right signals to support that position. I am not sure what message is sent to frontline clinical staff by the appointments of accountants to lead our two primary regulators, Monitor and the CQC.

It is our duty, and theirs, First to Do No Harm.

With Thanks to

Captain Guy Adams, CTC Aviation
Captain Manfred Mueller, Lufthansa
Captain Guy Hirst
Captain Martin Bromiley
Professor Emeritus James Reason
Professor Marc de Leval
Professor Peter Laussen
Dr. Edward Hickey
personal communication


[xii] Professor of Applied Psychology, Aston University 1976-84; Director of Human Technology 1984-93


