Validity, credibility, and applicability: the rise and rise of the surrogate

The elusiveness of outcome data of direct relevance to patients makes research in anaesthesia, pain, and critical care a challenge. Surrogate measures are used for several reasons. Instead of measuring the risk of organ injury, death, or long-term personality change after surgery and anaesthesia, we often measure arterial pressure, partial pressure of carbon dioxide, hospital stay, or short-term cognitive performance. Is the use of such proxy outcome measures justified? In modelling research, we often turn to physical or theoretical models, generating a single, idealized or ‘typical’ subject upon which to experiment. Is such substitution of reality with simulation reasonable? In basic science, the individual components of physiology or pharmacology are studied in relative isolation. Is it useful to know the isolated effects of x on y when performing our clinical work?

The answer to these questions is found in the twin issues of credibility (also termed face validity) and validity. Research related to anaesthesia and critical care abounds with surrogates that are replacements for outcomes, subjects, or methodologies. This is because answering the questions that we really wish to ask is often impractical for scientific, ethical, or financial reasons. It then becomes acceptable to substitute questions that have obtainable answers for the questions that we wanted to ask. The real answers, those that contain absolute truth, may be unobtainable.

When designing studies or reading research papers, it is important that we are sensitive to the issues of real interest being substituted by surrogates. We should always question the validity and applicability of the surrogate. The degree of a surrogate’s validity will dictate how well it represents the phenomenon under study. For example, in assessing a patient’s arterial pressure response to laryngoscopy, we have no fundamental interest in the arterial pressure, as such, but rather in its effects on vital organs, particularly the heart and brain. Thus, we think it reasonable to conduct the research with the implicit assumption that the demonstrated effects of laryngoscopy indicate the effects on the heart and brain. Even here, the details of the surrogate are important. Which aspects of arterial pressure should be measured: the absolute increase in pressure, the duration of increase, or both? Each factor may have a bearing upon how well the chosen surrogate maps onto the outcome of real interest. At best, the authors using such surrogates should discuss their relevance, validity, and limitations. At worst, readers should spot the use of these proxies and ask whether the authors’ conclusions are based acceptably in their findings or whether the assumptions made about inter-changeability of the surrogates and the real outcomes are not justified.

Validity and applicability have even greater impact when we are assessing research using models, be they computer simulation models, bench models, in vitro models, or animal models. How well does the response of an isolated porcine coronary artery represent the human, septic vascular bed? How well does a computational model of the respiratory system represent the complex, diseased, and intact human? How well does a construct using plastic tubing and a funnel represent the deformable, viscoelastic, mucous-covered human upper airway? The answer in all these cases depends upon the use to which the model is being put. The home-made airway model may very reasonably be examined to judge the effect of jet-ventilation on airway pressures, but would not make a credible subject for examining aerosol distribution delivered via various nebulizing devices. Similarly, experiments on isolated arteries or cells can identify a number of receptors that may be relevant, but will not give a full picture of comprehensive physiological response when the receptors are stimulated. Thus, validity is task-specific. Models will never be entirely faithful representations of reality, but may get close enough to the issues that matter so as to make their use credible.

The credible use of surrogates is important in that it draws attention to the relationship between the researcher and the consumer of research (i.e. the reader). A researcher’s model-based research is only successful if the model
is credible in the eyes of the consumer. Such credibility may be established by the researcher through comparing the performance of the model in tasks with empirical data, by describing the model in such detail so as to leave the reader in no doubt as to the veracity of the model, or by convincing the reader to assume the model’s validity through sleight-of-hand. It is of the last of these that the reader must beware, as whenever a model is presented as a surrogate for the organism/process of interest, the reader must ask if this model is valid for the task to which it is put. Occasionally, models are ‘sold’ as valid by citation of previous validation exercises. Again, the reader must beware; the model is only applicable to the task in hand (and thus credible) if it has been shown to be acceptably valid in the arena of interest.

Surrogates are heavily used in translational research as, at the boundaries between disciplines and between researchers, surrogates provide continuity. Here, imperfect inter-disciplinary communication may lead to misinterpretation or over-valuing of surrogates and imperfect conclusions may be reached. The path from treatment to outcome or from disease to end-organ effect is long and usually impossible to visualize as a single researcher. Here, a variety of surrogates may be necessary, often applied as a chain leading between the input (e.g. disease or treatment) and the outcome (e.g. end-organ effect). For example, researchers considering the impact of administered colloid solutions on outcome may consider outcome to be adequately represented by renal function, in turn, represented by serum creatinine measurements. Here, a series of assumptions is made to make the research feasible. It is important that the validity and credibility of such assumptions are addressed at every stage of research—at its inception, design, conduct, analysis, and publication.

Surrogates are also used to build our mental concepts for pathology and physiology. But this can then distort our understanding and research, if the surrogate then becomes the end point. Problems have been created by application of surrogate-derived models to clinical conditions. Sepsis is a well-known example, with numerous negative trials of a ‘magic bullet’ based on the assumption that suppression of a surrogate marker (e.g. TNF, IL-1 receptor, endotoxin, etc.) would result in improved outcomes.

This is not to say that the use of surrogates is wrong. Far from it! They are often the foundations upon which good research is built. The use of surrogates is essential for the progress of research, and therefore, the understanding of their strengths and limitations among the researchers and the consumers is crucial. It is vital that the researchers show restraint when estimating the potential of a surrogate. It is equally important that the readers challenge the use of surrogates on the grounds of validity. Above all, scientific journals (i.e. reviewers and editors) have a responsibility to ensure that the applicability and the limitations of a surrogate to a given research question are fully addressed before a paper is published.

J. G. Hardman*
I. K. Moppett
R. P. Mahajan

University Department of Anaesthesia
Queen’s Medical Centre
Nottingham NG7 2UH
UK
*E-mail: j.hardman@nottingham.ac.uk

References
5 Ferreira JC, Chipman DW, Kacmarek RM. Trigger performance of mid-level ICU mechanical ventilators during assisted ventilation: a bench study. Intensive Care Med 2008; 34: 1669–75