Control of *Pteridium aquilinum*: Meta-analysis of a Multi-site Study in the UK

GAVIN STEWART\textsuperscript{1,*}, EMMA COX\textsuperscript{2}, MIKE LE DUC\textsuperscript{2}, ROBIN PAKEMAN\textsuperscript{3}, ANDREW PULLIN\textsuperscript{1} and ROB MARRS\textsuperscript{2}

\textsuperscript{1}Centre for Evidence-Based Conservation, School of the Environment & Natural Resources, Bangor University, Bangor, Gwynedd, LL57 2UW, UK, \textsuperscript{2}Applied Vegetation Dynamics Laboratory, School of Biological Sciences, University of Liverpool, PO Box 147, Liverpool, L69 7ZB, UK and \textsuperscript{3}Macaulay Land Use Research Institute, Craigiebuckler, Aberdeen, AB15 8QH, UK

Received: 16 October 2007 Returned for revision: 3 January 2008 Accepted: 18 January 2008 Published electronically: 12 March 2008

- Background and Aims A great deal of money is spent controlling invasive weeds as part of international and national policies. It is essential that the funded treatments work across the region in which the policies operate. We argue that experiments across multiple sites are required to validate these programs as results from single sites may be misleading. Here, the control of *Pteridium aquilinum* (bracken) is used as a test example to address the following four questions. (1) Does the effectiveness of *P. aquilinum*-control treatments vary across sites? (2) Is the best treatment identified in previous research (cutting twice per year) consistent at all sites, and if not why not? (3) Is treatment performance related to *P. aquilinum* rhizome mass, litter cover or litter depth at the various sites? (4) Does successful *P. aquilinum* control influence species richness?
- Methods *Pteridium aquilinum*-control treatments were monitored for 10 years using six replicated experiments and analysed using meta-analysis. Meta-regressions were used to explore heterogeneity between sites.
- Key Results The effectiveness of treatments varied between sites depending on the measure used to assess *P. aquilinum* performance. In general, cutting twice per year was the most successful treatment but on some sites other, less expensive treatments were as good. The effectiveness of treatments at different sites was not related to rhizome mass, but the effectiveness of most applied treatments were inversely related to post-control litter. Effective treatment was also associated with high species richness.
- Conclusions It is concluded that successful development of national weed control programs requires multi-site experimental approaches. Here, meta-analyses demonstrate that variation in effectiveness between sites could be explained in part by pre-specified variables. Reliance on data from a single site for policy formulation is therefore clearly dangerous.

Key words: Bracken, *Pteridium aquilinum*, weed, control, herbicide, litter dynamics, species diversity, meta-analysis.

INTRODUCTION

A fundamental limitation in vegetation management and restoration ecology is the ability to predict, with some degree of precision, the likely outcome of a proposed treatment across a range of sites. This is essential to meet the objectives of both policy-makers and those responsible for implementing practical management schemes. However, it is well known that accurate prediction is difficult when treatments may have to be applied over wide environmental conditions, and where the restoration objectives may vary according to the ecosystem being treated. An obvious solution is to carry out multi-site experiments where the same treatments are applied at the same time; however, there are relatively few examples of such an approach (Pywell \textit{et al.}, 2002; Le Duc \textit{et al.}, 2003; Pakeman, 2004). This is most problematic when the treatments have to be delivered via an Integrated Land Management Approach (Milligan \textit{et al.}, 2004), where restoration includes the control of invasive species coupled with techniques to restore more appropriate vegetation, and is especially difficult where the invasive weed is problematic, and success is variable.

* For correspondence. E-mail g.stewart@bangor.ac.uk

The case-study problem

*Pteridium aquilinum* (bracken) is a serious invasive weed of upland and marginal land in many parts of the world (Pakeman and Marrs, 1992). In the UK, *P. aquilinum* is an especially problematic weed, causing problems in upland heaths and acid grasslands, where it often occurs in dense stands (summarized in Smith and Taylor, 1995). *Pteridium aquilinum* is difficult to control for a range of reasons, including an extensive rhizome system (Le Duc \textit{et al.}, 2003) and a high productivity that produces a dense frond cover and deep litter, which combine to reduce understorey vegetation (Marrs \textit{et al.}, 2000).

The current agricultural and conservation policy in the UK for *P. aquilinum* infestation is delivered through Agri-Environment Schemes (MAFF, 1993, 1996) and Biodiversity Action Plans (Anon, 1995a, b). Essentially, the policy is to reduce *P. aquilinum* infestations where possible and commonly to restore *Calluna vulgaris*-dominated heathland or acid grassland. *Pteridium aquilinum* is considered a mid-successional plant, usually occupying a niche between plago-climax heath/moor/grassland and woodland (Marrs \textit{et al.}, 2000, Marrs and Watt, 2006), although it can persist for long periods and appears in some cases to represent a highly resilient stable state.
In the UK, *P. aquilinum* occurs across a wide range of conditions and, given that succession can occur along many trajectories (Mitchell *et al.*, 2000; Suding *et al.*, 2004), there is potentially a large number of possible outcomes for *P. aquilinum* control/vegetation restoration schemes. In addition, *P. aquilinum* communities are known to have high resilience (Cox *et al.*, 2007, 2008; Stewart *et al.*, 2007) and *P. aquilinum* control often produces variable and conflicting results (Le Duc *et al.*, 2000; Cox *et al.*, 2007). It is, therefore, important from a national policy perspective to understand the response of *P. aquilinum* to control treatment in a variety of bio-geographical situations.

**Approach to this study**

There are a number of approaches that can be applied to analyse such multi-site studies, including univariate analysis of variance (Le Duc *et al.*, 2000; Pywell *et al.*, 2002; Pakeman, 2004; Cox *et al.*, 2007, 2008) and multivariate analysis (Cox, 2007). Recently there has been an increasing interest in the use of formal meta-analysis in ecological studies, specifically through the development of evidence-based frameworks and systematic review (Pullin and Knight, 2001, 2003; Sutherland *et al.*, 2004; Pullin and Stewart, 2006). Meta-analysis is a statistical approach to the analysis of a collection of results from individual studies or sites for the purpose of integrating the findings and exploring variation in effects (Hedges and Olkin, 1985; Arnqvist and Wooster, 1995; Gurevitch and Hedges, 1999, 2001; Osenberg *et al.*, 1999). The term has been used to describe any form of quantitative research synthesis, but also refers more specifically to a range of techniques that combine effect sizes using a weighting method. Meta-analytical approaches allow combination of heteroscedastic effect sizes and allow separation of within- and between-site variance, in contrast to ANOVA-type analyses, with a concurrent increase in power (Gurevitch and Hedges, 1999). Furthermore, the use of effect sizes and confidence intervals provide estimates of the precision and magnitude of an effect, in contrast to null-hypothesis significance tests (Shinichi and Cuthill, 2007).

Often, meta-analysis involves the comparison of experiments carried out by different researchers, in different situations and with heterogeneity in application methods, for example the review of asulam control of *P. aquilinum* (Stewart *et al.*, 2007). In such analyses it is often difficult to attribute variation in the success of treatments between studies to environmental variation because of confounding by the different methodologies employed by different researchers at different sites (Stewart *et al.*, 2005). However, data from a 10-year study of *P. aquilinum* control where six replicated experiments were carried out more-or-less simultaneously, with the same treatments being applied and results collected by the same team of observers (Le Duc *et al.*, 2000; Cox *et al.*, 2007) does not have these same problems. This dataset provides an opportunity to assess the value of meta-analytical approaches in ecological management and to assess the importance of inter-site variability with respect to treatment performance.

Here, meta-analytical approaches are used to address the following specific questions.

1. Does the effectiveness of *P. aquilinum*-control treatments vary across sites?
2. Is the best treatment identified in previous research (cutting twice per year, Marrs *et al.*, 1998; Cox *et al.*, 2007) consistent at all sites, and if not why not?
3. Is treatment performance related to *P. aquilinum* rhizome mass, litter cover or litter depth at the various sites?
4. Does successful *P. aquilinum* control influence species' richness?

On a more general note, we also assess the role of meta-analysis in analysing multi-site experiments in applied vegetation management.

**MATERIALS AND METHODS**

Six experiments were available for meta-analysis, as they had identical *Pteridium aquilinum* (L.) Kuhn control treatments applied; the experiments were set up at four different regional locations in the UK: Cannock Chase in the English Midlands; Hordron Edge in the North Peak Environmentally Sensitive Area, in the Peak District National Park, northern England; the Carneddau Estate, Snowdonia National Park, North Wales; and Sourhope in the Cheviot Hills, on the England/Scotland border. Sites are referred to hereafter as Cannock, Carneddau, Peak and Sourhope. Exact methodological details have been published elsewhere (Le Duc *et al.*, 2000, 2003) so only a brief summary is given here (see Table 1 for details).

**Experimental design and *P. aquilinum*-control treatments**

Use of an all-terrain vehicle to apply the treatments constrained experimental layout, and accordingly split-plot designs were used throughout. Each experiment had two or three replicated blocks with the six *P. aquilinum*-control treatments applied randomly at the main-plot level. The vegetation restoration treatments were applied randomly within the sub- and sub-sub-plot levels, and are not discussed further here. The experiment at Peak had a split-split-plot design from the start, and the one at Carneddau was changed to a split-split-plot design during its course to accommodate additional treatments (Table 1).

The six main-plot, *P. aquilinum*-control treatments were: (1) untreated (experimental control); (2) cut once per year, in June; (3) cut twice per year, in both June and August; (4) a single June cut in year one followed by herbicide (asulam) spraying in year two (‘cut and spray’); (5) asulam in year one only (‘spray’); (6) asulam in year one followed by a single June cut in year two (‘spray and cut’).
Table 1. Detailed description of experiments designed to test a range of \textit{P. aquilinum} control and vegetation restoration methods across the UK. Entries for experimental designs (all split- or split-split-plot randomized designs) represent: number of replicate blocks/number of plots per block/number of sub-plots per plot/number of sub-sub-plots per sub-plot. These codes are truncated from the right when lower experimental levels do not exist. All treatments are described in the text. Square brackets enclose treatment (or treatment start) date. All experimental treatments were applied in balanced designs with appropriate untreated contrasts, the untreated contrasts are not shown for clarity.

<table>
<thead>
<tr>
<th>Experiment Name</th>
<th>Sourhope Estate</th>
<th>North Peak ESA</th>
<th>Carneddau Estate</th>
<th>Cannock Chase</th>
</tr>
</thead>
<tbody>
<tr>
<td>Started</td>
<td>NT 861 202</td>
<td>NT 846 210</td>
<td>SK 213 870</td>
<td>SH 690 711</td>
</tr>
<tr>
<td>Design</td>
<td>2/6/2</td>
<td>2/6/2</td>
<td>3/6/2/3</td>
<td>3/6/3/2</td>
</tr>
<tr>
<td>Smallest plot size</td>
<td>10 × 18 m</td>
<td>10 × 18 m</td>
<td>10 × 5 m</td>
<td>10 × 5 m</td>
</tr>
<tr>
<td>Initial treatment</td>
<td>[Jul. 94]</td>
<td>[Jul. 93]</td>
<td>[Jul. 93]</td>
<td>[Jul. 93]</td>
</tr>
<tr>
<td>Sub-sub-treatments</td>
<td>–</td>
<td>–</td>
<td>Calluna seed (brush and litter) [Nov. 93]</td>
<td>–</td>
</tr>
<tr>
<td>Vegetation</td>
<td>U4e (U4e)</td>
<td>U4a (U4a)</td>
<td>U20 (H18)</td>
<td>U4a (U4a)</td>
</tr>
</tbody>
</table>

*National Vegetation Classification according to (Rodwell, 1991\textit{a}, \textit{b}, 1992; see Le Duc et al., 2000). NVC classes represented are: U2 = \textit{Deschampsia flexuosa} grassland; U2b = \textit{Vaccinium myrtillus} sub-community of \textit{Deschampsia flexuosa} grassland; U4a = typical sub-community of \textit{Festuca ovina}–\textit{Agrostis capillaris}–\textit{Galium saxatile} grassland; U4e = \textit{Vaccinium myrtillus}–\textit{Deschampsia flexuosa} sub-community of \textit{Festuca ovina}–\textit{Agrostis capillaris}–\textit{Galium saxatile} grassland; U20 = \textit{Pteridium aquilinum}–\textit{Galium saxatile} community; H18 = \textit{Vaccinium myrtillus}–\textit{Deschampsia flexuosa} heath; W16 = \textit{Quercus} spp. –\textit{Betula} spp. –\textit{Deschampsia flexuosa} woodland. Species nomenclature follows Stace (1997).
Combination treatments had only one application of each treatment, applied in years one and two. Generally, single cuts took place in June, second cuts and herbicide application in August.

Cutting was applied using a petrol-driven flail mower (Logic MFG series 300), trailed by a quad four-wheel-drive ATV (Kawasaki KLF 300) up to and including 1999, and subsequently using a hand-operated AEBI model HC55 flail mower. Asulam (Asulox, Bayer CropScience PLC) was sprayed in late August or early September at 4.4 kg active ingredient ha\(^{-1}\) (11 L Asulox ha\(^{-1}\)) in 400 L water ha\(^{-1}\) using a standard knapsack sprayer. In two experiments (Sourhope 1, Cannock 2; Table 1) plots treated with a single application of asulam were re-treated with asulam in 1996, 3 years after first treatment. On that occasion the herbicide was applied using a weed wiper (Rotowiper, Bisset Engineering International Ltd., New Zealand) trailed by the ATV. In June 1997 flooding prevented access to Sourhope 2, and only a single cut (August) was applied that year.

**Monitoring**

All experiments were monitored in June and August before the application of any *P. aquilinum*-control treatments. Quadrats (1 m\(^2\)) were placed at pre-selected random co-ordinates on 1-m grids within each sub- (sub-) plot. In June the cover of all species was estimated. In August a visual estimate of *P. aquilinum* cover was made in this 1 m\(^2\) quadrat and frond density and frond length were assessed in the central 0.25 m\(^2\); here, all fronds were cut at ground level, counted and length measured. Two quadrats were used in each sub-plot when the experiment was split to sub-plot level; when further divided, only one quadrant was used per sub-sub-plot.

**Data analysis and presentation**

Mean values for each plot in each block were obtained after appropriate transformation of raw data; frond density using \((Y + 0.5)^{0.5}\), frond length and cover using \(ln(Y + 1)\). The means were then back-transformed for meta-analysis. Change over time was averaged across years to avoid post hoc rationalization regarding the choice of temporal end point. Sub- and sub-sub-plot treatments were disregarded to prevent any confounding impacts. The comparative impacts of the applied treatments were analysed by meta-analysis and meta-regression (Cooper and Hedges, 1994; Deeks et al., 2001; Gurevitch and Hedges, 2001); these analyses were performed on three *P. aquilinum* variables: cover (%), mean frond density (number m\(^{-2}\)) and frond length (cm). Cohen’s \(D\) effect sizes (difference between population means divided by average population standard deviation; Deeks et al., 2001) were derived from the treatment and control means, standard deviations and sample sizes. Sensitivity analyses were used to verify the robustness of the effect size estimator in comparison to the Glass and Hedges’s effect size estimates (Cooper and Hedges, 1994) and there were no major differences between the effect size metrics. Response ratios were not used for this analysis as we were concerned with absolute rather than relative differences, particularly in relation to the meta-regressions. As the aim of this study was to determine the variation in effect between sites, a random effects model was used (Gurevitch and Hedges, 1999). This method uses the standardized mean difference method to express the size of the treatment effect in each study relative to the variability observed in that study (Deeks et al., 2001). Here, the DerSimonian and Laird method was used to perform this analysis (DerSimonian and Laird, 1986; Cooper and Hedges, 1994). The usual DerSimonian and Laird estimates of the effects sizes \((\theta_j)\) and mean and variance \((\tau^2)\) are given by Deeks et al. (2001) as:

\[
\tau^2 = Q - (k - 1)/\sum w_i - (\sum w_i^2/\sum w_i)
\]

where \(Q\) is the heterogeneity statistic with \(\tau^2\) set to zero if \(Q < k - 1\), and the weights \((w)\) are calculated by inverse variance, i.e.:

\[
w_i = 1/s.e.(\theta_i)^2 + \tau^2
\]

The pooled effect size for this DerSimonian and Laird method \((D)\) is given by:

\[
\theta_D = \sum w_i/\theta_i/\sum w_i
\]

with standard error:

\[
s.e.(\theta_D) = 1/\sqrt{\Sigma w_i}
\]

Questions 1 and 2 (treatment effectiveness across sites and assessing best treatment) were tested by inspection of Forrest plots of the estimated treatment effects from the studies along with their 95% confidence intervals, and by formal tests of homogeneity undertaken prior to each meta-analysis (Thompson and Sharp, 1999). Questions 3 and 4 (testing relationships between treatment performance and explanatory variables, including species richness) were tested by examining the associations of treatment effects with environmental variables that had been collected at the individual treatment plot level (*P. aquilinum* rhizome mass, litter cover, litter depth and species’ richness), previously postulated to affect both *P. aquilinum* control and restoration success (Marrs et al., 1998). These associations were examined using univariate random effects SMD meta-regression in Stata version 8.2 (Stata Corporation, USA, 2003) using the program Metareg (Sharp, 1998). The association of treatment effect derived from treatment–control comparisons \((n = 5)\) with explanatory variables \((n = 4)\) and outcome measures \((n = 3)\) required 60 regression analyses, necessitating the use of Bonferroni correction for each outcome measure, which were treated as independent metrics (Sankoh et al., 1997). The multiple comparison approach was preferred over the inclusion of treatment as a factor in order to retain independence, and to minimize the already high heterogeneity within the analyses. This allowed the exploration of heterogeneity by...
### Table 2. Pooled treatment effects based on DerSimonian and Laird SMD meta-analysis of the studies with 95% confidence intervals and tests of homogeneity based on a chi-square distribution of $Q$ (H$_2^2$ measure of heterogeneity). Significant results are indicated in bold ($P < 0.05$).

<table>
<thead>
<tr>
<th></th>
<th>Cut once per annum</th>
<th>Cut twice per annum</th>
<th>Cut in the first year, sprayed in the second</th>
<th>Sprayed once with herbicide</th>
<th>Sprayed in the first year, cut in the second</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>(A) Cover (%)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Untreated control</td>
<td>$D = 1.585 \ (0.855–2.315)$</td>
<td>$P &lt; 0.001$</td>
<td>$H^2 = 28.71$</td>
<td>$D = 1.592 \ (0.970–2.214)$</td>
<td>$P &lt; 0.001$</td>
</tr>
<tr>
<td>Cut once per annum</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cut twice per annum</td>
<td>$D = 2.180 \ (1.529–2.831)$</td>
<td>$P &lt; 0.001$</td>
<td>$H^2 = 19.01$</td>
<td>$D = 0.044 \ (0.286–0.373)$</td>
<td>$P = 0.001$</td>
</tr>
<tr>
<td><strong>(B) Frond length (cm)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Untreated control</td>
<td>$D = 2.581 \ (1.930–3.231)$</td>
<td>$P &lt; 0.001$</td>
<td>$H^2 = 20.39$</td>
<td>$D = 1.545 \ (0.799–2.291)$</td>
<td>$P &lt; 0.001$</td>
</tr>
<tr>
<td>Cut once per annum</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cut twice per annum</td>
<td>$D = 3.379 \ (2.836–3.921)$</td>
<td>$P &lt; 0.001$</td>
<td>$H^2 = 10.87$</td>
<td>$D = -0.514 \ (0.865 to –0.163)$</td>
<td>$P = 0.004$</td>
</tr>
<tr>
<td><strong>(C) Frond density (number m$^{-2}$)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Untreated control</td>
<td>$D = 0.347 \ (0.424–1.119)$</td>
<td>$P &lt; 0.001$</td>
<td>$H^2 = 48.47$</td>
<td>$D = 0.615 \ (0.323–1.554)$</td>
<td>$P &lt; 0.001$</td>
</tr>
<tr>
<td>Cut once per annum</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cut twice per annum</td>
<td>$D = 0.615 \ (0.323–1.554)$</td>
<td>$P &lt; 0.001$</td>
<td>$H^2 = 68.65$</td>
<td>$D = 0.045 \ (0.031–1.151)$</td>
<td>$P = 0.004$</td>
</tr>
<tr>
<td>Cut in the first year, sprayed in the second</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sprayed once with herbicide</td>
<td>$D = 0.321 \ (0.286–0.928)$</td>
<td>$P &lt; 0.001$</td>
<td>$H^2 = 9.39$</td>
<td>$D = -0.507 \ (0.972 to –0.168)$</td>
<td>$P = 0.005$</td>
</tr>
<tr>
<td>Sprayed once with herbicide</td>
<td>$D = 0.300 \ H^2 = 32.88$</td>
<td>$P &lt; 0.001$</td>
<td></td>
<td>$D = 0.128 \ (0.187–0.443)$</td>
<td>$P = 0.045$</td>
</tr>
</tbody>
</table>
Fig. 1. Forrest plots of individual site effect sizes from the six independent experiments (labelled 1–6) comparing five *P. aquilinum*-control treatments against untreated controls for three variables: *P. aquilinum* cover (%), mean frond length (cm), and mean frond density (number m⁻²). Treatment codes: C1 = cut once per year; SC = sprayed once and cut once; CS = cut once and sprayed with asulam once; S = sprayed once with asulam; and C2 = cut twice per year. Site codes: 1 = Cannock 1; 2 = Cannock 2; 3 = Carneddau; 4 = Peak; 5 = Sourhope 1; and 6 = Sourhope 2. Solid boxes represent the mean effect size of individual studies with 95% confidence intervals; the open diamond is the pooled effect size generated using standardized mean difference random-effects meta-analysis (the 95% CI is indicated by diamond width).
Frond length

**C1**

<table>
<thead>
<tr>
<th>Study</th>
<th>Standardized mean diff. (95% CI)</th>
<th>% weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2.06 (1.31, 2.82)</td>
<td>17.3</td>
</tr>
<tr>
<td>2</td>
<td>1.95 (0.85, 3.04)</td>
<td>17.9</td>
</tr>
<tr>
<td>3</td>
<td>3.11 (2.37, 3.85)</td>
<td>17.4</td>
</tr>
<tr>
<td>4</td>
<td>2.77 (2.06, 3.48)</td>
<td>17.7</td>
</tr>
<tr>
<td>5</td>
<td>2.21 (1.43, 2.98)</td>
<td>17.1</td>
</tr>
<tr>
<td>6</td>
<td>4.27 (3.06, 5.48)</td>
<td>12.7</td>
</tr>
<tr>
<td>Overall (95% CI)</td>
<td>2.58 (1.93, 3.23)</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Study</th>
<th>Standardized mean diff. (95% CI)</th>
<th>% weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.69 (0.98, 2.40)</td>
<td>16.5</td>
</tr>
<tr>
<td>2</td>
<td>-0.14 (-0.74, 0.47)</td>
<td>17.2</td>
</tr>
<tr>
<td>3</td>
<td>2.17 (1.54, 2.79)</td>
<td>17.1</td>
</tr>
<tr>
<td>4</td>
<td>1.70 (1.11, 2.29)</td>
<td>17.3</td>
</tr>
<tr>
<td>5</td>
<td>1.58 (0.88, 2.28)</td>
<td>16.6</td>
</tr>
<tr>
<td>6</td>
<td>2.37 (1.51, 3.23)</td>
<td>15.4</td>
</tr>
<tr>
<td>Overall (95% CI)</td>
<td>1.55 (0.80, 2.29)</td>
<td></td>
</tr>
</tbody>
</table>

**SC**

<table>
<thead>
<tr>
<th>Study</th>
<th>Standardized mean diff. (95% CI)</th>
<th>% weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.55 (-0.08, 1.18)</td>
<td>17.0</td>
</tr>
<tr>
<td>2</td>
<td>-0.43 (-1.06, 0.20)</td>
<td>17.0</td>
</tr>
<tr>
<td>3</td>
<td>2.54 (1.86, 3.23)</td>
<td>16.8</td>
</tr>
<tr>
<td>4</td>
<td>0.42 (-0.09, 0.94)</td>
<td>17.4</td>
</tr>
<tr>
<td>5</td>
<td>1.74 (1.00, 2.47)</td>
<td>16.7</td>
</tr>
<tr>
<td>6</td>
<td>3.83 (2.71, 4.95)</td>
<td>15.1</td>
</tr>
<tr>
<td>Overall (95% CI)</td>
<td>1.39 (0.32, 2.46)</td>
<td></td>
</tr>
</tbody>
</table>

**S**

<table>
<thead>
<tr>
<th>Study</th>
<th>Standardized mean diff. (95% CI)</th>
<th>% weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.25 (-0.36, 0.68)</td>
<td>17.0</td>
</tr>
<tr>
<td>2</td>
<td>0.32 (-0.31, 0.94)</td>
<td>17.0</td>
</tr>
<tr>
<td>3</td>
<td>2.98 (2.24, 3.72)</td>
<td>16.6</td>
</tr>
<tr>
<td>4</td>
<td>0.77 (0.24, 1.20)</td>
<td>17.4</td>
</tr>
<tr>
<td>5</td>
<td>2.80 (1.92, 3.66)</td>
<td>15.8</td>
</tr>
<tr>
<td>6</td>
<td>1.81 (1.03, 2.59)</td>
<td>16.3</td>
</tr>
<tr>
<td>Overall (95% CI)</td>
<td>1.46 (0.53, 2.39)</td>
<td></td>
</tr>
</tbody>
</table>

**C2**

<table>
<thead>
<tr>
<th>Study</th>
<th>Standardized mean diff. (95% CI)</th>
<th>% weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>4.11 (3.03, 5.20)</td>
<td>13.9</td>
</tr>
<tr>
<td>2</td>
<td>3.58 (2.59, 4.57)</td>
<td>15.3</td>
</tr>
<tr>
<td>3</td>
<td>4.13 (3.25, 5.02)</td>
<td>17.1</td>
</tr>
<tr>
<td>4</td>
<td>2.91 (2.20, 3.63)</td>
<td>20.3</td>
</tr>
<tr>
<td>5</td>
<td>2.48 (1.67, 3.30)</td>
<td>18.4</td>
</tr>
<tr>
<td>6</td>
<td>3.36 (2.36, 4.37)</td>
<td>15.1</td>
</tr>
<tr>
<td>Overall (95% CI)</td>
<td>3.38 (2.84, 3.92)</td>
<td></td>
</tr>
</tbody>
</table>

Fig. 1. Continued.
FIG. 1. Continued.

Frond density

C1

Study

Overall (95% CI)

Standardized mean diff.

% weight

C2

Study

Overall (95% CI)

Standardized mean diff.

% weight

SC

Study

Overall (95% CI)

Standardized mean diff.

% weight

S

Study

Overall (95% CI)

Standardized mean diff.

% weight

Stewart et al. — Control of Pteridium
considering factors in relation to each treatment comparison separately, the most parsimonious approach.

RESULTS

The use of meta-analysis is becoming increasingly prevalent in ecology (Hedges and Olkin, 1985; Arneqvist and Wooster, 1995; Gurevitch and Hedges, 1999, 2001; Osenberg et al., 1999) but not all ecologists are familiar with the output of such analyses. The results presented in Table 2 summarize the overall impact of different treatments across the six study sites, with the diagonal structure resulting from comparison of all permutations of five treatment comparisons. The pooled effect size \((D)\) is significant when the confidence intervals do not cross zero (Table 2, Figs 1 and 2). Effect sizes smaller than 0-2 are considered small by convention (Cohen, 1988), although there is no empirical basis for appraising ecological effect sizes for \(D\). Thus the majority of comparisons presented illustrate moderate-to-large statistically significant effects (Table 2). The statistical significance of variation between sites is tested using the \(Q\) statistic following a chi-square distribution under the null hypothesis that the true treatment effect is the same for all sites. The majority of the comparisons vary significantly between sites (see below). The Forrest plots (Figs 1, 2) illustrate the effectiveness of an applied treatment at each site, the overall pooled effect of treatment, and the variation in effect both within and between sites. Further information regarding interpretation of Forrest plots is available at (www.cebc.bangor.ac.uk, evidenced-based conservation, review guidelines).

**Question 1: does the effectiveness of \(P.\ aquilinum\)-control treatments vary across sites?**

There was significant variation\((P < 0.05)\) between sites for almost all treatments relative to their respective untreated control for all three \(P.\ aquilinum\) frond response...
variables (Table 2, Fig. 1); the two exceptions were (1) the cut twice per year treatment for frond length and (2) cut and spray for frond density. Cutting twice per year is the most effective treatment overall for reducing cover and height (biggest significant effect size, Table 2, Fig. 1), but spraying once is more effective in reducing frond density (Table 2, Fig. 1).

There was also significant variation ($P < 0.05$) between sites for other treatment comparisons but these varied with the outcome measurement. For cover (Table 2A) there was variation between sites when comparing the cut twice per year treatment versus spraying and cutting. For frond length, there was significant variation between sites for most treatment comparisons (Table 2B), the exceptions being cutting once or twice per year versus cutting and spraying. For frond density there were significant site effects for the untreated control versus all treatments except cut and spray (Table 2C). For the other frond density treatment comparisons the picture was variable, with the following being significant: (1) cutting once per year versus cut and spray and spraying, (2) cutting twice per year versus cut and spray, spray and cut versus spray and cut and spray.

**Question 2:** is the best treatment identified in previous research (cutting twice per year) consistent at all sites, and if not why not?

The effectiveness of cutting twice per year varies significantly ($P < 0.05$) between sites in comparison with uncut controls for cover and frond density, but not for frond length (Table 2, Fig. 2). The relative effectiveness of cutting twice per year does not generally vary significantly ($P < 0.05$) except in comparison with sprayed and cut

---

**Fig. 2.** Continued.
(cover and frond length), sprayed (frond length) and cut and sprayed (frond density; Table 2, Fig. 2). Thus, although the effectiveness of cutting varies across sites, it remains the most effective treatment in nearly all cases.

**Question 3:** Is treatment performance related to *P. aquilinum* rhizome mass, litter cover or litter depth at the various sites?

The effectiveness of *P. aquilinum* control was related to measured rhizome mass only for frond length in the cut-twice treatment (*P* < 0.05, Table 3A). This reduction in frond length associated with the cut-twice treatment produced an increase in effect size of 0.42 per unit increase in rhizome mass (Table 3A) but the result is not significant (at the alpha level of *P* < 0.0025) when the multiple comparisons are accounted for with Sidak’s or Bonferroni adjustment. The relationships of the control treatments with litter cover and depth were varied, with significant negative results found in seven and eight comparisons, respectively, although the coefficients were small (Table 3). Adjustment for multiple comparisons reduces the number of significant relationships to one and three for litter cover and depth, respectively (Table 3). Where a significant regression coefficient was found, the *P. aquilinum*-control treatment reduced the litter variable relative to the untreated experimental control.

**Question 4:** Does successful *P. aquilinum* control influence species richness?

The effectiveness of *P. aquilinum*-control treatment was significantly related to species’ richness (species’ richness ranged from 12–38 species per site) in 12 out of 15 comparisons (eight of 15 were significant following adjustment for multiple comparisons, Table 3D). Whilst the regression coefficients are low, these are statistically significant and
The results from the inter-site comparison of all treatments against the untreated controls showed three important results. First, there were significant reductions in at least one *P. aquilinum* performance measure relative to untreated controls for all applied treatments on most sites. Second, there were significant differences in response to untreated controls for all applied treatments on most sites. Third, different responses were found for some site × treatment comparisons depending on the *P. aquilinum* measure used (frond length, frond density, frond cover). Thus, in answer to our first question, the effectiveness of *P. aquilinum* control varies between sites, confirming the predictions of other workers (Pottier et al., 2005).

Comparisons of all treatments at all sites revealed that cutting twice within a year was usually the most effective treatment, but there were a few site × treatment combinations where other treatments were more effective at some sites depending on the measure used. Where *P. aquilinum* had been cut twice per year, the effective control obtained was at the expense of applying a total of 20 treatments over the 10-year period of the study. However, where other treatments were successful, fewer applications (one or two) were used, with obvious financial implications. Unfortunately, we cannot at this point predict where these treatments are likely to be superior, although the meta-regression gives some indications. Thus, the answer to our second question is that cutting twice within a year is generally the most effective treatment of those tested, but further research is required to identify the situations when this is not the case.

We also investigated the relationships between applied treatments and a range of other *P. aquilinum* variables (rhizome mass, litter cover, litter depth) that have been shown to be important in both *P. aquilinum* control and the subsequent regeneration on semi-natural communities (Marrs et al., 1998). The relationship with the rhizome mass was studied because ultimately the rhizomes are the main part of the *P. aquilinum*-control problems (Pakeman et al., 1994; Pakeman and Marrs, 1996; Pottier et al., 2005). Here, there was a positive relationship between rhizome mass and the effectiveness of the cut twice per year treatment, supporting both the hypotheses underlying questions 2 and 3. Litter variables were studied because they represent a barrier to the regeneration of new plant communities (Lowday and Marrs, 1992; Marrs et al., 2007). Our results support the hypothesis that most site × treatment combinations reduce the depth and cover experiments at different sites within the UK, which had all been run with common treatments and methodologies. The results allowed cross-site comparisons of applied treatments using a formal statistical approach. There are of course some drawbacks to this approach, one of which is the time frame over which the study effect size is derived. Here, we used a 10-year period, partly because there were 10 years of data available, but also because the results reflected a longer-term impact. As a result, shorter-term success of some treatments may have been underestimated, and this is likely in the treatments where asulam was applied because there is often a good initial effect followed by subsequent regrowth (Marrs et al., 1998).

The results presented here demonstrate the usefulness of the meta-analytical approach for investigating the significance of applied ecological restoration treatments in a multi-site study. Often, meta-analysis is used to compare treatment effects derived from different studies (Cooper and Hedges, 1994; Gurevitch and Hedges, 1999), with systematic reviews using meta-analysis to synthesize and summarize research findings to support evidence-based conservation (Stewart et al., 2005; Pullin and Stewart, 2006). Formal systematic reviews have recently been advocated in an attempt to avoid the biases associated with previous ecological meta-analyses (Leimu and Koricheva, 2004, 2005). Here, the approach was used to investigate hypotheses within a series of individual

**DISCUSSION**

The results from the inter-site comparison of all treatments against the untreated controls showed three important results. First, there were significant reductions in at least one *P. aquilinum* performance measure relative to untreated controls for all applied treatments on most sites. Second, there were significant differences in response between sites. Third, different responses were found for some site × treatment comparisons depending on the *P. aquilinum* measure used (frond length, frond density, frond cover). Thus, in answer to our first question, the effectiveness of *P. aquilinum* control varies between sites, confirming the predictions of other workers (Pottier et al., 2005).

Comparisons of all treatments at all sites revealed that cutting twice within a year was usually the most effective treatment, but there were a few site × treatment combinations where other treatments were more effective at some sites depending on the measure used. Where *P. aquilinum* had been cut twice per year, the effective control obtained was at the expense of applying a total of 20 treatments over the 10-year period of the study. However, where other treatments were successful, fewer applications (one or two) were used, with obvious financial implications. Unfortunately, we cannot at this point predict where these treatments are likely to be superior, although the meta-regression gives some indications. Thus, the answer to our second question is that cutting twice within a year is generally the most effective treatment of those tested, but further research is required to identify the situations when this is not the case.

We also investigated the relationships between applied treatments and a range of other *P. aquilinum* variables (rhizome mass, litter cover, litter depth) that have been shown to be important in both *P. aquilinum* control and the subsequent regeneration on semi-natural communities (Marrs et al., 1998). The relationship with the rhizome mass was studied because ultimately the rhizomes are the main part of the *P. aquilinum*-control problems (Pakeman et al., 1994; Pakeman and Marrs, 1996; Pottier et al., 2005). Here, there was a positive relationship between rhizome mass and the effectiveness of the cut twice per year treatment, supporting both the hypotheses underlying questions 2 and 3. Litter variables were studied because they represent a barrier to the regeneration of new plant communities (Lowday and Marrs, 1992; Marrs et al., 2007). Our results support the hypothesis that most site × treatment combinations reduce the depth and cover
of P. aquilinum litter, which indicated that the sites showed some reduction in litter and hence should be more amenable to colonization by other species.

We also showed that there were significant positive effects of P. aquilinum control on species’ richness, which answers question 4. Here, the regression coefficients are small, but they are highly significant. This is an important result because it indicates that one of the major outcomes of the P. aquilinum-control strategy is the development of a plant community with greater plant species’ diversity than under dense P. aquilinum cover. Details of the vegetation development of these experiments have been described elsewhere (Cox et al., 2007; Le Duc et al., 2007).

A major outcome of this study is the clear need for management experiments to be repeated in different places in order to develop evidence-based policy decisions, especially when such information is to be used to develop national guidelines and funding strategies (Sutherland et al., 2004; Anon, 2005; Pullin and Stewart, 2006). Too often, management conclusions are extrapolated from limited numbers of sites or single sites (e.g. Viggers and Hearn, 2005; Chamaillé-Jammes et al., 2007; Larsen and Guillemette, 2007). Our results here show that there is considerable site variation, possibly caused by differences in climatic regime, substrate, and past and current management practices, which all influence the starting position before management is applied and the desired end-point achieved (Marrs et al., 2000; Marrs and Watt, 2006).

There is also an obvious need to carry out further work to ascertain why P. aquilinum is so variable, and why in some places it is difficult to control whereas in others it is apparently less difficult (Pakeman, 2004). This is particularly important for P. aquilinum in view of its potential health and safety implications for livestock (Marrs and Watt, 2006), its predicted increase under future climate change scenarios, and its potential to increase in area and density as a result of land-use changes resulting from the reduced stocking rates encouraged in some agri-environment schemes (Pottier et al., 2005).

We re-emphasize the utility of meta-analyses in providing increased control of Type-II errors (i.e. the probability of accepting a wrong null hypothesis; Cohen, 1988), and in enabling assessment of the magnitude of an effect as well as its statistical significance (Glass et al., 1981; Rosenthal, 1984; Gurevitch and Hedges, 1999, 2001). In contrast to more traditional ANOVA-type analyses, meta-analysis allows the practical importance of an experimental effect to be assessed (Harris and Rosenthal, 1985). We suggest that the use of meta-analyses should not be restricted to research synthesis, but extended to the analysis of any applied multi-site study where inter-site variability is important.

ACKNOWLEDGEMENTS

We thank DEFRA and NERC of the UK for funding this work, and the land managers who permitted access to our experimental sites: Cannock, Staffordshire County Council (Sue Sheppard); Peaks (Jeremy Archdale and Neil Taylor); Carneddau (National Trust); and Sourhope (Macaulay Institute). We also thank Julia Koricheva, David Causton and an anonymous reviewer for their valuable and constructive, critical appraisal of this work.

LITERATURE CITED


