



Learning from clinical medicine to improve the use of surrogates in ecology

Philip S. Barton, Jennifer C. Pierson, Martin J. Westgate, Peter W. Lane and David B. Lindenmayer

P. S. Barton (philip.barton@anu.edu.au), J. C. Pierson, M. J. Westgate, P. W. Lane and D. B. Lindenmayer, Fenner School of Environment and Society, The Australian National Univ., Canberra, Australian Capital Territory, 0200, Australia.

Surrogates are used widely in ecology to detect or monitor changes in the environment that are too difficult or costly to assess directly. Yet most work on surrogates to date has been correlative, with little work on their predictive capacity or the circumstances under which they work. Our suggestion is to revisit and learn from research in the clinical medical sciences, including the causal statistical frameworks available to validate relationships between treatments, surrogate variables, and the outcome of interest. We adapt this medical thinking to ecology by providing a new framework that involves specification of the surrogate model, statistical validation, and subsequent evaluation in a range of spatial and temporal contexts. An inter-disciplinary surrogate concept will allow for a more rigorous approach to validating and evaluating proxy variables, thus advancing the selection and application of surrogates in ecology.

Synthesis

We draw together ideas from the medical sciences to define an explicit surrogate concept that has not previously been used in ecology. We present a new framework for specifying surrogate models involving validation using a causal framework, and subsequent re-evaluation in different spatial and temporal contexts - an approach closely aligned with that used by researchers in the clinical medical sciences. This rigorous method can advance the science underpinning the application of surrogates in ecology by shifting the focus away from correlative understanding to one that focuses instead on causation and prediction. An improved use of surrogates is imperative if we are to meet the challenge of properly measuring and understanding the multifarious and complex problems in contemporary ecology.

Why surrogates are important and improvement is needed

Surrogates are used in many areas of ecology and the environmental sciences as faster or cheaper substitute measures of populations, ecosystems or ecological processes (Caro and O'Doherty 1999, Feld et al. 2009, Lindenmayer and Likens 2011). The significant appeal of surrogates has led to their study and application in a wide range of fields, such as conservation planning (Rodrigues and Brooks 2007), and the monitoring of pollution (Goodsell et al. 2009), water quality (Niemi et al. 2004) and biodiversity (Noss 1990). Each of these fields informs important policy and management decisions, such as the location of reserves for the protection of biodiversity (Rodrigues and Brooks 2007), or the risk of disease transmission from water-borne bacteria (Sinclair et al. 2012). The need for informative surrogates in ecology and the broader environmental sciences is therefore important.

Despite the widespread application of surrogates, much of the progress in their selection and application has been based on correlative findings. For example, there has been

a substantial amount of work on which species (Caro 2010, Branton and Richardson 2011, Lindenmayer et al. 2014) or environmental variables (Faith and Walker 1996, Gollan et al. 2009, Barton et al. 2014) are associated with variation in the diversity of sets of taxa. In contrast, there has been limited progress made on how to use surrogates in a predictive context (but see McGeoch et al. 2002, Fleishman et al. 2005). In particular, few studies have examined changes in a surrogate variable 1) with specific reference to a desired outcome and 2) as a consequence of a particular treatment. An unclear understanding of a desired outcome reduces the potential to find a relevant and useful surrogate of that outcome. How a surrogate and outcome behave in response to a specific treatment is central to understanding how the former can predict the latter.

We suggest that the lack of conceptualisation of the surrogate concept as a variable in the pathway between treatment and outcome has retarded the selection and application of surrogates in several important areas, including some that are critical to the measurement and monitoring of ecosystems. This includes detecting and monitoring changes in biodiversity, air and water quality, or determining the success of

management interventions aimed at restoring populations of species or aspects of ecosystem function. It is in these areas that surrogates hold great promise to provide useful insight into otherwise complex or difficult-to-measure problems.

Here we present the surrogate concept as it is used in contemporary clinical medicine, especially in the fields of clinical pharmacology and therapeutics (Prentice 1989, Atkinson et al. 2001, Aronson 2005, Fleming and Powers 2012). The use of surrogates in clinical medicine has a history dating back several decades (Aronson 2005), and has helped provide information about patient responses to treatments or disease progression earlier, more quickly, and more cheaply than direct (and sometimes invasive) alternatives (NIH Definitions Working Group 2000, Atkinson et al. 2001, Aronson 2005). By adapting concepts and terminology from this established body of medical theory, we show how an explicit surrogate model that links a treatment to an outcome via a surrogate can be used in ecology. We suggest that by learning from medicine, ecologists will have at their disposal a more rigorous approach to the selection, validation and evaluation of surrogate variables. We describe how this can be done by first presenting the medical perspective to surrogacy and its accompanying conceptual model. We then present a new framework for how this knowledge can be applied to ecological problems.

The medical perspective

Surrogates, or surrogate endpoints, have been used in clinical medicine for several decades (Brotman and Prince 1988, Boone and Kelloff 1993, Aronson 2005). A surrogate endpoint is a specific biological or pathological process used in lieu of the clinical endpoint, the latter being 'a characteristic or variable that reflects how a patient feels, functions or survives following a therapeutic treatment (e.g. a drug) or during or after progression of disease (NIH Definitions Working Group 2000, Atkinson et al. 2001). In such scenarios, the clinical response of the patient is the true variable of interest, but its observation may be either impractical (improvement of symptoms may take too long) or even unethical (effects may potentially be adverse). For example, elevated blood pressure (the surrogate) is used as a physiological predictor of stroke risk (clinical endpoint) because it provides simpler, cheaper and faster knowledge of a patient than more complex scans or tests (Aronson 2005). Similarly, levels of blood serum thyroid stimulating hormone (the surrogate) is used as a biochemical predictor of hypothyroidism (the clinical endpoint) (Aronson 2005).

A surrogate endpoint in medicine is therefore conceptually analogous to a surrogate used in ecology. For example, changes in the species richness of a particular taxon (the surrogate) is commonly used as a predictor of changes in the diversity of a broader set of biodiversity (the true variable of interest) (Rodrigues and Brooks 2007, Westgate et al. 2014). Compared with medical surrogates, however, ecological surrogates often lack this clear conceptualisation. In the absence of this conceptualisation, it is difficult to communicate a standardised way to assess whether a surrogate predicts an outcome, or to determine how a surrogate performs in relation to a treatment effect. One way to address this shortcoming is through the use of an explicit

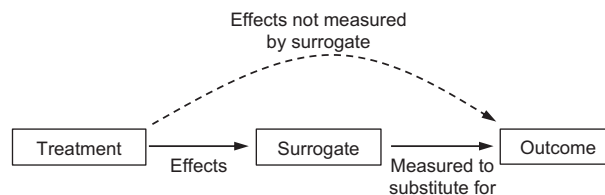


Figure 1. A conceptual model relating a treatment effect to a specified outcome via a surrogate variable (solid lines) (adapted from (Atkinson et al. 2001)). This surrogate concept acknowledges the independent effect of the treatment on the outcome, which may not be detected or represented by the surrogate (broken lines).

surrogate concept that links the treatment to the outcome via the surrogate, which we detail below.

An explicit surrogate concept and its advantages

We show in Fig. 1 an adaptation of the clinical surrogate concept as described by the National Institutes of Health Biomarkers Definitions Working Group (Atkinson et al. 2001). In this model, the authors clearly link a treatment (e.g. a therapeutic intervention) to an outcome (e.g. patient health) via a surrogate biomarker. Fig. 1 also identifies the effect of the treatment on the outcome that may not be accounted for by the surrogate.

Treatments can be any effect, but in clinical medicine they might include the administration of a drug, or cessation of exposure to a pathogenic agent. Placing the surrogate in the pathway of effects between the treatment and the outcome enables the examination of the potential effect of the surrogate and its ability to predict the outcome (Atkinson et al. 2001, Fleming and Powers 2012). Having a known mechanistic or causal link between the treatment and surrogate, or between the surrogate and outcome, greatly increases the likelihood of the surrogate being useful (Aronson 2005). The choice of outcome measures is understood to be one of the most important considerations in designing a clinical trial (Fleming and Powers 2012). This involves clearly defining what variable is of true interest and what level or degree of change in this variable is informative. Such a precise outcome increases the usefulness of information provided by the surrogate in terms of when a goal or outcome is likely to have been reached.

Depicting the relationship between treatment, surrogate and outcome variables in the format of Fig. 1 identifies the pathways of potential effects. This provides clear advantages to the assessment of surrogates: namely, the systematic examination of the strength of association between 1) treatment and outcome, 2) treatment and surrogate, and 3) surrogate and outcome. This is critical to the statistical validation of surrogates, which we discuss next.

When is a surrogate valid?

Validation of surrogates in medicine is more explicit than validation in ecology. For example, a general definition used for validation of an environmental indicator is that "it is well founded and achieves the overall objectives or it produces the intended effects" (Bockstaller and Girardin 2003, p. 641) (Table 1). This might be achieved if it is scientifically

Table 1. Glossary of terms, their definitions and some examples of similar or redundant terms.

Term	Definition	Similar or redundant terms
Treatment	An effect of interest; can be anthropogenic, natural or experimental	An intervention; a contrast or comparison of interest
Surrogate	An ecological element or process that is used to represent another aspect of an ecological system	A proxy variable or indicator (Lindenmayer and Likens 2011)
Outcome	An ecological element or process of specific interest	An objective (sensu McGeoch 1998) or endpoint (sensu Atkinson et al. 2001)
Causal framework	A conceptual model of the pathway of effects from treatment to outcome via a surrogate and other covariates (Joffe and Greene 2009)	Path analysis and structural equation modelling can also be considered as ‘frameworks’ for the systematic examination of variables within a model system (Grace et al. 2010)
Validation	A statistical assessment of a surrogate. Rejection of the null hypothesis of “no relationship between the surrogate and the outcome, and confirmation of the ability of a surrogate to predict the effect of a treatment on the outcome” (Aronson 2005, Weir and Walley 2006)	A variable is validated if it is “well founded and achieves the overall objectives, or it produces the intended effects” (Bockstaller and Girardin 2003)
Evaluation	A broader statistical assessment of a surrogate in a range of circumstances; can be determined by testing in different ecosystems and over time	Robustness (sensu McGeoch 1998, Barton et al. 2014) and generality (sensu Chown and McGeoch 2011)
Specificity	The degree to which a surrogate is associated with a specific outcome (Aronson 2005)	Precision and accuracy (Niemi and McDonald 2004)
Consistency	Persistence of an association between the surrogate and outcome in different circumstances (Aronson 2005)	As for evaluation (see above)

designed, the information it supplies is relevant, and if it is useful to end users (Bockstaller and Girardin 2003, p. 641). By contrast, surrogate validation in clinical medicine involves the explicit assessment of how a surrogate behaves in conjunction with the other variables (Buyse et al. 2010). A surrogate is generally regarded as valid when it allows correct inference to be drawn regarding the effect of a treatment on the outcome (Weir and Walley 2006). For this to occur, a strong association needs to be known or established between the surrogate and the outcome (Begg and Leung 2000). Thus, a valid medical surrogate is “a response variable for which a test of the null hypothesis of no relationship to the treatment is also a valid test of the corresponding null hypothesis of no relationship between the outcome and treatment” (Prentice 1989, p. 432). However, this definition is thought to be true only if the directionality of relationships is explicit (VanderWeele 2013). Alternatively, then, a valid surrogate can be identified if “knowing the effect of a treatment on the surrogate allows prediction of the effect of treatment on the outcome” (Joffe and Greene 2009, p. 530). To achieve validation of a surrogate, a candidate surrogate must therefore be shown to be able to predict the effect of a treatment on the outcome. From a medical perspective, this means that the surrogate must be correlated with the outcome of interest (e.g. survival or pathological improvement) in different patients (Buyse et al. 2010). Importantly, quantifying the magnitude and direction of effects is critical to the validation process, as a surrogate response to a treatment needs to represent the size and direction of the outcome response to a treatment. Ecologists should adopt these explicit hypotheses and definitions for a valid surrogate.

Clearly, a detailed statistical understanding of the interrelationships between treatment, surrogate and outcome variables is required for validation of the surrogate concept outlined in Fig. 1. One of the main ways that statisticians have approached the problem of surrogate validation is with statistical causal frameworks (Box 1) that build

upon the model shown in Fig. 1. Although established for use in clinical medical sciences, we suggest these frameworks provide a well-developed foundation for the validation of surrogates in ecology. This is because the different statistical approaches, and their advantages and limitations, have already been examined in detail, and there is a rich available literature for application to ecological problems (Prentice 1989, Buyse and Piedbois 1996, Buyse and Molenberghs 1998, Begg and Leung 2000, Molenberghs et al. 2002, Taylor and Wang 2002, Weir and Walley 2006, Gilbert et al. 2008, Buyse et al. 2010, Li et al. 2011, Ghosh 2012).

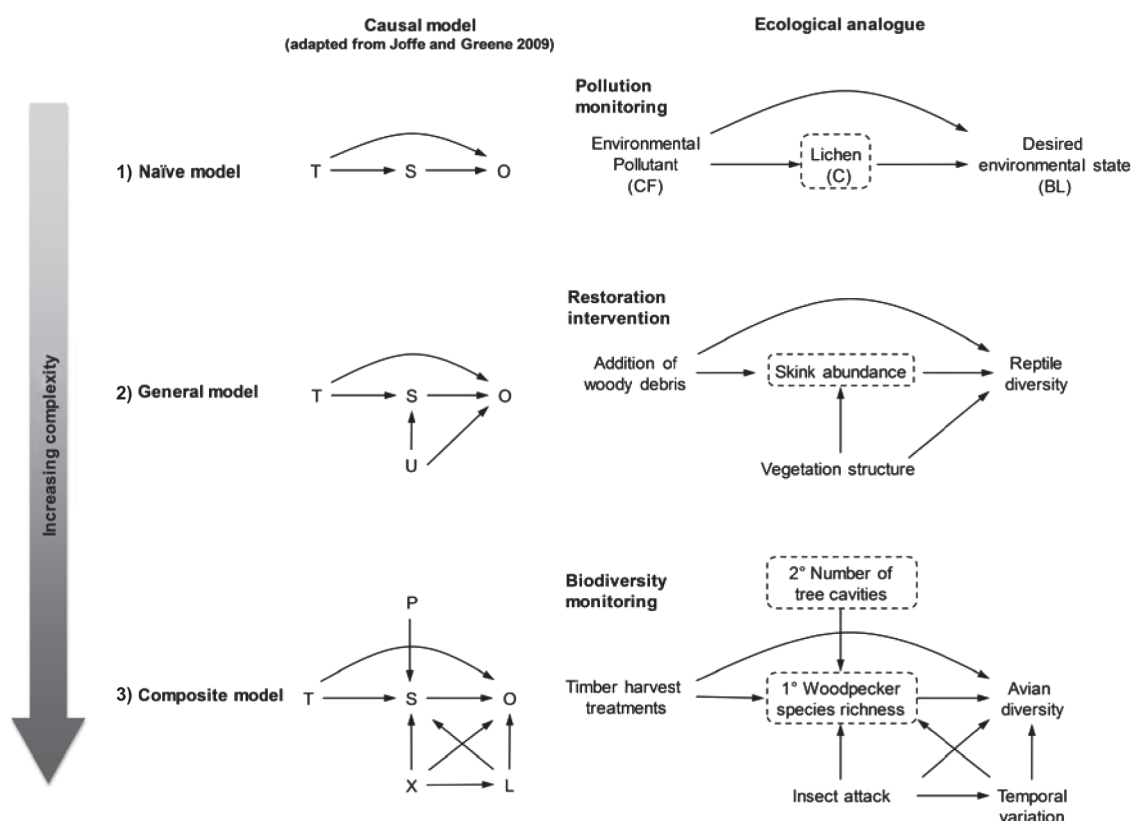
A framework for the development of surrogates in ecology

Building on the concepts and terminology described above, we next outline a new framework to improve the development of surrogates in ecology. The framework includes 1) specification of the surrogate model (Fig. 1), 2) validation of the surrogate (Box 1), and 3) evaluation of the surrogate (Box 2). We outline our framework below.

(1) Specification of the surrogate model

For the medical surrogate concept to be applicable to ecology, attention must be given to describing a measurable outcome following a treatment. Yet, the precise description of goals, objectives or endpoints (Table 1) remains difficult in many fields of applied ecology, such as biodiversity monitoring (Lindenmayer and Likens 2010) and adaptive management (Westgate et al. 2013). The importance of clear objectives also has been highlighted by Chown and McGeoch (2011) in steps 1–3 of their bioindicators selection process. The use of treatment surrogates, however, necessitates the clear identification of an outcome so that the validity of a surrogate can be assessed (Fig. 1). Linking the treatment effect to the objective, via a surrogate variable, reveals the basic surrogate model to be examined. The selection of candidate

Box 1. Causal frameworks and surrogate validation



Surrogate validation involves the examination of the strength of associations between treatment, surrogate and outcome variables. Quantifying the magnitude and direction of effects and associations is critical to the validation process, as a surrogate response to a treatment should represent the size and direction of the outcome response to a treatment. The widespread use of surrogates in clinical medicine, and the strong need for accuracy and rigour, has resulted in a considerable body of research dedicated to the statistical validation of surrogates (Prentice 1989, Buyse and Molenberghs 1998, Colburn 2000, Arkinson et al. 2001, Molenberghs et al. 2002, Taylor and Wang 2002, Weir and Walley 2006, Joffe and Greene 2009, VanderWeele 2013). Central to this has been the development of causal frameworks, which have been designed to examine the links between treatments and their effect on patients, and how well surrogate variables reflect these links (Joffe and Greene 2009, Buyse et al. 2010, VanderWeele 2013). Importantly, the use of causal frameworks to examine surrogate models in medicine is conceptually analogous to the use of path analysis and structural equation modelling used to establish understanding of ecological systems (Shipley 2009, Grace et al. 2010). This involves the specification of a model that considers the direct and indirect effects of a set of variables of interest. The causal frameworks used in medicine range in complexity, but build upon the basic 1) naïve model which links the effect of the treatment (T) to the outcome (O), the treatment to the surrogate (S), and the surrogate to the outcome (see also Fig. 1). Models can incorporate additional sources of variability, with the 2) general model also incorporating the effects of other covariates (U) on the surrogate and response variables. The 3) composite model has further flexibility in terms of linking additional covariates (denoted by X and L) and their interrelated effects on the surrogate and response variables. The ideal situation is for these covariates to be measured and incorporated into the analyses. The statistical methods that might be used to determine associations among variables within a causal framework might include 1) conditional independence of variables, 2) analysis of direct and indirect effects with path analysis, 3) meta-analysis, and 4) principal stratification of variables (Prentice 1989, Joffe and Greene 2009, Shipley 2009, VanderWeele 2013). Each of these approaches has different data requirements and we encourage readers to consult with statisticians and to explore these methods in the literature.

We give three ecological analogues to illustrate how each causal framework might be used to validate a surrogate for three different situations (see diagram above). 1) A naïve model can apply to the problem of air pollution monitoring, such as the use of lichens as indicators of atmospheric contaminants (Conti and Cecchetti 2001). In this scenario, the concentration of the contaminant is measured in the atmosphere (C) and in the tissue of lichens (CF), and is assessed relative to the desired background level (BL). 2) A general causal model might be applied to an ecosystem restoration problem, such as the application of an intervention aimed at increasing animal diversity in a range of vegetation types (Manning et al.

(Continued)

Box 1. Continued

2013). Here, skink abundance can be measured as a surrogate for overall reptile diversity, and assessed in different types of vegetation (see diagram above). 3) A composite causal model might be applied to a biodiversity monitoring problem such as the use of woodpeckers as indicators of general avian diversity under a range of disturbance pressures (Drever et al. 2008). Notably, the composite model not only allows for incorporation of multiple covariates (e.g. areas of insect attack, different survey years), but also incorporates primary and secondary surrogate variables (see diagram above). In this case, the number of tree cavities is a secondary surrogate for the primary surrogate of woodpecker species richness. In all models, unexplained variation in surrogate or outcome responses to treatments could be attributable, in part, to unmeasured and/or unknown covariates. Further, many situations in ecology are likely to involve a composite model, but might begin with testing a naïve or general model. As additional covariates are measured, and understanding of the system is developed, an adaptive approach to surrogate validation might be taken.

surrogate variables should, at least in the first instance, be based on an understanding of mechanisms that link the surrogate to the objective (Pierson et al. 2015), as well as other *a priori* suitability criteria (*sensu* McGeoch 1998) such as cost-effectiveness, ease of measurement, or potential usefulness.

(2) Validation of the surrogate

This step involves statistically quantifying relationships among variables within a causal framework that identifies the pathways of potential effects from the treatment to the surrogate and outcome, and establishing where variation is explained. Unexplained variation in variable responses might be attributable, in part, to unmeasured and/or unknown covariates. This is conceptually analogous to the use of path analysis or structural equation modelling used to establish a scientific understanding of models (Shipley 2009, Grace et al. 2010). Once a surrogate model is constructed, various methods can be used to examine the effect of the treatment on the surrogate and outcome. We outline example frameworks and list some of the statistical tools and approaches available to validate surrogates in the medical sciences in Box 1. Using these tools, a test of the null hypothesis can then be made of no relationship between the response of the surrogate and the response of the outcome to a treatment. This part of our framework is conceptually analogous to steps 4–6 of the bioindicators selection procedure outlined by Chown and McGeoch (2011, Table 18.2); i.e. the accumulation of data and statistical tests of the relationship between the surrogate and the outcome of interest.

(3) Evaluation of the surrogate

The evaluation of surrogates is recognised as a challenging task even in medicine (Buyse et al. 2010), where research can be conducted in a strictly controlled environment. By contrast, the evaluation of surrogates in ecology is especially challenging due to the larger number of variables and the limited ability to control for them. Just as anatomy and physiology will differ between patients, important differences will also occur in the composition and dynamics of communities and ecosystems, and responses can be unpredictable and highly variable. For this reason, evaluation is a critical part of developing confidence in a proposed surrogate, including surrogate specificity and consistency (Box 2, Table 1), and learning when it works in different ecosystems and how it

behaves over time (Barton et al. 2014, Pierson et al. 2015). This part of our framework is similar to steps 7–9 of the bioindicators selection procedure outlined by Chown and McGeoch (2011, Table 18.2) and (McGeoch 1998); i.e. the repeated testing of the surrogate under different conditions to assess robustness and generality (Table 1).

Of course, certain kinds of ecological research are better suited than others to the medical surrogacy approach. This includes research with clear-cut management or restoration interventions, well-defined and measurable surrogate variables, and a reduced set of known covariates. Below, we retrofit our framework to an existing biodiversity monitoring program to show how it might be re-framed conceptually from a medical surrogacy perspective. We also note that such a surrogacy approach should include a learning process where new information is incorporated and fed back into the surrogate model to aid in the scientific understanding of the problem.

Application of the framework

Our framework can be applied to a range of ecological problems that are similar to medical scenarios that measure change in a variable of interest over time, and in an experimental setting. This includes, for example, various agri-environment schemes established to measure changes in taxon diversity over time and in response to specific land-use changes or interventions (Kleijn et al. 2006, Perkins et al. 2011). These schemes have several key analogies with a medical case: monitoring sites = patients (subjects); land management intervention = therapeutic intervention (treatment); taxon diversity at sites = surrogate biomarker (surrogate); increases in taxon diversity = improved health of patients (outcome).

An exemplar agri-environment scheme is the Australian Government's Environmental Stewardship Program [ESP] (Zammit et al. 2010) which aims to maintain and improve the condition and extent of targeted Matters of National Environmental Significance on private land by using market-based approaches (Lindenmayer et al. 2012). Specifically, the ESP has targeted remnant patches of critically endangered White Box-Yellow Box-Blakely's Red Gum Grassy Woodland and Derived Native Grassland (BGGW) ecological community. The ESP targets BGGW on farmland with private landholders paid under contract to undertake key management actions to improve the condition and extent

Box 2. Evaluation of ecological surrogates

A major challenge in determining the generality of an ecological surrogate, and its applicability in different situations, is finding the limitations of surrogate effectiveness. Whereas validation of a surrogate is critical for determining the ability of a surrogate to predict the outcome in a specific and narrowly defined situation (Box 1), evaluation of a surrogate is critical to understanding its generality across different situations. The enormous complexity and vast array of interacting variables in ecosystems makes establishing generality particularly difficult, and paradoxically is what often motivates the use of surrogates in the first place – to provide a simpler way to measure responses of diverse and complex systems.

Systematic evaluation of an ecological surrogate can be achieved through repeated tests of surrogacy relationships in contrasting ecosystems and over time. For example, Pierson et al. (2015) evaluated the performance of hollow-bearing trees to act as a surrogate for arboreal marsupials. This surrogacy relationship has both a strong mechanistic basis and has been previously validated (Lindenmayer et al. 2014). Pierson et al. (2015) evaluated this surrogacy relationship in contrasting woodland and forest ecosystems (i.e. different spatial situations), and over time (i.e. different temporal situations). Although they found that hollow-bearing trees were a good predictor of marsupial richness in different ecosystems (good spatial generality), they found no strong predictive relationship over time (limited temporal generality). This suggests that, despite a strong causal link between the surrogate (hollow-bearing trees) and outcome (marsupial occurrence), its reliability over time can be highly variable, and might therefore be of limited reliability as a monitoring tool (Pierson et al. 2015). Evaluation might also compare different surrogates to determine their performance in different situations. These data would also need to be weighed against the cost and logistics of collecting additional data on each surrogate variable, and how each surrogate might perform in isolation or in combination with others.

of woodland over time and, in turn, increase the abundance of particular groups of woodland-associated taxa over time (Zammit et al. 2010).

- In this context, specification of the surrogate model involves placement of the surrogate taxon in the pathway of effects between the treatment (land management) and the true outcome (woodland condition), as shown in Fig. 2. The key land management treatments in the ESP include 1) changed domestic livestock grazing on woodland patches, 2) reduced firewood collection, 3) control of invasive plants and animals, and 4) increase in vegetation cover either through promoting natural regeneration or by direct replanting. Surrogate variables used in the ESP are the presence, abundance and diversity of native birds and reptiles, and the structure and composition of vegetation (Lindenmayer et al. 2012). These entities are measured as a proxy for the true outcome of changes in woodland ‘condition’, which is based on an empirical understanding of the attributes of a high

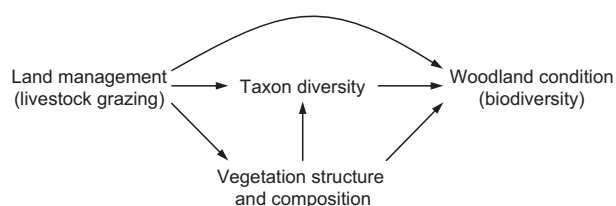


Figure 2. A schematic representation of the use of taxon diversity as a surrogate for woodland condition responses to land management in a large-scale agri-environment scheme (Lindenmayer et al. 2012). Following a general causal model (Box 1), reduced livestock grazing is applied to woodland remnants on farms, and faunal diversity responses are measured in lieu of broader biodiversity within each woodland site. Grazing can affect faunal diversity either directly or indirectly through altered vegetation structure and composition. Additional covariates can be added to the model if further sources of variation require examination.

quality woodland remnant in terms of structural complexity, native plant composition, and vertebrate fauna (Prober and Thiele 2005, Gibbons et al. 2008, Lindenmayer et al. 2012).

- Validation of the surrogate can be achieved with reference to the ‘general’ causal model shown in Box 1, and the systematic testing of associations between treatment and surrogate, treatment and outcome, and surrogate and outcome (Fig. 2). In our example, statistical examination of associations between treatments and surrogate variables is occurring through a major monitoring program on ESP sites. Paired control sites were established on each farm, where ESP-related management interventions were not undertaken: that is, ‘business-as-usual’ farm management practices were employed (e.g. traditional levels of livestock grazing). The paired control and stewardship sites were matched on the basis of vegetation type, land-form, patch size and patch connectivity (Lindenmayer et al. 2012). Ongoing studies are currently testing the effect of management treatments on different sets of taxa (associations between treatment and surrogate), as well as covariates such as vegetation structure and composition. Further tests, however, are required to determine levels of congruence in diversity between different subsets of taxa (associations between the surrogate and outcome), and how vegetation attributes might modify this cross-taxon congruence.
- Evaluation of the surrogate. Testing of the effectiveness of the surrogate variables in different circumstances, and identifying the limitations of each surrogate variable, is occurring through the examination of treatment-surrogate associations over time, and in different regions within the ESP. Monitoring of sites in the ESP has been underway for four years and initial results suggest that sites contracted under the ESP program have, over time, begun to support more tree regeneration, more native plant species, and less bare ground than control sites (Sato et al. unpubl., Lindenmayer et al. unpubl.). This

suggests that surrogate variables of vegetation 'condition' (i.e. the outcome) appear to be informative over time and across a wide geographic area. Further work is required to evaluate faunal surrogate taxa in the ESP.

Conclusions

We have drawn together ideas from the medical sciences to define an explicit surrogate concept that has not previously been used in ecology. This is relevant to several areas of applied ecology, including biodiversity monitoring, conservation management, and ecosystem restoration. Applied ecology has already benefited from ideas and concepts established in the medical sciences. This includes, for example, the shift towards evidence-based conservation (Pullin and Knight 2001, Fazey et al. 2004, Sutherland et al. 2004). The evaluation of different aspects of surrogacy might also be improved through rigorous synthesis and meta-analysis of the accumulated literature (Rodrigues and Brooks 2007, Mellin et al. 2011, Westgate et al. 2014). In the health sciences, this has been facilitated by groups such as the Cochrane Collaboration – a global network of health practitioners and researchers (The Cochrane Collaboration 2014). This also has been attempted in the environmental (<www.environmentalevidence.com>) and conservation (<www.conservationevidence.com>) sciences. Such networks allow for the building of evidence, and the publication of systematic reviews of the evidence. This is an important avenue for further advancement in surrogate ecology (Westgate et al. 2014). Our definition of a surrogate, and our new framework for specifying surrogate models, completing validation using a causal framework, and subsequent re-evaluation in different spatial and temporal contexts, is closely aligned with the approach already used by researchers in the clinical medical sciences. This rigorous approach can advance the science underpinning the application of surrogates in ecology by shifting the focus away from correlative understanding to one that focuses instead on causation and prediction. An improved use of surrogates is imperative if we are to meet the challenge of properly measuring and understanding the multifarious and complex problems in contemporary ecology.

Acknowledgements – Funding for this work was provided by an ARC Laureate Fellowship to DBL.

References

- Aronson, J. K. 2005. Biomarkers and surrogate endpoints. – *Br. J. Clinical Pharmacol.* 59: 491–494.
- Atkinson, A. J. et al. 2001. Biomarkers and surrogate endpoints: preferred definitions and conceptual framework. – *Clin. Pharmacol. Therapeutics* 69: 89–95.
- Barton, P. S. et al. 2014. Robustness of habitat-based surrogates of animal diversity: a multi-taxa comparison over time. – *J. Appl. Ecol.* 51: 1434–1443.
- Begg, C. B. and Leung, D. H. Y. 2000. On the use of surrogate end points in randomized trials. – *J. R. Stat. Soc. Ser. A* 163: 15–24.
- Bockstaller, C. and Girardin, P. 2003. How to validate environmental indicators. – *Agric. Systems* 76: 639–653.
- Boone, C. W. and Kelloff, G. J. 1993. Intraepithelial neoplasia, surrogate endpoint markers, and cancer chemoprevention. – *J. Cellular Biochem. Suppl.* 17F: 37–48.
- Branton, M. and Richardson, J. S. 2011. Assessing the value of the umbrella-species concept for conservation planning with meta-analysis. – *Conserv. Biol.* 25: 9–20.
- Brotman, B. and Prince, A. M. 1988. Gamma-glutamyltransferase as a potential surrogate marker for detection of the non-A, non-B carrier state. – *VoxSanguinis* 54: 144–147.
- Buyse, M. and Molenberghs, G. 1998. Criteria for the validation of surrogate endpoints in randomized experiments. – *Biometrics* 54: 1014–1029.
- Buyse, M. and Piedbois, P. 1996. On the relationship between response to treatment and survival time. – *Stat. Med.* 15: 2797–2812.
- Buyse, M. et al. 2010. Biomarkers and surrogate end points – the challenge of statistical validation. – *Nat. Rev. Clin. Oncol.* 7: 309–317.
- Caro, T. 2010. Conservation by proxy: indicator, umbrella, keystone, flagship and other surrogate species. – Island Press.
- Caro, T. M. and O'Doherty, G. 1999. On the use of surrogate species in conservation biology. – *Conserv. Biol.* 13: 805–814.
- Chown, S. L. and McGeoch, M. A. 2011. Measuring biodiversity in managed landscapes. – In: Magurran, A. and McGill, B. (eds), *Biological diversity: frontiers in its measurement and assessment*. Oxford Univ. Press, pp. 252–264.
- Colburn, W. A. 2000. Optimizing the use of biomarkers, surrogate endpoints, and clinical endpoints for more efficient drug development. – *J. Clin. Pharmacol.* 40: 1419–1427.
- Conti, M. E. and Cecchetti, G. 2001. Biological monitoring: lichens as bioindicators of air pollution assessment – a review. – *Environ. Pollut.* 114: 471–492.
- Drever, M. C. et al. 2008. Woodpeckers as reliable indicators of bird richness, forest health and harvest. – *Biol. Conserv.* 141: 624–634.
- Faith, D. P. and Walker, P. A. 1996. Environmental diversity: on the best-possible use of surrogate data for assessing the relative biodiversity of sets of areas. – *Biodivers. Conserv.* 5: 399–415.
- Fazey, I. et al. 2004. Can methods applied in medicine be used to summarize and disseminate conservation research? – *Environ. Conserv.* 31: 190–198.
- Feld, C. K. et al. 2009. Indicators of biodiversity and ecosystem services: a synthesis across ecosystems and spatial scales. – *Oikos* 118: 1862–1871.
- Fleishman, E. et al. 2005. Using indicator species to predict species richness of multiple taxonomic groups. – *Conserv. Biol.* 19: 1125–1137.
- Fleming, T. R. and Powers, J. H. 2012. Biomarkers and surrogate endpoints in clinical trials. – *Stat. Med.* 31: 2973–2984.
- Ghosh, D. 2012. A causal framework for surrogate endpoints with semi-competing risks data. – *Stat. Probabil. Lett.* 82: 1898–1902.
- Gibbons, P. et al. 2008. Rapidly quantifying reference conditions in modified landscapes. – *Biol. Conserv.* 141: 2483–2493.
- Gilbert, P. B. et al. 2008. Evaluating a surrogate endpoint at three levels, with application to vaccine development. – *Stat. Med.* 27: 4758–4778.
- Gollan, J. R. et al. 2009. Testing common habitat-based surrogates of invertebrate diversity in a semi-arid rangeland. – *Biodivers. Conserv.* 18: 1147–1159.
- Goodsell, P. J. et al. 2009. Evidence necessary for taxa to be reliable indicators of environmental conditions or impacts. – *Mar. Pollut. Bull.* 58: 323–331.
- Grace, J. B. et al. 2010. On the specification of structural equation models for ecological systems. – *Ecol. Monogr.* 80: 67–87.
- Joffe, M. M. and Greene, T. 2009. Related causal frameworks for surrogate outcomes. – *Biometrics* 65: 530–538.
- Kleijn, D. et al. 2006. Mixed biodiversity benefits of agri-environment schemes in five European countries. – *Ecol. Lett.* 9: 243–254.

- Li, Y. et al. 2011. Causal assessment of surrogacy in a meta-analysis of colorectal cancer trials. – *Biostatistics* 12: 478–492.
- Lindenmayer, D. B. and Likens, G. E. 2010. The science and application of ecological monitoring. – *Biol. Conserv.* 143: 1317–1328.
- Lindenmayer, D. B. and Likens, G. E. 2011. Direct measurement versus surrogate indicator species for evaluating environmental change and biodiversity loss. – *Ecosystems* 14: 47–59.
- Lindenmayer, D. B. et al. 2012. A novel and cost-effective monitoring approach for outcomes in an Australian biodiversity conservation incentive program. – *PLoS ONE* 7: e50872.
- Lindenmayer, D. B. et al. 2014. An empirical assessment and comparison of species-based and habitat-based surrogates: a case study of forest vertebrates and large old trees. – *PloS ONE* 9: e89807.
- Manning, A. D. et al. 2013. Bringing forward the benefits of coarse woody debris in ecosystem recovery under different levels of grazing and vegetation density. – *Biol. Conserv.* 157: 204–214.
- McGeoch, M. A. 1998. The selection, testing and application of terrestrial insects as bioindicators. – *Biol. Rev.* 73: 181–201.
- McGeoch, M. A. et al. 2002. The verification and application of bioindicators: a case study of dung beetles in a savanna ecosystem. – *J. Appl. Ecol.* 39: 661–672.
- Mellin, C. et al. 2011. Effectiveness of biological surrogates for predicting patterns of marine biodiversity: a global meta-analysis. – *PLoS ONE* 6: e20141.
- Molenberghs, G. et al. 2002. Statistical challenges in the evaluation of surrogate endpoints in randomized trials. – *Contr. Clin. Trials* 23: 607–625.
- Niemi, G. J. and McDonald, M. E. 2004. Application of ecological indicators. – *Annu. Rev. Ecol. Evol. Syst.* 35: 89–111.
- Niemi, G. et al. 2004. Rationale for a new generation of indicators for coastal waters. – *Environ. Health Perspect.* 112: 979–986.
- NIH Definitions Working Group 2000. Biomarkers and surrogate endpoints in clinical research: definitions and conceptual model. – In: Downing, G. J. (ed.), *Biomarkers and surrogate endpoints*. Elsevier, pp. 1–9.
- Noss, R. F. 1990. Indicators for monitoring biodiversity – a hierarchical approach. – *Conserv. Biol.* 4: 355–364.
- Perkins, A. J. et al. 2011. Adaptive management and targeting of agri-environment schemes does benefit biodiversity: a case study of the corn bunting *Emberiza calandra*. – *J. Appl. Ecol.* 48: 514–522.
- Pierson, J. C. et al. 2015. Can habitat surrogates predict the response of target species to landscape change? – *Biol. Conserv.* 184: 1–10.
- Prentice, R. L. 1989. Surrogate endpoints in clinical trials – definition and operational criteria. – *Stat. Med.* 8: 431–440.
- Prober, S. M. and Thiele, K. R. 2005. Restoring Australia's temperate grasslands and grassy woodlands: integrating function and diversity. – *Ecol. Manage. Restor.* 6: 16–26.
- Pullin, A. S. and Knight, T. M. 2001. Effectiveness in conservation practice: pointers from medicine and public health. – *Conserv. Biol.* 15: 50–54.
- Rodrigues, A. S. L. and Brooks, T. M. 2007. Shortcuts for biodiversity conservation planning: the effectiveness of surrogates. – *Annu. Rev. Ecol. Evol. Syst.* 38: 713–737.
- Shipley, B. 2009. Confirmatory path analysis in a generalized multilevel context. – *Ecology* 90: 363–368.
- Sinclair, R. G. et al. 2012. Criteria for selection of surrogates used to study the fate and control of pathogens in the environment. – *Appl. Environ. Microbiol.* 78: 1969–1977.
- Sutherland, W. J. et al. 2004. The need for evidence-based conservation. – *Trends Ecol. Evol.* 19: 305–308.
- Taylor, J. M. G. and Wang, Y. 2002. Surrogate markers and joint models for longitudinal and survival data. – *Contr. Clin. Trials* 23: 626–634.
- The Cochrane Collaboration 2014. The Cochrane Collaboration. Trusted evidence. Informed decisions. Better health. <www.cochrane.org/>. Accessed 3 March 2014.
- VanderWeele, T. J. 2013. Surrogate measures and consistent surrogates. – *Biometrics* 69: 561–581.
- Weir, C. J. and Walley, R. J. 2006. Statistical evaluation of biomarkers as surrogate endpoints: a literature review. – *Stat. Med.* 25: 183–203.
- Westgate, M. J. et al. 2013. Adaptive management of biological systems: a review. – *Biol. Conserv.* 158: 128–139.
- Westgate, M. J. et al. 2014. Global meta-analysis reveals low consistency of biodiversity congruence relationships. – *Nat. Comm.* doi: 10.1038/ncomms4899.
- Zammit, C. et al. 2010. Using markets for woodland conservation on private land: lessons from the policy-research interface. – In: Lindenmayer, D. B. et al. (eds), *Temperate woodland conservation and management*. CSIRO Publishing, pp. 297–307.