www.nature.com/pr

COMMENT

# Black swans and ambitious overgeneralization in newborn intensive care

David K. Stevenson[1]✉, Ronald J. Wong[1], Gary M. Shaw[1], Nima Aghaeepour[2], Ivana Maric[1], Lawrence S. Prince[1], Jonathan D. Reiss[1] and Michael Katz[1]



*Pediatric Research* (2022) 92:357–359; https://doi.org/10.1038/s41390-021-01771-5

Sometimes progress begins as an anomalous event, an amazing observation that changes one's perspective on what is possible and what is not. However, it would be wrong to assume that such an occurrence would always arise by a random accident. Indeed, an intentional act, maybe even one thought to be futile based on prior experience, in response to a desperate situation might lead to an unexpected result, thus relieving the desperation. Recall Shakespeare's comment in Hamlet: "Diseases desperate grown, by desperate appliance are relieved, or not at all". The purpose of this commentary is to suggest that those in the field of neonatology should consider a more systematic, rather than a haphazard, approach to lowering the "threshold of viability" for the initiation of newborn intensive care, including randomized controlled trials, if greater generalization of practice protocols for this purpose is desired.

Another metaphor that has been applied in such a circumstance is the "black swan" theory (Fig. 1), referring to an amazing event with a significant impact on the course of subsequent events, and often rationalized in hindsight without affirming experimentation to justify the generalization of the experience.[1] Of course, universality, and certainly knowledge of cause and effect, cannot be achieved without the latter. Indeed, without experimentation, there is a risk for "ambitious overgeneralization", for applying what is gleaned from one unique circumstance to others, which may seem similar, but differ in consequential ways, which, at first, might be overlooked. An example of the phenomenon was described by William Silverman as the widespread application of apparently safe, life-saving oxygen therapy, which caused retinopathy of prematurity when applied too broadly and without attention to dose to treat premature infants to reduce mortality and morbidity.[2] Another recent example of ambitious overgeneralization could be the application of phototherapy to increasingly smaller, more immature, and more translucent premature infants without a randomized, controlled trial to evaluate its safety and efficacy in such infants. Only recently, because of a trial that focused on the use of phototherapy to protect against bilirubin-induced injury in such infants, in particular, those weighing between 501 and 750 g, was a potential risk of sustained light exposure identified.[3]

Medical care of the most immature infants, those born at <24 weeks' gestation and often weighing little more than a pound, provides another case in point. The apparent "threshold of viability" has moved ever earlier in gestation during the past 50 years with the introduction of mechanical ventilation, artificial surfactants, and other technical innovations. Whereas survival after birth at 22 weeks' gestation was once perceived as a "black swan" event in prior decades, it has now become much more commonplace. The expectation for survival in some parts of the world has rapidly changed. From 2014 to 2019, survival among all infants born at 22 weeks' gestation in the US Vermont Oxford Network hospitals increased from 6 to 17% of all liveborn infants (and 29% of resuscitated infants). During that time, the rate of resuscitation more than doubled, from 26 to 58%.[4] Today, perhaps the rare and generally unexpected survival of a newborn <22 weeks' gestation could be characterized as a "black swan" event, with occasional case reports of such early survival now showing up in the literature.[5] Will such anecdotes of successful resuscitation prompt more general application of resuscitative efforts at this earlier gestational age?

Of course, the definition of "success" might also differ from the perspective of the parents, neonatologists, nurses, or others in society. Much remains uncertain about why some infants born this early survive and others do not, and even less is known about their conditions of survival and long-term sequelae extending well into adulthood. Individual infants at the same postconceptional age might differ dramatically in their actual maturity, just as adolescents of the same chronological age differ with respect to the time that they experience puberty. Moreover, individual variation in organ maturity may further complicate the vulnerabilities of particular infants, making the application of a generalized approach (based simply on gestational age) fraught with challenges, as the interaction between an infant's biologic capacities and the environment (treatment strategies) gradually unfolds and is revealed in the form of the various possible outcomes. Taking a slightly different perspective, outlier institutions reporting >50% survival of infants born at 22 weeks' gestation could be characterized as "black swans", rather than the infants themselves. Should the reports from such "black swan" centers prompt widespread application of the treatments used at these centers in other neonatal intensive care units (NICUs) in order to achieve similar results?

There are practical consequences for the uncertainty in the medical decision-making about neonatal intensive care that have been described previously.[6] Although the "threshold of viability"

[1]Department of Pediatrics, Stanford University School of Medicine, Stanford, CA 94305, USA. [2]Department of Anesthesiology, Perioperative and Pain Medicine, Stanford University School of Medicine, Stanford, CA 94305, USA. ✉email: dstevenson@stanford.edu

Received: 26 July 2021 Revised: 14 September 2021 Accepted: 18 September 2021
Published online: 2 October 2021

**Fig. 1 Black swan.** The rare occurrence of a black swan in a sea of white swans.

referred to older gestational ages in 1987 when the commentary was published, the philosophical issues related to the resuscitation of premature infants at the "threshold of viability" and decision-making about their care in the NICU have not really changed over the last two decades. The previous article discusses the paradox of the physician's oath to preserve life and avoid suffering and brings into relief the conflict between these two commitments in the practice of newborn intensive care: a "wait until certainty" approach virtually guaranteeing suffering for some infants, not only during their course in the NICU, but perhaps even lifelong. Perhaps, saving a life trumps any morbidity (there are many people who believe this axiom), but the consequences of such an approach must be understood (including equity considerations in the face of limited resources, increased socioeconomic disparities, and increased suffering despite survival). However, alternative approaches have been offered: the "statistical prognostic strategy" and the "individualized prognostic strategy". With respect to the initiation of intensive care, the former simply estimates a particular infant's chances of survival based on all the factors known to affect the eventual outcome and consideration of the caregiver's experience and the current state and availability of technology. The attendant decision is categorical in nature and is generalized without any individual exceptions. For example, for an individual infant, intensive care would not be initiated if the birthweight were <450 g, intensive care would not be undertaken in any case. The purpose of this example is not to suggest that a birthweight alone should be what determines the "threshold of viability", but only that some agreed-upon criteria are the basis for a "yes" or "no" decision about proceeding with intensive care, independent of any other individual characteristics or responses to intensive care. On the face of it, such an approach ensures less suffering caused by intensive care in the NICU and later in life, but it also conflicts with the principle of saving a life, if possible. Moreover, such an approach might also stall the advancement of intensive care technology and innovations that could improve newborn intensive care overall. Some neonatologists have assumed this disposition, and they are consequently late adopters of any new technologies or approaches that might lower the "threshold of viability".

The "individualized prognostic strategy", on the other hand, provides for some "reasonable restraint" of the application of newborn intensive care while allowing for the predominance of the principle of saving a life and advancing newborn intensive care. That is, in addition to whatever criteria is applied for the initiation of intensive care, other individual characteristics and responses to care would and should also be considered. The individualized prognostic strategy, while initiated for as many patients as possible, as in the "wait until certainty" approach, is re-examined at intervals based on both statistical information and an infant's response to therapy. Intensive care is continued for infants who have shown improvement. For the infants with continued poor development and prognoses, intensive support is withdrawn. This re-evaluation has the purpose of minimizing unnecessary suffering. With the "wait until certainty" approach that maximizes a physician's commitment to saving a life, increasing the burden of suffering in many cases would be ensured. Therefore, taking a hybrid approach of the individualized prognostic strategy might achieve the desired balance between the twin objectives of the physician's oath. Although the proposed algorithm for implementing the "individualized prognostic approach" in the older article might be outdated, the notion of "individualized prognostication" remains a sound one. Indeed, the approach begs for improvements in individual prognostication and perhaps a whole new taxonomy of prematurity (not simply defined by gestational age or birthweight), which can differentiate and predict the clinical trajectories, including likely complications of treatment, even before birth. From a statistical perspective, "statistical" and "individualized" prognostic strategies are not that different. They both form a decision based on a prediction of the survival probability; in the former case in a single-stage process, and in the latter a multistage process. With the help of improved statistical techniques, including machine learning, combined with deeper knowledge about the biological disposition of a particular pregnancy and fetus, better predictions seem feasible.[7,8] Also, recent developments in artificial intelligence (e.g., long short-term memory neural networks)[9] have provided us with unique capabilities to study "black swan" events in large medical databases. However, these sophisticated tools themselves are challenging to use from an ethical perspective and must ultimately be guided by principles carefully designed by society.[10] A statistical model attempting to predict the outcome of an individual infant born at the "threshold of viability" would likely be developed on the basis of a small number of "black swan" cases (that could be biased). Although it is possible that advances in transfer learning[11] could help alleviate the problem of limited data, the prediction value obtained by any statistical algorithm should only serve as a guide to the caretakers and parents. Here, once we obtain an accurate prediction for each neonate as an output of a machine-learning algorithm, that value should quantify two, sometimes conflicting, objectives: minimization of mortality and minimization of suffering, making the decision-making even more difficult. Any building of models to facilitate such decisions could be one of the most consequential uses of statistics and should be undertaken with the utmost care. In the end, such computational tools will only serve to support shared decision-making involving parents and caretakers.

There is one more important point that must be made in this context. Rather than a haphazard approach to lowering the "threshold of viability" because of different institutional policies or healthcare team dispositions in regard to the initiation of newborn intensive, a systematic approach to understanding what can be achieved in this regard, including randomized, controlled trials should be considered, if generalization is desired. Although the "wait until certainty" and the "statistical prognostic strategy" approaches would make the assumption of equipoise or the implementation of a randomized, controlled clinical trial to test a particular intervention algorithm difficult, if not impossible, the "individualized prognostic strategy" might lend itself to such a systematic approach. At least, the experiment to test whether the "threshold of viability" might be lowered in a manner consistent with the physician's complex oath to save lives and avoid suffering could be undertaken, with the understanding that the preceding hypothesis might be accepted or rejected. However, at least if an individual infant's response to whatever resuscitative and treatment efforts were systematically applied, this would have been considered in the final decision whether to continue with neonatal intensive care.

In conclusion, a dependence on clinical anecdotes or case series of ad hoc attempts to resuscitate and provide intensive care to increasingly immature infants will never do more than identify "black swan" events. Only a properly randomized study will provide a legitimate rationalization for a universal approach to the initiation of newborn intensive care. On the other hand, who is to say that there may not be "red swans" (or even less common events)? Sometimes the presence or absence of something is only a matter of frequency. Where the line is drawn for the "threshold of viability" might be better understood as the line where equipoise can be assumed by caring healthcare practitioners and parents—below which experimentation can be undertaken, and above which treatment should be implemented.

## REFERENCES

1. Taleb, N. N. *The Black Swan: The Impact of the Highly Improbable* 1st edn (Random House, 2007).
2. Silverman, W. A. Ambitious overgeneralisation. *Paediatr. Perinat. Epidemiol.* **16**, 288–289 (2002).
3. Morris, B. H. et al. Aggressive vs. conservative phototherapy for infants with extremely low birth weight. *N. Engl. J. Med.* **359**, 1885–1896 (2008).
4. Rysavy, M. A. et al. An immature science: intensive care for infants born at ≤23 weeks of gestation. *J. Pediatr.* **233**, 16–25.e11 (2021).
5. Ahmad, K. A., Frey, C. S., Fierro, M. A., Kenton, A. B. & Placencia, F. X. Two-year neurodevelopmental outcome of an infant born at 21 weeks' 4 days' gestation. *Pediatrics* **140**, e20170103 (2017).
6. Fischer, A. F. & Stevenson, D. K. The consequences of uncertainty. An empirical approach to medical decision making in neonatal intensive care. *JAMA* **258**, 1929–1931 (1987).
7. Villar, J. et al. Association between preterm-birth phenotypes and differential morbidity, growth, and neurodevelopment at age 2 years: results from the INTERBIO-21st Newborn Study. *JAMA Pediatr.* **175**, 483–493 (2021).
8. Espinosa, C. et al. Data-driven modeling of pregnancy-related complications. *Trends Mol. Med.* https://doi.org/10.1016/j.molmed.2021.01.007 (2021).
9. Hochreiter, S. & Schmidhuber, J. Long short-term memory. *Neural Comput.* **9**, 1735–1780 (1997).
10. Russell, S., Hauert, S., Altman, R. & Veloso, M. Robotics: ethics of artificial intelligence. *Nature* **521**, 415–418 (2015).
11. Torrey, L. & Shavlik, J. In *Handbook of Research on Machine Learning Applications and Trends: Algorithms, Methods, and Techniques* (eds Olivas, E. S., Guerrero, J. D. M., Martinez-Sober, M., Magdalena-Benedito J. R. & Lopez, A. J. S.) 242–264 (IGI Global, 2009).

## ACKNOWLEDGEMENTS

We would like to acknowledge Matthew A. Rysavy, MD, PhD, whose presentation to the Stanford Prematurity Research Center on June 23, 2021 entitled "Black Swans: Learning from Deviations in Premature Infant Care", inspired this commentary.

## AUTHOR CONTRIBUTIONS

All authors have met the *Pediatric Research* authorship requirements. D.K.S., R.J.W., G.M.S., N.A., and M.K. have made substantial contributions to conception and design. D.K.S., R.J.W., G.M.S., N.A., I.M., L.S.P., J.D.R., and M.K. have made substantial contributions to drafting the article or revising it critically for important intellectual content and have approved the final version to be published.

## FUNDING INFORMATION

This work was supported in part by the Charles and Marie Robertson Foundation, the Christopher Hess Research Fund, the Charles B. and Ann L. Johnson Endowed Fund, and the Prematurity Research Fund.

## COMPETING INTERESTS

The authors declare no competing interests.

## CONSENT STATEMENT

No patient consent was required for this commentary.

## ADDITIONAL INFORMATION

**Correspondence** and requests for materials should be addressed to David K. Stevenson.