Editorial

Putting the 'point' back into the ritual: a binary approach to difficult airway prediction

All happy families are alike; each unhappy family is unhappy in its own way...
–Leo Tolstoy: (opening line of 'Anna Karenina')

In this issue of the journal, Nørskov et al. add to the weight of evidence appearing superficially to suggest that we 'cannot predict the difficult airway' [1]. This is part of their series based on very large datasets: the evidence to support this mantra seems robust [2–4]. Although they conclude with the hope that airway prediction can be improved [1], the root cause of our apparent collective inability to predict airway difficulty was most elegantly explained by Yentis nearly 14 years ago [5]. He applied the phrase 'pointless ritual' to the act of airway prediction. Yentis' article should be compulsory reading for all anaesthetists: it is a logically balanced, beautifully phrased, perfectly crafted piece. However, it is also wrong. Its argument is based on a false premise, as are so many airway research studies searching for a 'predictive test'. The purpose of this editorial is to offer a correction and in doing so, we touch upon both diagnosis and principles of managing the difficult airway.

There is little to be gained by summarising Yentis' editorial – it is easier and more informative simply to read it. He shows how difficult airway prediction can be considered a diagnostic test, the efficacy of which is conventionally measured by sensitivity, specificity, positive and negative predictive values (PPV and NPV respectively), etc. The clinical signs of difficult airway patients overlap with those of normal airway patients and, because real difficult airways are so rare, the mathematical reality is that no predictive test will ever fulfil conventionally acceptable thresholds for sensitivity, specificity, PPV and NPV (Fig. 1).

The data of Nørskov et al. at first seem to support this nihilism [1–4]. With a sensitivity of approximately 9% and a PPV of approximately 30% (Table 1), a difficult airway seems to be rarely detected and if detected, incorrect as a diagnosis more often than it is correct. Few budding inventors would rush to patent a hypothetical diagnostic test with this level of performance. Nørskov's group alone has in total presented results of > 500,000

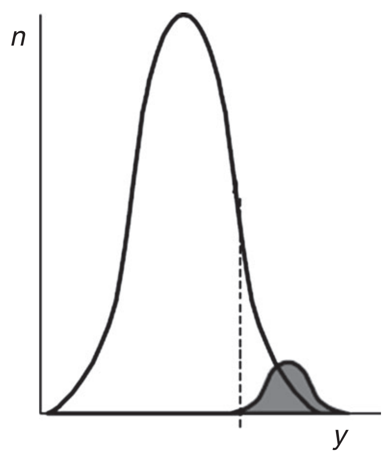

Figure 1 Adapted from Yentis [5]. A schematic representation of why he argued it is impossible to distinguish the difficult airway patients (dark area) from easy airway patients (clear area). The y-axis is number, and the x-axis an arbitrary measure for the clinical signs of difficulty. A hypothetical cut-off (dotted line) that reliably identifies all 'difficult' patients nevertheless will include a larger number of 'easy' airway patients. This yields poor test sensitivity, specificity, PPV, etc.

This editorial accompanies an article by Nørskov et al., Anaesthesia 2017; 72: doi: 10.1111/anae.13701.

Table 1 Adapted with simplified data from Table 1 of Nørskov et al. [2]. The statistical values for the performance of the predictive test are listed below.

	Difficult intubation		
	Yes	No	Total
Predicted difficult			
Yes	300	700	1000
No	3000	179,000	182,000
Total	3300	179,700	183,000

Sensitivity = 300/3300 = 9%; Specificity = 179,000/179,700 = 99.6%.
PPV = 300/1,000 = 30%; NPV = 179,000/182,000 = 98%.
Likelihood odds ratio = (300/700)/(3000/179,000)) = 25.6 (2,560%).

patients with broadly the same results. Case closed?

Correcting the false premise

Not quite. The first and perhaps most powerful point is that any diagnostic test can only apply to a homogenous condition [6]. Any universal test for 'cancer' will perform poorly for the simple reason that cancer is heterogeneous. There is no common test for say, breast cancer and bladder cancer and melanoma: we do not combine signs like haematuria or nipple retraction or skin discolouration to yield an aggregate score. Rather, diagnostic tests in this field are correctly specific only to the individual cancer type. Similarly, there is no airway-related commonality between a thin patient with normal jaw structure and isolated tempero-mandibular disease (little or no mouth opening), and an obese patient with good mouth opening but a severely receding chin, lingual tonsils and a beard. So a prior expectation of finding a universal test is inappropriate.

Because a difficult airway is not a single entity, giving weight to different airway signs by scoring simply cements the problem and does not solve it. Predictive power remains poor [7–10]. In reality, it is the patients with the easy airway that share the common signs. Or, to paraphrase Tolstoy: '*all easy airways are alike; difficult airways are difficult each in their own way...*'. Anaesthetists instinctively already appreciate this in their practice as they view their patients as individuals requiring ad personam management, rather than as samples randomly drawn from the population. Nørskov et al.'s control group anaesthetists performed as well as the intervention group because they used their clinical instincts to predict specific airway difficulties, rather than apply to apply a fixed algorithm to all patients, regardless [1].

A second point is that the act of predicting the airway cannot be dissociated from the act of managing the airway. In cancer or an infectious disease, diagnosis is in many respects an isolated intellectual act (e.g. Hodgkins vs. non-Hodgkins lymphoma; falciparum vs. non-falciparum malaria, etc); something that determines but is unaffected by the results of treatment. Failure of response may require rechecking the pathology slides, but does not of itself change the diagnosis. By contrast, a difficult airway is defined solely as an entity resistant to 'routine' airway management. Strictly speaking, it is only those patients falsely predicted to be 'easy' in whom routine airway management is unsuccessful who are truly 'difficult'. The true status of those patients successfully managed by advanced techniques (because they were predicted difficult), may never be known.

Moreover, predicting airway difficulty colours subsequent management. Even if using routine techniques, practitioners may adopt greater care (e.g. more careful laryngoscopy or subconsciously stronger grip during mask ventilation), thus making such patients 'easy' [11]. In reverse, in patients predicted to be 'easy', practitioners may inappropriately adopt a more relaxed, suboptimal approach, resulting in more unanticipated difficult cases [12]. These putative factors are difficult to investigate. To our knowledge the (unethical but key) experiment has not been done, of letting unsupervised novices use only basic airway techniques to manage patients predicted as 'difficult'. Only this would establish if airway prediction really is 'pointless'.

The true predictive utility of a binary approach

As the logical, inevitable conclusion of his analysis, Yentis provocatively proposed a YETI test for airway prediction - simply assigning all patients as 'easy' [5]. Tongue-incheek, he showed how, according to conventional statistical thresholds, YETI would yield test sensitivity, specificity and PPV little different from any more sophisticated predictive test (the provocative conclusion being: 'why bother predicting?').

YETI is correct on purely statistical grounds for hypothetical randomly drawn samples, but fails if applied to an individual patient presenting to us. It would be deeply unimpressive if YETI were applied to a morbidly obese patient presenting after extensive facial reconstructive surgery, a litany of failed intubations and a video of impossible tracheal intubation. Statisticians refer to this as the 'posterior probability'; the extra information that informs real-life predictions. Another descriptor for this is 'plausible'. Whereas terms like 'probable' or 'possible' relate to the relative frequency of the outcome (common or rare, respectively), plausibility better encapsulates the role of judgement based on reasoning, regardless of rarity. A later defence citing the mantra 'we cannot predict the difficult airway' would unlikely persuade the courts, were this hypothetical patient to be disastrously managed only with routine techniques.

It is now increasingly recognised that conventional statistical approaches for common diseases cannot be applied to rare (orphan) disease [13]. The patients in question are potential victims, not beneficiaries, of our standard statistical thresholds for diagnostic tests, and of the stringent requirements imposed by regulatory bodies for large randomised controlled trials before treatments are approved. Specialised approaches have to be adopted for meaningful progress [14].

Extending this concept to anaesthesia, the 'difficult airway' could and should be regarded as a syndrome, composed of very many individual 'rare diseases'. Some of our patients in fact do have orphan genetic disorders that underlie their difficulty, and airway management-related hypoxia has been registered as a syndrome of interest in Oxford University's rare disease initiative (http://www.rarediseases.ox.ac.uk/home). Rather than aim to develop a universal test to predict 'difficulty' that fulfils standard statistical thresholds we should apply our clinical skills to diagnosing the individual rare disease in an ad personam approach. Thus a diagnosis for a rare disease may be entertained as plausible, even if it remains remote in quantitative terms. In turn, we need to maintain the range of specific skills necessary optimally to manage each of these 'rare diseases'.

There is little commonality between each of our 'rare airway diseases' other than that they are 'not easy'. Therefore, one approach is to diagnose and exclude the homogenous 'easy' group that shares the common clinical signs using a binary categorisation, leaving the heterogenous cohort of the 'difficult'. Figure 2 displays the results of this approach. It takes a plot similar that used by Yentis (Fig. 1), but looks at the same data from a different angle. A binary approach divides patients into two groups (red line): those 'predicted easy' (E + E') and those 'predicted difficult' (D + D').

Those 'predicted easy' consist of: (i) those truly easy, who come

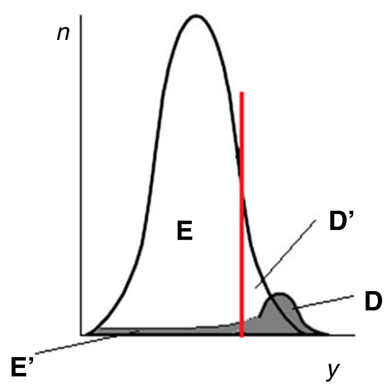

Figure 2 Similar to Fig. 1, the figure is adapted to show that some difficult patients in fact have few or no clinical signs of difficulty. The red line cut-off divides the 'predicted easy' from the 'predicted difficult'. Within the former are the truly easy (E) and also some unanticipated difficult (E'). Within the latter are the truly difficult (D) and also many unanticipated easy (D'). Although the conventional test sensitivities, specificities, PPV remain poor, the proportion of truly difficult patients to the right of the red line is very much higher than the proportion to the left of the line. Expressed another way, to the right of the red line patients are all plausibly difficult (even if some turn out not to be). To the left of the line, patients are all implausibly difficult (even if some turn out actually to be difficult).

to no harm with conventional airway management (E; the vast majority); and (ii) a tiny minority (E') who are unexpectedly found difficult, and in whom advanced airway management needs to be applied after induction, often with some urgency. Group E' is the focus of all unanticipated difficult airway guidelines [15, 16], and also should be a prime focus of difficult airway research (e.g. what is it that makes the trachea of a healthy, thin, edentulous patient with normal jaw structure and good mouth opening difficult to intubate?) [17].

Those 'predicted difficult' (D + D') should logically be subjected to some form of advanced or carefully-planned airway management from the start. The anaesthetist should later be able to explain how the airway management plan differed for this group vs. their routine plan for those 'predicted easy'. For some of these patients (D), these alternative or advanced methods will be fully justified since these patients will have truly difficult airways. However, for others (D'; perhaps the majority in this group), it may transpire that although advanced techniques were successful, they may not have been necessary (e.g. a post-intubation laryngoscopy reveals a Cormack-Lehane grade 1 view). Nevertheless, these patients will remain with cogent clinical signs of airway difficulty, and later anaesthetists may continue to judge that, despite these reassuring findings in the past, subsequent presentation for surgery still requires advanced airway management (e.g. how could one be certain things had not changed?). In essence, the red line in Fig. 2 is one that separates difficulty that is plausible from that which is implausible, regardless of its rarity.

When viewed in this way, the data offers more than crumbs of comfort. Table 1 (from Nørskov et al. [2]) shows the conventionally poor diagnostic sensitivity of 9% and PPV 30%. However, when our binary approach is applied, we see that of every 60 patients 'predicted easy', only around one of 60 will be a surprise (E'). Of every 60 patients 'predicted difficult', about 18 of them will be so (D). Therefore, in our 'predicted difficult' cohort of D+D', true difficulty arises 18 times as commonly (1800%) as in the 'predicted easy' cohort. In fact, in our practice, we think the numbers are even better (of every 100 or so 'predicted easy' patients, just one is a surprise (E'); of every 100 or so 'predicted difficult' patients, around 35 will be truly difficult (D); ~3500% difference). This is a simply staggering performance for any diagnostic test of a rare disease. This is a simplified approach to the notion of the 'likelihood (or diagnostic) odds ratio', touched upon in previous correspondence and elsewhere [18, 19]. Actual values for the data are given in Table 1, and offer some statistical explanation of why we might regard some types of difficulty as 'plausible' vs. 'implausible'.

From prediction to management

We now discuss how this powerful predictive tool might, or should, influence management. Since diagnosis and management are inextricably linked, it follows that management should reflect the difficulty. It makes no sense to predict an airway as difficult, yet to treat it from the outset as any other easy airway. A later inspection of the records should indicate to a dispassionate observer – from the techniques used alone – whether the anaesthetist regarded the airway as plausibly difficult or not.

If this point is accepted, it then follows that there should be defined certain airway management methods classed as 'routine' (e.g. intravenous induction with or without neuromuscular blockade followed by supraglottic airway device, SAD, or tracheal intubation using a Macintosh laryngoscope) vs. 'advanced' (e.g. a range of awake or sedated tracheal intubation methods perhaps using specialised laryngoscopes with or without dedicated pre-oxygenation techniques). To paraphrase Tolstoy again: '*all easy airways may be treated alike; each difficult airway should be approached in its own way*'. An SAD may be recommended as a rescue device in Difficult Airway Society guidelines [15, 16, 20], but it does not follow logically from this that all airways predicted difficult can be managed with an SAD from the outset.

It is beyond the scope of this article to define what best these advanced management plans should be. However, the proposed binary approach acts as a screening test, identifying the cohort in whom specialised techniques should be applied from the start. Each anaesthetist will, or should, have their own binary cut-off (as well as their own methods of 'routine' vs. 'advanced'), and this imaginary red line (Fig. 2) may shift depending on context [21]. The same patient judged close to the line might be managed initially with near-routine methods at midday in a busy theatre suite, but with advanced methods if out-of-hours alone in an isolated operating room [22]. Whereas experts might disagree about the 'best' way to manage a specific case, there is always consensus that the case is plausibly difficult [23]. The chosen technique will depend on the specific 'rare disease' or scenario, individual experience, available equipment, and local standardised protocols [24, 25].

A binary predictive approach will enable individual anaesthetists to gauge their own performance through log-books or audits, and adjust the cut-off appropriately. If they experience too many unexpected difficulties, they may shift their personal cut-off to the left; too many 'unnecessary' advanced techniques and they may shift again back right. In this scheme, YETI sits far to the right; a complete abdication of clinical responsibility, it never makes a diagnosis of difficulty (0%), so statistically compares very poorly with our binary approach (~3500%). Some anaesthetists might prefer to advocate a third 'grey' category of the 'uncertain', but this is binary by default. Since binary means 'easy or not', those in the grey category are 'or not' (i.e. a third, in-between category simply moves the red line to the left).

These proposals, like NAP4, challenge what NAP4 itself recognised are among the current deficiencies in airway practice [26, 27]. All too often, anaesthetists recognise clinical signs indicating the airway may be 'not easy', but then fail to translate these into management plans distinct from routine. This is underlined by the two cases that were tragic drivers to NAP4. Gordon Ewing weighed 124 kg (BMI > 40 kg.m^{-2}); however, at both the hospital where he died after failed tracheal intubation and at the hospital where he (a month previously) underwent a successful anaesthetic, the initial management was with routine methods [28]. Similarly, Elaine Bromiley was known to have had congenital fused cervical vertebrae; yet, again, initial anaesthetic management was no different from routine for an 'easy' airway (e.g. there was even no pre-oxygenation) [29]. At the time, these clinical signs were not felt to dictate a diagnosis of 'difficult', but few could claim that they were ever hallmarks of 'easy'; i.e. difficulty was plausible. In a binary approach, learning from these incidents should cause us to shift our personal cut-offs more than a little to the left (Fig. 2).

Conclusions

As a basic duty as diagnosticians, anaesthetists should predict whether an airway will be easy or not. This should logically guide initial management: routine for the easy, advanced for the plausibly difficult. In this way, the diagnostic process can support and justify the use of advanced techniques. Using the same data as Yentis [5] (Fig. 1), we have reached the opposite conclusion by (a) viewing the data distributions from another angle (Fig. 2); (b) re-defining the entity as a syndrome composed of many rare diseases; (c) regarding predictive tests as screening rather than diagnostic. Osler described medical diagnosis as '*the art of probability*' [30], but this is all more than statistical trickery; clinical management dictated by a binary approach gives the ritual of airway prediction a point.

Competing interests

JP is the Scientific Officer of the Difficult Airway Society (DAS) – the views expressed are his own, and not those of DAS. He is also an editor of *Anaesthesia*. No other external funding or competing interests declared.

J. J. Pandit
Consultant/Professor and Fellow
Department of Anaesthesia
Oxford University Hospitals
Oxford, UK/
St John's College
Oxford, UK
Email: jaideep.pandit@dpag.ox.ac.uk
T. Heidegger
Head of Department/Associate Professor
Spitalregion Rheintal Werdenberg Sarganserland,
Grabs, Switzerland/
University of Bern
Bern, Switzerland

Keywords: airway management; predictive tests; risk; safety; statistics; tracheal intubation

References

1. Nørskov AK, Wetterslev J, Rosenstock CV, et al. Prediction of difficult mask ventilation using a systematic assessment of risk factors versus existing practice – a cluster randomised clinical trial in 94,006 patients. *Anaesthesia* 2017; **72**: doi: 10.1111/anae.13701.
2. Nørskov AK, Rosenstock CV, Wetterslev J, Astrup G, Afshari A, Lundstrøm LH.

Diagnostic accuracy of anaesthesiologists' prediction of difficult airway management in daily clinical practice: a cohort study of 188,064 patients registered in the Danish Anaesthesia Database. *Anaesthesia* 2015; **70**: 272–81.
3. Nørskov AK, Wetterslev J, Rosenstock CV, et al. Effects of using the simplified airway risk index vs usual airway assessment on unanticipated difficult tracheal intubation – a cluster randomized trial with 64,273 participants. *British Journal of Anaesthesia* 2016; **116**: 680–9.
4. Rosenstock CV, Nørskov AK, Wetterslev J. Lundstrøm LH; Danish Anaesthesia Database. Emergency surgical airway management in Denmark: a cohort study of 452,461 patients registered in the Danish Anaesthesia Database. *British Journal of Anaesthesia* 2016; **117** (Suppl. 1): i75–i82.
5. Yentis S. Predicting difficult intubation – worthwhile exercise or pointless ritual? *Anaesthesia* 2002; **57**: 105–9.
6. Feinstein AR. Misguided efforts and future challenges for research on ''diagnostic tests''. *Journal of Epidemiology and Community Health* 2002; **56**: 330–2.
7. Oates JD, Macleod AD, Oates PD, Pearsall FJ, Howie JC, Murray GD. Comparison of two methods for predicting difficult intubation. *British Journal of Anaesthesia* 1991; **66**: 305–9.
8. Yamamoto K, Tsubokawa T, Shibata K, Ohmura S, Nitta S, Kobayashi T. Predicting difficult intubation with indirect laryngoscopy. *Anesthesiology* 1997; **86**: 316–21.
9. Nath G, Sekar M. Predicting difficult intubation – a comprehensive scoring system. *Anaesthesia and Intensive Care* 1997; **25**: 482–6.
10. el-Ganzouri AR, McCarthy RJ, Tuman KJ, Tanck EN, Ivankovich AD. Preoperative airway assessment: predictive value of a multivariate risk index. *Anesthesia and Analgesia* 1996; **82**: 1197–204.
11. Law JA, Broemling N, Cooper RM, et al. The difficult airway with recommendations for management–part 2–the anticipated difficult airway. *Canadian Journal of Anesthesia* 2013; **60**: 1119–38.
12. Lundstrøm LH, Møller AM, Rosenstock C, et al. A documented previous difficult tracheal intubation as a prognostic test for a subsequent difficult tracheal intubation in adults. *Anaesthesia* 2009; **64**: 1081–8.
13. Department of Health. *UK Strategy for Rare Diseases*. London: Williams Lea, 2013.
14. Forman J, Taruscio D, Llera VA, et al. The need for worldwide policy and action plans for rare diseases. *Acta Paediatrica* 2012; **101**: 805–7.
15. Frerk C, Mitchell VS, McNarry AF, et al. Difficult Airway Society 2015 guidelines for management of unanticipated difficult intubation in adults. *British Journal of Anaesthesia* 2015; **115**: 827–48.
16. Marshall SD, Pandit JJ. Radical evolution: the 2015 Difficult Airway Society guidelines for managing unanticipated difficult or failed tracheal intubation. *Anaesthesia* 2016; **71**: 131–7.
17. Shiga T, Wajima Z, Inoue T, Sakamoto A. Predicting difficult intubation in apparently normal patients: a meta-analysis of bedside screening test performance. *Anesthesiology* 2005; **103**: 429–37.
18. Calder I. Useless ritual? *Anaesthesia* 2002; **57**: 612.
19. Grimes DA, Schulz KF. Refining clinical diagnosis with likelihood ratios. *Lancet* 2005; **365**: 1500–5.
20. Mushambi MC, Kinsella SM, Popat M, et al. Obstetric Anaesthetists' Association and Difficult Airway Society guidelines for the management of difficult and failed tracheal intubation in obstetrics. *Anaesthesia* 2015; **70**: 1286–306.
21. Huitink JM, Bouwman RA. The myth of the difficult airway: airway management revisited. *Anaesthesia* 2015; **70**: 244–9.
22. Hung O, Murphy M. Context-sensitive airway management. *Anesthesia and Analgesia* 2010; **110**: 982–3.
23. Cook TM, Morgan PJ, Hersch PE. Equal and opposite expert opinion. Airway obstruction caused by a retrosternal thyroid mass: management and prospective international expert opinion. *Anaesthesia* 2011; **66**: 828–36.
24. Heidegger T. Airway management: standardization, simplicity, and daily practice are the keys to success. *Anesthesia and Analgesia* 2010; **111**: 1073–4.
25. Heidegger T, Gerig HJ, Ulrich B, Kreienbühl G. Validation of a simple algorithm for tracheal Intubation: Daily practice is the key to success in emergencies—an analysis of 13,248 intubations. *Anesthesia and Analgesia* 2001; **92**: 517–22.
26. Royal College of Anaesthetists. *4th National Audit Project: Major Complications of Airway Management in the UK*. London: RCoA, 2011.
27. O'Sullivan E, Laffey J, Pandit JJ. A rude awakening after our fourth 'NAP': lessons for airway management. *Anaesthesia* 2011; **66**: 331–4.
28. Sheriffdom of Glasgow and Strathkelvin. *Determination of Sheriff Linda Margaret Ruxton in Fatal Accident Inquiry in the Death of Gordon Ewing*. 2010 FAI 15. https://www.scotcourts.gov.uk/search-judgments/judgment?id=328e86a6-8980-69d2-b500-ff0000d74aa7 (accessed 14/09/2016).
29. Bromiley M. *The Case of Elaine Bromiley.* www.chfg.org/wp-content/uploads/2010/11/ElaineBromileyAnonymousReport.pdf (accessed 14/09/2016).
30. Osler W. *Aequanimitas: Valedictory Address*, University of Pennsylvania, (1 May 1889). In: Silverman ME, Murray TJ, Bryan CS, eds. *The Quotable Osler*. Philadelphia, PA: American College of Physicians, 2003.

doi:10.1111/anae.13720

