CENTRE FOR EVIDENCE-BASED CONSERVATION

Systematic Review No. 2

Does burning of UK sub-montane, dry dwarf-shrub heath maintain vegetation diversity?

Review Report

Compiled by
Stewart G.B., Coles C.F. & Pullin A.S.

Centre for Evidence-Based Conservation
School of Biosciences
The University of Birmingham
Edgbaston
Birmingham B15 2TT, UK

July 2004

This review should be cited as: Stewart, G.B., Coles, C.F., Pullin, A.S. (2004) Does burning of UK sub-montane, dry dwarf-shrub heath maintain vegetation diversity? Systematic Review No. 2. Centre for Evidence-Based Conservation.

SUMMARY

Background
Burning is a common vegetation management practice in upland UK, principally used to stimulate new growth of grasses or heather. Regular burning is used to maintain heather in the young, highly productive age-class and to create a mosaic of vegetation of different ages, composition and structure, particularly on grouse moors which tend to be burnt on an eight- to fifteen-year cycle. Burning is known to alter the vegetation composition, pattern, physical and age structure, nutrient status and carrying capacity for herbivores, as well as the associated fauna. However, from a conservation perspective, the effects of burning on species diversity in sub-montane dry dwarf shrub heath have not been systematically evaluated.

Objective
To assess the effectiveness of burning in conserving diversity of sub-montane dry dwarf-shrub vegetation communities.

Search Strategy
The following computerised English language databases were searched: English Nature's "Wildlink", JSTOR, ISI Web of Knowledge (comprising BIOSIS previews, CAB Abstracts, Derwent Innovations Index, INSPEC, ISI Current Contents, ISI Proceedings, ISI Web of Science), Index to Theses Online. Additionally, the reference lists of articles were searched and selected authors, recognised experts and current practitioners in the field of upland ecology were contacted for further references.

Selection Criteria
Primary, quantitative studies of burning on sub-montane dry dwarf-shrub heaths in Great Britain and Ireland (NVC types: H9, H10, H12, H16, H18, H21 & H22) with appropriate controls or other unburnt comparators. The outcome was any change in floristic composition.

Data collection and analysis
Inclusion decisions, quality assessment and data extraction were duplicated, and consensus achieved by discussion or a third party. Some authors were contacted for missing data. The primary measure of effect was the difference in mean species richness between the treatment and control groups. Simpson's index (1-D) provided a measure of evenness. Random effects meta-analyses weighted by variance, were performed where data could be pooled. Potential sources of heterogeneity were examined in meta-regression analyses.

Main Results
Seven articles were included. Five of these articles contributed data for meta-analysis (involving 13 datasets for species richness and 12 datasets for Simpson's index, four of which were not independent). The pooled weighted mean difference ranged from (95% CI, p): 1.16 (-2.90 to 5.23; p = 0.5755) to 0.30 (-2.58 to 1.98; p = 0.7959) for species richness and -0.04 (-0.12 to 0.04; p = 0.3005) to -0.05 (-0.10 to 0.005; p = 0.0771) for Simpson's index.

There was heterogeneity in all meta-analyses. Regression analyses found significant associations between effect size and stand age at time of burning for species richness and Simpson's index although species richness was not statistically significant in the multivariate regression model. Post burn time and data quality were not significantly associated with effect size. Funnel plots of the datasets and statistical tests for asymmetry indicated that there was no relationship between effect size and study precision.

Conclusions

The available evidence is mainly based upon short-term data sets. The only study designed to sample more than one burning rotation suggests the effects of serial burning on floristic diversity vary according to stand age at time of burning and the time elapsed since the last burn. There is some evidence that serial burning reduces diversity in old stands but insufficient evidence exists with which to make firm conclusions regarding the effectiveness of rotational burning in maintaining vegetation diversity of sub-montane dry dwarf-shrub heaths.

The effects of single burning cycles are also variable. There is evidence that burning old stands can reduce diversity, therefore land managers should beware of burning old stands where the maintenance of species diversity is the objective. This suggests that a juxtaposition of stands burnt prior to the degenerate stage and unburnt stands is required to maximise floristic diversity in a dry heath moorland mosaic at a landscape scale. More research is required to provide evidence concerning site-specific factors and burning-grazing-moisture interactions for the delivery of site specific management.

Does burning of UK sub-montane, dry dwarf-shrub heath maintain vegetation diversity?

1. Background

Rotational burning is a routinely-practiced form of traditional vegetation management in upland areas of the UK. It is mainly used to promote new grass or heather growth (Gimingham 1960; Gimingham 1972; Gimingham 1985; Hobbs 1981). Grouse moors tend to be burnt on an eight- to fifteen-year cycle to maintain heather in the young, highly productive age-class creating a heathland mosaic of different age, composition and structure (Gimingham 1960; Gimingham 1972; Gimingham 1985; Backshall *et al.* 2001). These heathlands are valued for their biodiversity and as extensive cultural landscapes (Webb 1998). They are recognized as an important habitat at a European level by the EU Habitats Directive and are subject to a UK biodiversity action plan. Although upland heathland is dependent on management to prevent succession to scrub or woodland, many upland heaths may suffer from overgrazing and environmentally damaging burning regimes (BAP Tranche 2; Volume (VI); Page 217).It is thought that poorly managed muirburn (ie large-scale and too frequent in operation) reduces the habitat quality of upland heath by causing a simplification of structure, loss of lower plant assemblages and erosion of peat (BAP Tranche 2; Volume (VI); Page 217). In a recent conservation status assessment, 24% of the area of upland Sites of Special Scientific Interest (SSSIs) was found to be in an unfavourable condition. This has been attributed to inappropriate moor burning (English Nature 2003).

Conversely, the Moorland Association (an organisation representing many sporting stakeholders whose members manage 750,000 of the estimated 800,000 acres of heather moor in England and Wales) believe that the use of moors for sporting interest is beneficial in ecological terms and that burning management can increase floristic diversity. “A skilfully burnt moor will have a mosaic of heather and other moor plants of differing ages and the rich variety of wildlife it attracts.” (Moorland Association 2004). There is therefore a clear need for evidence concerning the effects of burning on the conservation value of heathland to determine if burning management is compatible with conservation objectives. The loss of plant species is of particular concern therefore the review focused on the effect of burning on floristic diversity.

The evidence from experience and primary studies on the effects of burning and grazing on the composition and structure of moorland vegetation has been reviewed in numerous publications (Marrs and Welch 1991, Hester and Sydes 1992, Cadbury 1992, Thompson *et al.* 1994, 1995, Armstrong and Milne 1995, Shaw *et al.* 1996, Backshall *et al.* 2001, Tucker 2004). The impact of fire on post burn succession in dry heath is dependent largely on the age of *Calluna* as this determines whether regeneration is from seed or vegetative (Tucker 2004). Young *Calluna* in the pioneer stages can undergo rapid vegetative regeneration (Hobbs *et al.* 1984) whereas older stands regenerate more slowly (Miller and Miles 1970, Hobbs and Gimingham 1984, 1987 Berdowski and Siepel 1988). *Calluna* has a large seedbank which allows regeneration by seed when vegetative regeneration is limited however, recovery by seed germination is slow in comparison to vegetative regrowth and may allow colonisation, establishment and even long term dominance of species other than

Calluna. (Hobbs and Gimingham 1984). Despite the large research effort in UK upland plant ecology, there has been no formal, systematic evaluation of the evidence for the effectiveness of burning in achieving the objective of maintaining heathland vegetation diversity from a conservation perspective.

2. Objective of the review

To systematically collate and synthesise published and unpublished evidence in order to address the question "does burning of UK sub-montane, dry dwarf-shrub heath maintain vegetation diversity?"

Secondary objectives were:

- To examine whether the effect of burning on heathland diversity is influenced by stand age at the time of burning.

- To examine whether the effect of burning on heathland diversity is influenced by post burn time.

- To examine whether the effect of burning on heathland diversity is influenced by data quality.

- To explore other possible reasons for any heterogeneity in the effect of burning on heathland diversity.

3. Methods

3.1. Question formulation

The execution of a systematic review requires that a specific question is posed containing three key elements;

1. A subject – e.g. a definable population, species, or community type
2. An intervention – e.g. a management action
3. The desired outcome – e.g. the objective or target

To construct a suitable and appropriate question we met with representatives of the English Nature 'Uplands Group', part of the Terrestrial Wildlife Team. The resulting question, "does rotational burning of UK sub-montane, dry dwarf-shrub heath maintain vegetation diversity?" has these three elements as follows;

Subjects

Subject: sub-montane, dry dwarf-shrub heath in Great Britain and Ireland (NVC types: H9 *Calluna vulgaris – Deschampsia flexuosa* heath, H10 *Calluna vulgaris – Erica cinerea* heath, H12 *Calluna vulgaris – Vaccinium myrtillus* heath, H16 *Calluna vulgaris – Arctostaphylos uva-ursi* heath, H18 *Vaccinium myrtillus – Deschampsia flexuosa* heath, H21 *Calluna vulgaris – Vaccinium myrtillus – Sphagnum capillifolium* heath, H22 *Vaccinium myrtillus – Rubus chamaemorus* heath).

Intervention
Any form of burning was considered for inclusion. Studies examining burning over more than one rotation were noted, and where possible distinction was made between accidental burning and management or experimental burning. The control group is unburnt for >20 years resulting in the following comparison: burnt site compared with a site unburnt for >20 years that must have been subject to burning less recently than the burnt site where the burnt site is unburnt for >20 years. Studies without a valid comparator were excluded.

Outcome(s)
Articles with information on floristic composition were included. Diversity was quantified using species richness and Simpson's index (1-D) to provide outcome measures where: $D = N(N-1) / \sum n(n-1)$, N = total number of all species, n = total number of a given species. For 1-D, equitability takes a value between 0 and 1, with 0 being complete evenness.

3.2. Search Strategy
Studies, in the English language, were identified through computerised searches of English Nature's "Wildlink", JSTOR, ISI Web of Knowledge (comprising BIOSIS previews: 1969 to 2003, CAB Abstracts: 1973 to 2003, Derwent Innovations Index: 1963 to 2003, INSPEC: 1969 to 2003, ISI Current Contents: 1997 to 2003, ISI Proceedings: 1990 to 2003, ISI Web of Science: 1975 to 2003), Index to Theses Online (1970 to 2003). The search strategies used (Appendix one) combined synonyms for moor burning with target habitat and species terms.

To enhance the search we provided 52 recognised experts and current practitioners in the field of upland ecology with the provisional short list of articles and asked them to list any they thought we had missed. We also searched the bibliographies of articles accepted for full text viewing and contacted authors of articles that did not present data that were inferred to exist from methodological descriptions, or that presented data that could not easily be extracted.

3.3. Study inclusion criteria
Two reviewers assessed articles for broad relevance by screening of title and abstract. Three reviewers subsequently studied the full texts of selected articles for relevance and admitted them to the review by consensus. Inclusion criteria were derived from the review question. Thus, to be included in the review, articles had to report on primary studies conducted on the subject, describe the method of sampling and include measurements of floristic composition on sites that had been subjected to burning. Studies were excluded that had no control or unburnt comparators, or that reported on the effects of unique, accidental fires.

3.4. Study quality assessment
The methodological quality of the included studies was assessed by one reviewer and one third of the assessments were corroborated by a second reviewer. Quality assessment was undertaken using a fit for purpose study quality assessment instrument (Appendix 2). Primary assessment (with overriding weight) was given to study design using a hierarchy of evidence adapted from Stevens & Milne (1997)

and Pullin and Knight (2003). Studies that were considered flawed in their design were excluded from the review. Other criteria were used to further examine selection bias, performance bias, assessment bias and attrition bias, with higher weighting given to comparison of groups at baseline, assessment of replication, rigour of abundance measures and representativeness. The assessments of study quality are described in the table of included studies (Appendix 3), and were recorded for use in meta-regression analyses.

3.5. Data extraction

Data on species richness and Simpson's index from all included studies were extracted by one reviewer and one third of the studies were corroborated by a second reviewer. Data was extracted using a pilot tested data extraction form. Data presented only in graphs and figures were extracted where possible, but on occasion this proved impossible to undertake with accuracy. Attempts were made to contact selected authors through an open-ended request in order to obtain missing information or for clarification.

Additional information related to study methodology or quality that was extracted is shown in Table 1. Characteristics related to the populations that were extracted included: stand age at time of burning, community type & Initial floristic composition, location and altitude. Characteristics of the intervention that were extracted included: post burn time for treatment sites and equivalent monitoring times for control sites.

Species richness and Simpson's index were extracted from the articles on a replicate by replicate basis in order to derive a mean, standard deviation and sample size for treatment and control groups. These replicates were combined to form datasets. We developed a set of *a priori* rules to increase the repeatability of dataset formation whereby data were pooled at the smallest possible sample sizes, to ensure heterogeneity was minimised and to maximise the number of points available for meta-analysis, provided that independence was maintained. Pooling was by stand age at time of burning or post burn time where stand age was uniform or unknown. One article presented data on widely differing timescales that shared the same baseline. The data was extracted for each point in time but this data must be examined using sensitivity analysis and cannot be combined in meta-analysis due to problems with independence. The table of included studies (Appendix 3) provides a description of the outcome data reported from each dataset and the rationale behind its extraction.

3.6. Data analysis and synthesis

Handling of missing main outcome data.
Where sufficient main outcome data for meta-analysis was not presented (mean, sample size, standard deviation for treatment and control), available data was extracted for non-quantitative evaluation.

Handling of results of studies (main outcome) with more than one treatment arm.
In the studies with more than one relevant intervention group and a common control group, such as those comparing groups with different post burn times to an unburnt

control, raw results (the mean, sample size, standard deviations) from all relevant experimental groups were combined in order to obtain a measure of treatment effect. This enables the inclusion of all relevant data in the primary meta-analysis, although data quality is compromised slightly.

Choice of measure of effect and meta-analyses of main outcome.
The chosen measures of treatment effect were species richness and Simpson's index. These measures can both be treated as continuous data and analysed according to differences in mean treatment effects and their standard deviations. Random effects meta-analyses based on weighted mean difference were performed throughout using statsdirect. Sensitivity analyses were performed where appropriate.

Assessment of heterogeneity and investigation of reasons for heterogeneity.
Heterogeneity was assessed 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).

In addition to aspects of study quality, two potential sources of heterogeneity were specified *a priori* as secondary objectives of the review. We hypothesised that: (1) the effect of burning differs according to stand age at time of burning; and (2) the effect of burning differs according to post burn time. The association of these factors with estimated effects were examined by performing random effects meta-regression analyses in Stata version 8.0 (Stata Corporation, USA) using the program Metareg (Sharp 1998).

Investigation of publication and other biases.
Funnel plots (plots of effect estimates versus the inverse of their standard errors) were drawn. Asymmetry of the funnel plot may indicate publication bias and other biases related to sample size, though it may also represent a true relationship between trial size and effect size. A formal investigation of the degree of asymmetry was performed using the method proposed by Egger et al (1997).

4. Results

4.1 Review statistics
Of 24,484 hits (including replicates) in computerised databases, we judged 256 retrieved articles to have possibly relevant subjects, interventions and outcomes, based upon their titles only, or titles plus abstracts where the latter were available. This was inclusive of articles where there was insufficient information to make a decision without reference to the full texts. An additional 24 relevant articles were retrieved by searching bibliographies of articles that were obtained for full text viewing. One study found in this way was included in the final review.

Responses from experts and practitioners provided 14 potentially relevant articles (based upon assessment of titles only) that had not been found previously. None of these articles were accepted for the final review. One author provided raw data (Welch *pers. comm.*) to replace those in an article that had been selected for inclusion (Welch et al. 1994). Of 16 authors of articles that inferred the existence of useful data that were not presented, nine could not be contacted, four did not reply and three indicated that the data were unavailable. No new data were obtained as a result of contacting three authors of articles from which data could not readily be extracted.

Of the total of 279 articles retrieved for full text viewing, 76 were excluded because they presented only secondary evidence. The target biotype (the subject) was not represented in 36, 75 did not directly measure populations that had been subjected to a relevant intervention (burning of known timing) as part of a management regime and 162 did not present quantitative data on relevant measured outcomes. Thus, a number of studies were excluded on multiple criteria. Of 31 primary studies that presented data on the target biotype with a rotational burning intervention and relevant measured outcomes, 23 had no appropriate controls or comparators and were excluded. This included post-burn succession time series data, for which the lack of controls or other appropriate comparators meant that burning was only one of several factors that were potentially responsible for the observed patterns. The final review incorporated 13 datasets on species richness and 12 datasets on Simpson's index based on information from five articles; with a further two articles containing some information but lacking the variance data necessary for quantitative synthesis.

4.2 Study quality

Five articles are included in meta-analysis. The article conducted by Hobbs and Gimingham (1984) has been treated as five independent datasets, since the results for different stand ages at time of burning have been reported independently with their own comparators. Likewise, the article by Stevenson *et al*. (1996) has two independent datasets since stands with different post burn times have been reported on separately with their own comparators. Data from Welch *pers comm.* has been extracted in four datasets (from a timeseries) but they share the same baseline and must not be combined in meta-analysis for this reason. One article (Mallik and Gimingham 1983) presents data for species richness but not Simpson's index whilst the remainder report on both outcomes. Two additional studies (King 1960; Miller 1979) present outcome data but no measure of variance precluding their inclusion in meta-analyses. See characteristics of included studies tables (Appendix 3) for details of each study.

Design, methods and data quality.

The datasets included in formal meta-analysis were heterogenous with respect to design, methods and data quality. The most important dichotomy was between time series (Hobbs and Gimingham 1984; Mallik and Gimingham 1983) and site comparisons (all other studies). Baseline comparisons within site comparisons varied with respect to factors recorded but none of them presented baseline data indicating that the treatment and control sites were homogenous with respect to defined factors (Table 1) prior to burning. There was some heterogeneity within treatment and control arms in all the site comparisons although the time series of Hobbs and Gimingham (1984) (A-D) and Mallik and Gimingham (1983) had very low intra-treatment variation. No articles measured both the grazing and soil moisture co-interventions although Welch (*pers. comm.*) and Elliot (1953) recorded grazing and Stevenson *et al.* (1996) recorded soil moisture. Objective parameters of abundance were utilised by all the studies included in meta-analysis but differing parameters of abundance were used to examine floristic composition at a range of scales. There was also variation in the number and proportion of species monitored.

Populations.

The included datasets represent a range of heathland community types. Five datasets are H16, five (four non independent) are H12, one is H10 whilst the remaining two are intermediate between H9 / H12 and H10 / H16 respectively. Stand age at time of burning is variable ranging from seven to 36 years, although it is unknown for seven datasets. The altitude of the sites ranges from 180m to 600m. Twelve of the sites are Scottish (with a NE bias) with one from the Peak District in England.

Interventions.

The post burn time of datasets (treatment) varied from one year to 19 years but eight had a post burn time of <5 years whilst only two independent datasets had a post burn time of >10 years. Only one study (two data sets) (Stevenson *et al.* 1996) presented a comparison of serial burning with no or very limited burning that allows the effect of continued rotational burning to be ascertained.

4.3 Outcome of the review

Effect of burning on heathland diversity

The effect of burning on heathland diversity (as measured by species richness and Simpson's index) have been combined, where appropriate, to produce pooled estimates as described in the methods section.

Effect on species richness: The effect sizes and variance of the 13 data sets with species richness data are shown in a Forrest plot (Fig. 1). Pooling all 13 datasets was not valid because the Welch data sets are not independent. Two pooled estimates of effect were generated using the two most extreme measures of effect from the non independent data (Welch B and Welch D). For 10 datasets, including Welch B combined, the pooled estimate was 1.16 (95% confidence interval (CI), -2.90 to 5.23; $p = 0.5755$) indicating that there is no significant overall effect of burning on species richness. The CIs are wide and there is substantial heterogeneity in results which could be observed statistically ($Q = 117.99$ on nine degrees of freedom, $p < 0.0001$). A bias assessment funnel plot (Fig. 2) of the 13 data sets does not look asymmetrical, and the weighted regression test for asymmetry (Egger 1997) was not statistically significant (asymmetry intercept (95% CI) = -0.40 (-5.40 to 4.59) ($p = 0.86$)). There is, therefore, no evidence of bias using this method.

Sensitivity analysis was performed by substituting Welch D for Welch B and the results are very similar to those reported above. For 10 datasets, including Welch D combined, the pooled estimate was -0.30 (95% CI, -2.58 to 1.98; $p = 0.7959$). There is substantial heterogeneity in results ($Q = 23.80$ on nine degrees of freedom, $p < 0.0046$). The weighted regression test for asymmetry (Egger 1997) was not statistically significant (asymmetry intercept (95% CI) = -0.31 (-1.50 to 2.12) ($p = 0.70$)).

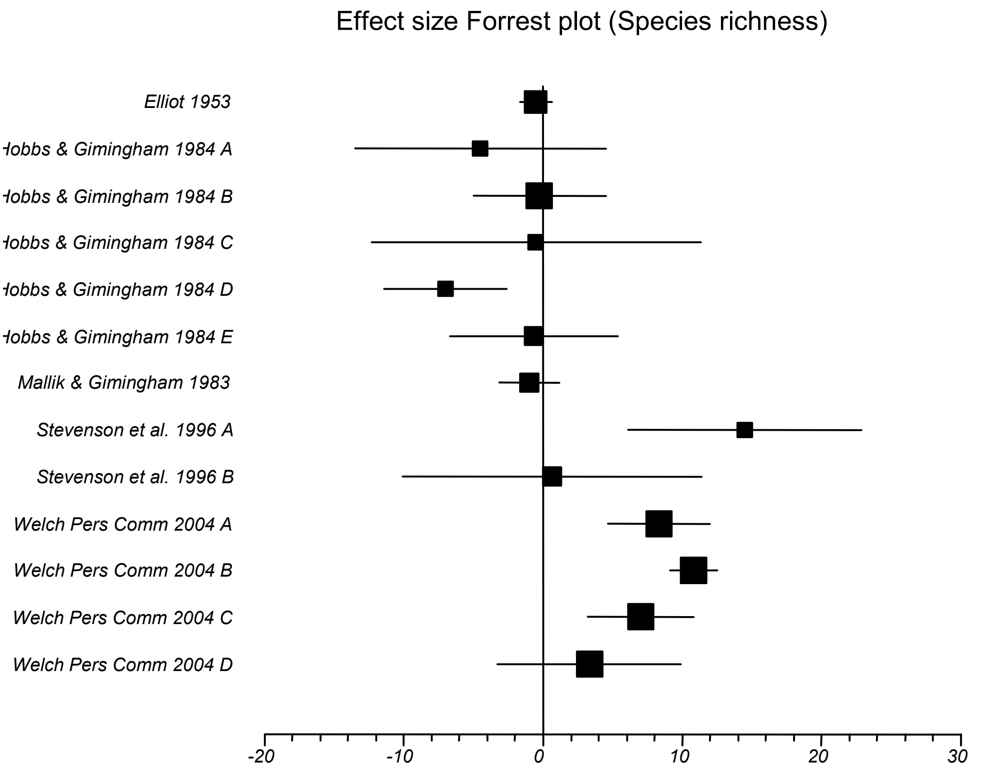

Figure 1. Forrest plot of effect sizes included in species richness meta-analyses.

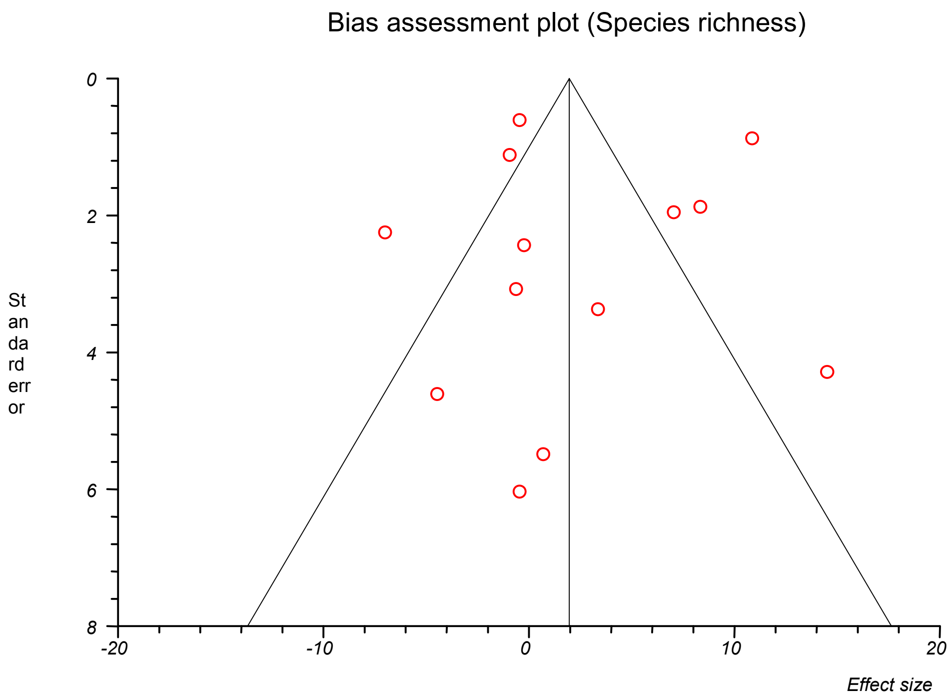

Figure 2. Funnel plot of the relationship between species richness effect size and study precision.

Effect on Simpson's index: The results of analysis on Simpson's index are very similar to those reported for species richness above. The effect sizes and variance of the 12 data sets with Simpson's index data are shown in a Forrest plot (Fig 3). As before, pooling all 12 datasets was not valid therefore two pooled estimates of effect were generated using the two most extreme measures of effect from the non independent data (Welch A and Welch D). For nine datasets including Welch A combined, the pooled estimate was -0.04 (95% CI, -0.12 to 0.04; p = 0.3005) indicating that there is no significant overall effect of burning on species diversity. The CIs are wide and there is substantial heterogeneity in results which could be observed statistically (Q = 275.46 on eight degrees of freedom, p < 0.0001). A bias assessment funnel plot (Fig. 4) of the 12 data sets does not look symmetrical. However, the weighted regression test for asymmetry (Egger 1997) was not statistically significant (asymmetry intercept (95% CI) = -2.72 (-8.13 to 2.70) (p = 0.27)). There is, therefore, no evidence of bias using this method.

As with species richness, sensitivity analysis was performed by substituting Welch D for Welch A and the results remained very similar. For nine datasets including Welch D combined, the pooled estimate was -0.05 (95% CI, -0.10 to 0.005; p = 0.0771). There is substantial heterogeneity in results (Q = 108.47 on eight degrees of freedom, p < 0.0001). The weighted regression test for asymmetry (Egger 1997) was not statistically significant (asymmetry intercept (95% CI) = -3.04 (-6.62 to 0.53) (p = 0.0841)).

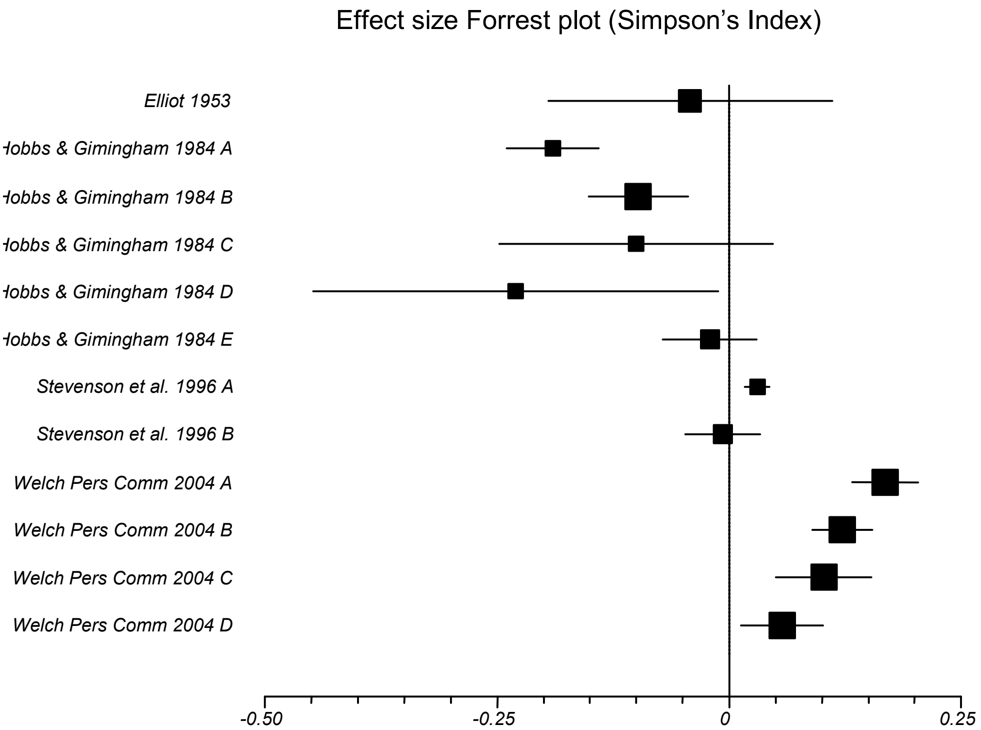

Figure 3. Forrest plot of effect sizes included in Simpson's index meta-analyses.

Bias assessment plot (Simpson's index)

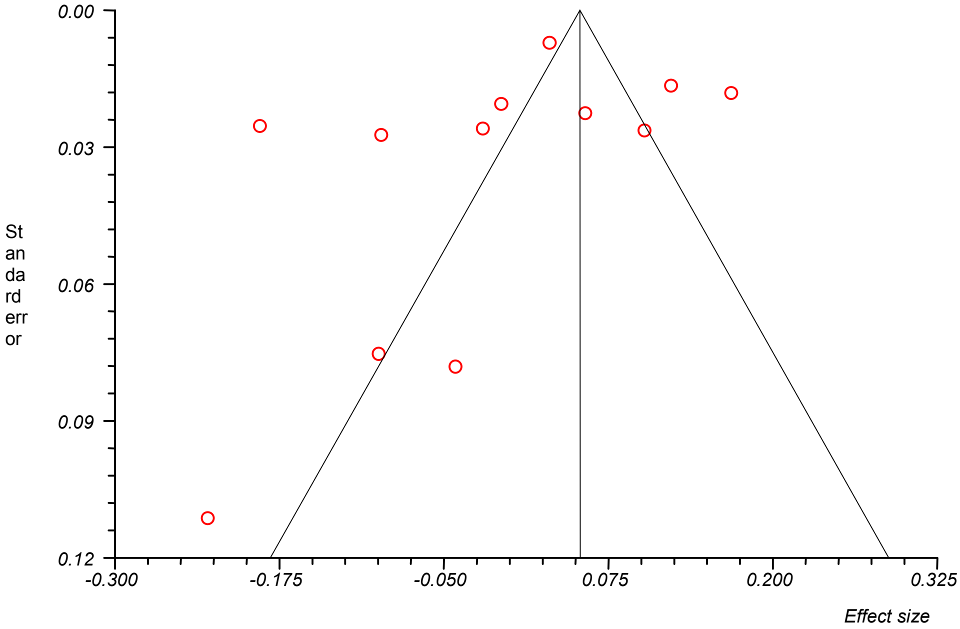

Figure 4. Funnel plot of the relationship between Simpson's index effect size and study precision.

Univariate and multivariate regression analysis on potential effect modifiers: Meta-regression was undertaken where data were available on all pre-specified factors (n = 6 for species richness, n = 5 for Simpson's index). Univariate meta-regression suggested no significant association between estimates of effect for species richness or Simpson's index and the pre-specified factors: post burn time and data quality. However, there was an association of stand age at time of burning with both species richness ($p < 0.05$) and Simpson's index ($p < 0.01$) (Table 1) suggesting that post burn diversity decreases as stand age before burning increases. Multivariate meta-regression controlling for all variables resulted in stand age becoming non significant for species richness, however, it remained significant in the case of Simpson's index. Post burn time was dropped from the Simpson's index meta-regression as it had no relationship to effect size once stand age and data quality were accounted for.

Table 1. The effect of potential effect modifiers examined using Meta-regression.

Analysis	Variable	Coefficient	Standard error	z	P	Lower CI	Upper CI
Species richness univariate	Stand age at time of burn	-1.60	0.74	-2.16	0.031	-3.06	-0.15
	Post burn time	-0.46	2.15	-0.21	0.831	-4.68	3.76
	Data quality	-0.93	1.97	-0.47	0.637	-4.80	2.94
Species richness multivariate	Stand age at time of burn	-1.80	0.97	-1.85	0.064	-3.70	0.11
	Post burn time	-1.14	2.06	-0.55	0.579	-5.17	2.89
	Data quality	0.49	2.88	0.17	0.865	-5.16	6.14
Simpson's Index univariate	Stand age at time of burn	-0.03	0.007	-4.60	0.000	-0.04	-0.02
	Post burn time	0.12	0.06	1.95	0.052	-0.0008	0.25
	Data quality	-0.06	0.03	-1.95	0.052	-0.12	0.0004
Simpson's index multivariate	Stand age at time of burn	-0.02	0.009	-2.53	0.011	-0.04	-0.005
	Post burn time	dropped					
	Data quality	-0.02	0.02	-1.20	0.229	-0.07	0.02

Sub group analysis was undertaken to examine all the data available on post burn time. (The meta regression data was constrained by availability of data on stand age at time of burning as this was often unreported). The pooled effect size was negative for sites with a post burn time of <5years and positive for sites with a post burn time of >5 yrs but neither were significant. Further sub group analysis was not undertaken on data quality due to confounding effects with post burn time (all low quality data has a short post burn time). Similar confounding interactions and the high risk of type 2 errors associated with multiple *post hoc* testing prevented further sub group analysis as a means of exploring possible explanations for the heterogeneity in the results.

5. Discussion

The main aim of this review was to determine the effectiveness of burning for conserving heathland dwarf-shrub diversity. We have attempted this by comparing sites burnt within the last 20 years to sites unburnt for at least the last 20 years. In total, 13 data sets (12 for Simpson's index) were included in formal meta-analysis. Two of these data sets compared serially burnt sites to sites unburnt for 70 to 190 years, whilst the remainder examined one burning cycle on moorland blocks managed by regular burning as far as can be ascertained. There was no significant effect of burning on heathland diversity when sites were pooled for either species richness or Simpson's index. The overall effect size is negative for Simpson's index and varies with sensitivity analysis from positive to negative for species richness. The pooled effects have large confidence intervals and the analyses exhibit significant heterogeneity meaning that the effects of burning on heathland diversity are variable. A degree of funnel plot asymmetry may be suggested by visual inspection of the Simpson's index Funnel plot, but the Egger test provided no evidence of a significant relationship between trial size and effect estimate indicating that none of the analyses were subject to bias. The variability in the effect of burning is unsurprising given the heterogeneity of the populations and intervention. It is consistent with other work that suggests that initial floristic composition, stand age at time of burning and post burn time have an important bearing on the outcome of burning. It is interesting to note that species richness and Simpson's index exhibit similar patterns. This indicates that where species richness is high, evenness is high and probably reflects the low diversity of heathland communities where the dominance of a single species often results in floristic poverty (Rodwell 1991).

A secondary aim of this review was to examine whether there was any relationship between the effect of burning on heathland diversity and stand age at time of burning, post burn time and data quality. There was a significant relationship between stand age at time of burning and the magnitude of the treatment effect whereby older stands experienced greater loss of diversity than younger stands. This relationship was significant in univariate regression models for both species richness and Simpson's index but not in the multivariate model for species richness. This result should be interpreted with caution for a number of reasons: There are a large number of covariates relative to the sample size. This could lead to spurious claims of association (Sterne, Egger and Smith 2001). In addition the association is observational and may be confounded by other unknown or unmeasured factors. The post burn recovery time is low for all the studies with stand age data so the longer term implications of burning older stands remain unknown. Again, this result is consistent with previous work. Stevenson *et al.* (1996) provides the best evidence that this relationship persists. Two serially burnt stands aged >19 years when last burnt have lower species richness than their unburnt counterparts. Unfortunately, the analysis above does not account for this as replicates were stratified on post burn time and full information on the study is unavailable pending its publication (see characteristics of included studies tables: Appendix 3).

We were unable to detect a clear relationship between post burn time, data quality and the magnitude of the treatment effect in the meta-regression analysis. However, the number of samples was limited by information on stand age. Subgroup analysis on all the post burn time data identified a trend whereby diversity declined in response to burning in the short term (<5 years) but increased in the longer term (>5yrs). However this was not significant. It is clear that more samples would be required to fully elucidate the relationship between effect size, stand age at time of burning and post burn time. Only two studies (Stevenson *et al* 1996*;* Welch *pers. comm*. 2004) provided data on the effects of burning beyond 10 years. This requires rectifying if the long term effects of burning on diversity are to be ascertained.

Initially, we hoped to investigate other potential sources of heterogeneity in this review, examining factors related to the populations, interventions and study quality. However, these potential sources of heterogeneity were not investigated due to the small sample sizes, lack of information, confounding of factors and the high risk of type 2 errors associated with numerous *post hoc* analyses. The effects of burning are often modified by grazing interactions. Predictions of succession under different management regimes often fail to separate the two variables (*e.g*. Rodwell 1991). In the current analysis no articles measured both the grazing and soil moisture co-interventions although Welch (*pers. comm.*) and Elliot (1953) recorded grazing and Stevenson *et al.* (1996) recorded soil moisture. It is clear that more work is required to inform management involving either factorial experiments or integrated monitoring if the complex interactions between these factors are to be understood.

The two articles providing non meta-analysable data (King 1960; Miller 1979) do not substantially alter the findings of the pooled studies. Miller is concerned with *Calluna* productivity rather than diversity. King (1960) provides further evidence that burning old stands (>20 years) results in loss of diversity. Unburnt stands have 11 more species than their burnt counterparts and Simpson's index is higher. The post burn time is 9 years providing another indication that the relationship between stand age and diversity may persist beyond the short term.

6. Reviewers' Conclusions

6.1 Implications for conservation practice

This review indicates that the effect of burning on heathland diversity is variable. We found evidence that older stands experienced greater loss of diversity than younger stands when subject to burning. These results should be interpreted with caution but there is some indication that land managers should beware of burning old stands where the maintenance of species diversity is the objective. This suggests that in order to maximise floristic diversity in a dry heath moorland mosaic at a landscape scale, some stands should be burnt prior to the mature growth phase, presumably on variable rotations, and other stands should remain unburnt.

The evidence base is insufficient to provide robust implications for practice regarding management of the multiple interactions between grazing, soil moisture and burning. Likewise, there is insufficient information to tailor management to account for site specific factors such as community type, altitude and location.

6.2 Implications for research

Of any wild landscape found in Europe, heather moorland is arguably the single most intensively researched (Thompson and Miles 1995). It is therefore surprising that only one study (two datasets) allowed examination of the effect of serial rotational burning on heathland diversity in the sub-montane zone of Britain. The study by Stevenson *et al.* (1996) identified and compared serially burnt stands to stands infrequently burnt and unburnt for the last 70-190 years using a combination of paleoecological techniques to determine fire history with contemporary vegetation analysis. Stevenson *et al* (1996) make their own recommendations concerning further work. It is clear that further studies of this nature are required in other geographic locations to provide an evidence base for evaluation of about the effective of serial burning on dwarf-shrub diversity to support decision making for upland management in Britain. Power analysis based on constancy data from all five sites indicates that a sample size of 12 sites would be required to have a 95% chance of detecting the effect with a significance of at least 5% (0.95 effect size for species richness). Thus, if possible, samples sizes should be increased to reduce the probability of type 1 errors.

Where a single burning rotation is considered, the quality of the datasets included in this review is variable with many lacking important methodological details. This is likely in part to be due to the constraints that the workers were operating within. Many characteristics considered crucial for excluding bias, such as clearly stated randomisation, large sample sizes, homogenous replicates and comparable baselines require long term experimental studies or monitoring within an integrated framework rather than short term studies or *ad hoc* monitoring. Given the belief that heather moorland has been intensively researched, funding for such work may not have been a priority. However, it is clear from this review that there is a requirement for needs-led research to fill this knowledge gap. In particular, head-to-head comparisons of the influence of potential effect modifiers, sources of heterogeneity and co-interventions such as community type, location, altitude, grazing and drainage are required as the samples sizes required to investigate these effects in meta-regression are far in excess of the actual numbers available. Pilot studies should be undertaken prior to the design of these experiments to provide effect size estimates to allow power analysis in order to determine sample size. These cannot be derived from the meta-analysis as there are insufficient numbers of head to head comparisons for meaningful analysis.

7. Acknowledgements

This project was supported by English Nature. Angus MacDonald (Scottish Natural Heritage) and David Welch (Centre for Ecology and Hydrology, Banchory) provided additional information about included studies; Arri Coomarasamy, Khalid Khan (Birmingham Women's Hospital) and Rod Taylor (University of Birmingham) provided statistical advice and support; Mick Rebane (English Nature) commented on an earlier version of this review. We thank all the experts and practitioners that responded to our enquiries.

8. References

Armstrong, H.M. and Milne, J.A. (1995) The effects of grazing on vegetation species composition in Thompson, D.B.A., Hester, A.J. and Usher, M.B., (eds). *Heaths and moorland: cultural landscapes*. HMSO, Edinburgh.

Backshall, J, Manley, J and Rebane, M (eds.) 2001. *The Upland Management Handbook*. English Nature, Peterborough.

BAP Tranche 2; Volume (VI); Page 217. http://www.ukbap.org.uk.

Berdowski, J.J.M. and Siepel, H. (1988) Vegetative regeneration of *Calluna vulgaris* at different ages and fertilizer levels. *Biological Conservation* **46**, 85-94.

Cadbury, C.J. (1992) Grazing and other management of upland vegetation: a review with special reference to birds. Royal Society for the Protection of Birds, Sandy.

Egger M, Davey-Smith G, Schneider M, Minder C. (1997) Bias in meta-analysis detected by a simple graphical test. *British Medical Journal* **315**, 629-34.

Elliott, R. J. (1953) *The effects of burning on heather moors of the south Pennines. PhD Thesis*, University of Sheffield.

English Nature (2003) *England's best wildlife and geological sites. The condition of Sites of Special Scientific Interest in England in 2003*. English Nature External Relations Publication. English Nature, Peterborough.

Gimingham, C. H. (1985). *Muirburn*. In: *Vegetation management in northern Britain, proceedings of seminar, Peebles, 1985*. (ed. R.B. Murray) British Crop Protection Council Monograph no. 30. Pp. 71-75. BCPC, Croydon.

Gimingham, C.H. (1972) *Ecology of Heathlands*. Chapman and Hall, London.
Green, R.E. (2002). Diagnosing causes of population declines and selecting remedial actions. in: Norris, K. and Pain, D.J. (eds). *Conserving Bird Biodiversity. General principals and their application.* Cambridge University Press, Cambridge.

Gimingham, C.H. (1960) Calluna. *The Journal of Ecology,* **48**, 455-483.

Hester, A. and Sydes, C. (1992) Changes in the burning of Scottish heather moorland since the 1940s from aerial photographs. *Biological Conservation* **60**, 25-30.

Hobbs, R. J. and Gimingham, C. H. (1984) Studies on Fire in Scottish Heathland Communities II. Post-Fire Vegetation Development. *Journal of Ecology*, **72**, 585-610.

Hobbs, R.J. (1981) *Post-fire succession in heathland communities*. Ph.D. thesis, University of Aberdeen.

Hobbs, R.J. and Gimingham, C.H. (1987) Vegetation, fire and herbivore interactions in heathland. *Advances in Ecological Research* **16**, 87-173.

King, J. (1960) Observations on the Seedling Establishment and Growth of Nardus stricta in Burned Callunetum. *Journal of Ecology*, **48**, 667-677.

Mallik, A. U. and Gimingham, C. H. (1983) Regeneration of Heathland Plants Following Burning. *Vegetatio*, **53**, 45-58.

Marrs, R.H. and Welch, D. (1991) *Moorland wilderness: the potential effects of a reduced grazing pressure.* Report to the Department of the Environment, London.

Miller, G. R. (1979). Quantity and Quality of the Annual Production of Shoots and Flowers by Calluna Vulgaris in North-East Scotland. *Journal of Ecology*, **67**, 109-129.

Miller, G. R. and Miles, J. (1970). Regeneration of Heather (Calluna-Vulgaris) at Different Ages and Seasons in Northeast Scotland. *Journal of Applied Ecology* **7**(1), 51-60.

Moorland Association. (2004) http://www.moorlandassociation.org.

Pullin, A.S. and Knight, T.M. (2003). Support for decision making in conservation practice: an evidence-based approach. *Journal for Nature Conservation*, **11**, 83-90.

Rodwell, J.S. (1991) *British Plant Communities volume 2: Mires and heaths*. Cambridge University Press, Cambridge.

Sharp, S. (1998) Meta-analysis regression: statistics, biostatistics, and epidemiology. *Stata Technical Bulletin*.**42**, 16-22.

Shaw, S.C., Wheeler, B.D., Kirby, P., Philipson, P. and Edmonds, R. (1996). Literature review of the historical effects of burning and grazing of blanket bog and upland wet heath. *English Nature Research Report* **172**. English Nature, Peterborough.

Sterne, J.A.C., Egger, M. and Smith, G.D. (2001) Investigating and dealing with publication and other biases. In Egger, M., Smith, G.D. and Altman, D.G. (eds). *Systematic Reviews in health care: meta-analysis in context*. BMJ publishing group, London.

Stevens, A. and Milne, R. (1997) *The effectiveness revolution and public health*. In: *Progress in Public Health* (ed. G. Scally) pp.197-225. Royal Society of Medicine Press, London.

Stevenson, A. C., Rhodes, A.N., Kirkpatrick, A.H. & MacDonald, A.J. (1996) *The determination of fire histories and an assessment of their effects on moorland soils and vegetation*. Scottish Natural Heritage Research, Survey and Monitoring Report no.16, SNH, Edinburgh.

Thompson, D. B. A., MacDonald, A.J., Marsden, J.H. and Galbraith, C.A. (1995). Upland heather moorland in Great Britain: A review of international importance, vegetation change and some objectives for nature conservation. *Biological Conservation* **71**(2), 163-178.

Thompson, D.B.A. and Miles, J. (1995) Heaths and Moorland: Some conclusions and questions about environmental change. In Thompson, D.B.A., Hester, A.J. and Usher, M.B., (eds). *Heaths and moorland: cultural landscapes*. HMSO, Edinburgh.

Thompson, D.B.A., MacDonald, A.J. and Hudson, P.J. (1994) Upland Moors and heaths. In Sutherland, W.J. and Hill, D.A. (eds). *Habitat management*. Cambridge University Press, Cambridge.

Thompson, S,G, and Sharp, S.J. (1999). Explaining heterogeneity in meta-analysis: a comparison of methods. *Statistics in Medicine*, **18**(20), 2693-708.

Tucker, G. (2004) Review of the impacts of heather and grassland burning in the uplands on soils, hydrology and biodiversity. Report to English Nature.

Webb, N.R. (1998) The traditional management of European heathlands. *Journal of Applied Ecology*. **35**, 987-990.

Welch, D. Scott, D. Moss, R. and Bayfield, N.G. 1994. Ecology of blaeberry and its management in British moorlands. Institute of Terrestrial Ecology**:** 39.

Appendix 1. Full list of terms used in electronic searches

We searched each information source using a Boolean-type strategy employing both specific and general terms (e.g. burn* AND Calluna; * denotes a wildcard for use with truncated terms). Search terms were not combined using the Boolean operator 'OR', because we found that in the Web of Science database, such strategies produced consistently fewer hits than the sum of separate searches using only the 'AND' operator. These strategies were used for every source to ensure strict comparability of results.

1. burn* AND calluna
2. burn* AND dwarf AND shrub*
3. burn* AND empetrum
4. burn* AND erica
5. burn* AND heath*
6. burn* AND upland*
7. burn* AND vaccinium
8. fire* AND calluna
9. fire* AND dwarf AND shrub*
10. fire* AND heath*
11. fire* AND upland*
12. moor* AND burn*
13. moor* AND fire*
14. muirburn*
15. burn* AND bog*
16. burn* AND mire*
17. burn* AND peat
18. burn* AND Eriophorum
19. burn* AND Scirpus
20. burn* AND Trichophorum
21. burn* AND Molinia
22. burn* AND Sphagnum
23. fire* AND bog*
24. fire* AND mire*
25. fire* AND peat
26. fire* AND Eriophorum
27. fire* AND Scirpus
28. fire* AND Trichophorum
29. fire* AND Molinia
30. fire* AND Sphagnum.

Appendix 2. Quality Assessment Instrument to provide an estimate of bias surrounding extracted data.

Bias	Generic data quality features	Specific data quality features	Quality element	Quality score
Selection and Performance bias	Study Design	NA	Randomized controlled Trial	**80**
			Controlled Trial	**60**
			Site comparison	**40**
			Time series	**20**
	Baseline comparison (heterogeneity between treatment and control arms with respect to defined confounding factors before treatment)	Factors: Diversity (Species richness, Simpson's index). Community type & Initial floristic composition, Location, Altitude. Stand age at time of burning	Treatment and control arms homogenous	**6**
			Some heterogeneity between treatment and control arms	**3**
			Treatment and control arms not comparable with respect to confounding factors OR insufficient information	**0**
	Intra treatment variation	Factors: Stand age at time of burn, Post burn time Community type, Location, Altitude.	No heterogeneity within treatment and control arms	**4**
			Some heterogeneity within treatment and control arms	**2**
			Replicates within treatment and control arms not comparable	**0**
	Measurement of intervention and Co-interventions	Factors: Monitoring times, Grazing, Drainage.	All Factors other than intervention equal in treatment and control arms	**2**
			Some heterogeneity between treatment and control arms	**1**
			Factors not equal or unreported	**0**
Assessment bias	Measurement of outcome	Replication, parameter of abundance	Well replicated objective parameter of abundance used	**4**
			Replicated objective parameter of abundance used	**2**
			Unreplicated observations or subjective parameter of abundance used	**0**
		Number / proportion of species monitored	All species monitored in both treatment and control	**2**
			Most species monitored in treatment and control	**1**
			Limited number of species monitored	**0**
Attrition bias	Assessment of treatment effect on sample number	NA	No losses to follow up	**2**
			Minor (<20%) losses to follow up	**1**
			Major (>20%) losses to follow up	**0**

Appendix 3. Characteristics of included studies

<table>
<tr><th>Study</th><th colspan="3">Study Elliot 1953</th></tr>
<tr><td>Methods</td><td colspan="3">Site comparison based on seven measurements from four stands with frequency measures based on 50 random quadrats in each. The five burnt measures are combined as a treatment arm and compared to the two unburnt (for 20 years) measures as a control arm.</td></tr>
<tr><td>Population</td><td colspan="3">Stand age at time of burning: unknown.
Community type / Initial floristic composition: H9 / H12. A difficult separation especially with only seven species. Calluna vulgaris and Deschampsia flexuosa are dominant with associated ericoids and limited Eriophorum angustifolium.
Location: Toadsmouth moor, Peak District.
Altitude: 325m.</td></tr>
<tr><td>Intervention and Co interventions</td><td colspan="3">Post burn time (treatment): 5 years (mean).
Equivalent monitoring time (control): monitored at same time as treatment stands.</td></tr>
<tr><td rowspan="3">Outcomes</td><td></td><td>Simpson’s index</td><td>Species richness</td></tr>
<tr><td>Treatment</td><td>0.388</td><td>3</td></tr>
<tr><td>Control</td><td>0.43</td><td>3.5</td></tr>
<tr><td>Study design</td><td colspan="3">Site Comparison: 40</td></tr>
<tr><td>Baseline Comparison</td><td colspan="3">The location, altitude and presumably community type were the same for all stands but no information is presented concerning diversity, initial floristic composition or stand age at time of burning: 3</td></tr>
<tr><td>Intra treatment variation</td><td colspan="3">Community type, location and altitude are similar but post burn time is variable in the treatment arm and stand age at time of burning is unreported: 2</td></tr>
<tr><td>Measurement of intervention and Cointerventions</td><td colspan="3">Light sheep grazing (presumed equal) and monitoring times are equal in treatment and control arms (no information about soil moisture variation): 1</td></tr>
<tr><td>Replication & parameter of abundance</td><td colspan="3">Replicated objective parameter of abundance used (x and n presented): 2</td></tr>
<tr><td>Number / Proportion of species monitored</td><td colspan="3">7 species monitored. It is assumed that this represents most but not all species on species poor sites: 1</td></tr>
<tr><td>Attrition bias</td><td colspan="3">No losses to follow up: 2</td></tr>
<tr><td>Sum of Data quality</td><td colspan="3">51</td></tr>
<tr><td>Notes</td><td colspan="3">Data based on table 2. Three treatment sites (five measurements) burnt 2-9 years before monitoring. One control sites (two measurements) burnt 20 years before monitoring. The measurements from the same sites were taken in successive seasons and are therefore not fully independent but they provide a useable and germane measure of variability within the site.</td></tr>
</table>

<table>
<tr><th>Study</th><th>Hobbs and Gimingham 1984</th></tr>
<tr><td>Methods</td><td>Time series based on cover abundance in 128 subplots from 13 stands with before and after treatment data. There are five datasets based on combining two pioneer stands (A), four building stands (B). two mature stands (C), two degenerate stands (D) and three species poor stands (E). Before burning is the control arm and post burning is the treatment arm.</td></tr>
<tr><td>Population</td><td>Stand age at time of burning: A 7.1, B 13.9, C 21, D 35.8, E 10.6. (mean years).
Community type / Initial floristic composition: A-D H16, E H12.
Location: A-D Muir of Dinnet, NE Scotland. E Glen Dye, NE Scotland.
Altitude: A-D 180m, E 400m.</td></tr>
<tr><td>Intervention and Co interventions</td><td>Post burn time: A-D 2 years, E 3 years.
Equivalent monitoring time (control): NA</td></tr>
<tr><td>Outcomes</td><td><table>
<tr><td></td><td colspan="5">Simpson’s index</td><td colspan="5">Species richness</td></tr>
<tr><td>Dataset</td><td>A</td><td>B</td><td>C</td><td>D</td><td>E</td><td>A</td><td>B</td><td>C</td><td>D</td><td>E</td></tr>
<tr><td>Treatment</td><td>0.705</td><td>0.715</td><td>0.725</td><td>0.59</td><td>0.716</td><td>27.5</td><td>21</td><td>15.5</td><td>9</td><td>12.33</td></tr>
<tr><td>Control</td><td>0.895</td><td>0.813</td><td>0.825</td><td>0.82</td><td>0.737</td><td>32</td><td>21.25</td><td>16</td><td>16</td><td>13</td></tr>
</table></td></tr>
<tr><td>Study design</td><td>Time series: 20</td></tr>
<tr><td>Baseline Comparison</td><td>Treatment and control arms homogenous (time series): 6</td></tr>
<tr><td>Intra treatment variation</td><td>Community type, location and altitude, post burn time and stand age at time of burning very similar for replicates in datasets A-D: 4. Dataset E has variable stand age at time of burning: E 2.</td></tr>
<tr><td>Measurement of intervention and Cointerventions</td><td>No information about grazing or soil moisture: 0</td></tr>
<tr><td>Replication & parameter of abundance</td><td>Well Replicated objective parameter of abundance used (x and n presented): 4</td></tr>
<tr><td>Number / Proportion of species monitored</td><td>Total species number presented for species richness: 2
Limited number of species reported on / interpretable after burning for Simpson’s index: 0</td></tr>
<tr><td>Attrition bias</td><td>No losses to follow up: 2</td></tr>
<tr><td>Sum of Data quality</td><td><table>
<tr><td></td><td>A</td><td>B</td><td>C</td><td>D</td><td>E</td></tr>
<tr><td>Simpson’s index</td><td>36</td><td>36</td><td>36</td><td>36</td><td>34</td></tr>
<tr><td>Species richness</td><td>38</td><td>38</td><td>38</td><td>38</td><td>36</td></tr>
</table></td></tr>
<tr><td>Notes</td><td>Complex article with lots of data. Table 2 provides preburn frequency data for datasets 1-4, whilst figure 1 provides post burn data after 2 years (3 years for some replicates). Data is available for forbs in addition to the figure 1 data (see figure 2) but it cannot be read off accurately. Table five provides pre burn data for dataset 5 combining three stands of variable age before burning. Figure seven provides post burn information for four species and total species number.</td></tr>
</table>

<table>
<tr><th>Study</th><th colspan="5">Stevenson et al. 1996</th></tr>
<tr><td>Methods</td><td colspan="5">Site comparison based on a comparison of stands unburned for 70-190 years and serially burnt stands (determined using paleoecology) at five sites. Species constancy assessed using 50 quadrats per site in two randomized blocks. There are two datasets based on sites with short recovery times (Trochy hill and Gallow hill) (A), and long recovery times (Auldallan hill, Blacklow hill and Happas farm) (B). The unburned stands form control arms and the serially burnt stands form treatment arms.</td></tr>
<tr><td>Population</td><td colspan="5">Stand age at time of burning: unknown.*
Community type / Initial floristic composition: A H10, B H10 (with some affinities to H16 and H21).
Altitude: A 325m, B 243m.</td></tr>
<tr><td>Intervention and Co interventions</td><td colspan="5">Post burn time (treatment): A authors state short recovery time therefore we are using 4 years, B 19 years.
Equivalent monitoring time (control): monitored at same time as treatment stands.</td></tr>
<tr><td rowspan="4">Outcomes</td><td></td><td colspan="2">Simpson’s index</td><td colspan="2">Species richness</td></tr>
<tr><td>Dataset</td><td>A</td><td>B</td><td>A</td><td>B</td></tr>
<tr><td>Treatment</td><td>0.965</td><td>0.943</td><td>30</td><td>21</td></tr>
<tr><td>Control</td><td>0.935</td><td>0.95</td><td>15.5</td><td>20.33</td></tr>
<tr><td>Study design</td><td colspan="5">Site Comparison: 40</td></tr>
<tr><td>Baseline Comparison</td><td colspan="5">The location, altitude and presumably community type were similar in treatment and control for all stands but no information is presented concerning diversity, initial floristic composition or stand age at time of burning (This is unsurprising given the timescale): 3</td></tr>
<tr><td>Intra treatment variation</td><td colspan="5">Post burn time is similar within treatment arms but stand age at time of burning is unreported* and Community type, location and altitude are variable: 2</td></tr>
<tr><td>Measurement of intervention and Cointerventions</td><td colspan="5">Soil moisture is variable between replicates but is balanced in treatment and control arms, However there is no information about grazing: 1</td></tr>
<tr><td>Replication & parameter of abundance</td><td colspan="5">Replicated objective parameter of abundance used (x and n presented): 2</td></tr>
<tr><td>Number / Proportion of species monitored</td><td colspan="5">Constancy data presented in appendix 4 represents the majority but not all species (eg Carex binervis, Vaccinium vitis-idaea data not included): 1</td></tr>
<tr><td>Attrition bias</td><td colspan="5">No losses to follow up: 2</td></tr>
<tr><td>Sum of Data quality</td><td colspan="5">51</td></tr>
<tr><td>Notes</td><td colspan="5">This is the only study to look at fully rotational burning. Unfortunately, the long time scale means that the baseline comparison is hard or impossible to make for some factors. Likewise, stand age at time of burning and grazing history are unreported and inevitably there is some variation between replicates from different sites. *Stand age at time of burning is available (table 3.6 p 32) but pooling was by post burn time as this was missed during data extraction. This data should be reanalysed during any future amendment to account for this. Auldallan and Blacklaw have stand ages of >19 yrs, whereas Trochy, Happas and Gallow hill have stand ages of <5 years. Blacklaw hill experiences a loss of 6 species and Simpson’s index declines by -0.03 when comparing burnt to unburnt, thus although all other sites show an increase in Species richness and Simpson’s index the mean value for the two older stands is negative. (This is based on constancy data- the authors state that old unburnt stands have higher species richness than old serially burnt stands but the raw data is unavailable pending publication).</td></tr>
</table>

<table>
<tr><th>Study</th><th colspan="9">Welch Pers comm. (2004) raw data substituted for Welch et al. (1994)</th></tr>
<tr><td>Methods</td><td colspan="9">Site comparison based on frequency using 60 permanent positions and 10 pins per 1m frame. Six transects were burnt in 1983 (treatment arm), whilst two have been unburnt since the 1950s (control arm). Data has been extracted from monitoring in 1993 (A), 1996 (B), 1999 (C) and 2002 (D). These are not independent datasets as they share the same baseline.</td></tr>
<tr><td>Population</td><td colspan="9">Stand age at time of burning: unknown.
Community type / Initial floristic composition: H12.
Location: Glen Clunie.
Altitude: 600m (?).</td></tr>
<tr><td>Intervention and Co interventions</td><td colspan="9">Post burn time (treatment): A 10 years, B 13 years, C 16 years, D 19 years.
Equivalent monitoring time (control): monitored at same time as treatment stands.</td></tr>
<tr><td rowspan="4">Outcomes</td><td></td><td colspan="4">Simpson’s index</td><td colspan="4">Species richness</td></tr>
<tr><td>Dataset</td><td>A</td><td>B</td><td>C</td><td>D</td><td>A</td><td>B</td><td>C</td><td>D</td></tr>
<tr><td>Treatment</td><td>0.888</td><td>0.887</td><td>0.867</td><td>0.862</td><td>23.33</td><td>25.83</td><td>23</td><td>23.33</td></tr>
<tr><td>Control</td><td>0.72</td><td>0.765</td><td>0.765</td><td>0.805</td><td>15</td><td>15</td><td>16</td><td>20</td></tr>
<tr><td>Study design</td><td colspan="9">Site Comparison: 40</td></tr>
<tr><td>Baseline Comparison</td><td colspan="9">The location, altitude and community type, diversity and initial floristic composition were similar for all stands prior to treatment (baseline taken from 1981) but stand age at time of burning has not been ascertained: 3</td></tr>
<tr><td>Intra treatment variation</td><td colspan="9">Replicates are similar in all known respects but stand age at time of burning has not been ascertained: 2</td></tr>
<tr><td>Measurement of intervention and Cointerventions</td><td colspan="9">Sheep and Red deer grazing has been variable in time but is assumed to be constant across the site and monitoring times are equal in treatment and control arms (no information about soil moisture variation): 1</td></tr>
<tr><td>Replication & parameter of abundance</td><td colspan="9">Well replicated objective parameter of abundance used (x and n presented): 4</td></tr>
<tr><td>Number / Proportion of species monitored</td><td colspan="9">All species monitored in treatment and control: 2</td></tr>
<tr><td>Attrition bias</td><td colspan="9">No losses to follow up: 2</td></tr>
<tr><td>Sum of Data quality</td><td colspan="9">54</td></tr>
<tr><td>Notes</td><td colspan="9">The datasets are not independent because they share the same baseline. Stand age and soil moisture status are not known but measurement of outcome is of high quality as a result of experimental protocol and contacting the authors for raw data.</td></tr>
</table>

<table>
<tr><th>Study</th><th colspan="3">Mallik and Gimingham 1983</th></tr>
<tr><td>Methods</td><td colspan="3">Time series based on cover abundance in 128 subplots from 3 stands with before and after treatment data. Before burning is the control arm and post burning is the treatment arm. There is only before and after data on species richness.</td></tr>
<tr><td>Population</td><td colspan="3">Stand age at time of burning: 15 years.
Community type / Initial floristic composition: H16.
Location: Muir of Dinnet, NE Scotland.
Altitude: 180m.</td></tr>
<tr><td>Intervention and Co interventions</td><td colspan="3">Post burn time: 1 year.
Equivalent monitoring time (control): NA.</td></tr>
<tr><td rowspan="3">Outcomes</td><td></td><td>Simpson’s index</td><td>Species richness</td></tr>
<tr><td>Treatment</td><td>NA</td><td>21.33</td></tr>
<tr><td>Control</td><td>NA</td><td>22.33</td></tr>
<tr><td>Study design</td><td colspan="3">Time series: 20</td></tr>
<tr><td>Baseline Comparison</td><td colspan="3">Treatment and control arms homogenous (time series): 6</td></tr>
<tr><td>Intra treatment variation</td><td colspan="3">Community type, location and altitude, post burn time and stand age at time of burning very similar for replicates: 4.</td></tr>
<tr><td>Measurement of intervention and Cointerventions</td><td colspan="3">No information about grazing or soil moisture: 0</td></tr>
<tr><td>Replication & parameter of abundance</td><td colspan="3">Well Replicated objective parameter of abundance used (x and n presented): 4</td></tr>
<tr><td>Number / Proportion of species monitored</td><td colspan="3">All species monitored for treatment and control: 2</td></tr>
<tr><td>Attrition bias</td><td colspan="3">No losses to follow up: 2</td></tr>
<tr><td>Sum of Data quality</td><td colspan="3">38</td></tr>
<tr><td>Notes</td><td colspan="3">Data based on table 3. No data presented on abundance just presence absence therefore cannot ascertain Simpson’s index.</td></tr>
</table>

Study	King 1960		
Methods	Site comparison based on five unburnt stands and one burnt stand with cover (Domin Scale) and constancy measures. Replication methodology is unclear especially with regard to the burnt stand.		
Population	Stand age at time of burning: >20 years Community type / Initial floristic composition: H12. Location: Cheviot hills (SE Scotland) Altitude: 495m		
Intervention and Co interventions	Post burn time (treatment): 9 years Equivalent monitoring time (control): monitored at same time as treatment stands		
Outcomes		Simpson's index	Species richness
	Treatment	0.961	30
	Control	0.974	41
Study design	Site Comparison, 40		
Baseline Comparison	The location, altitude, stand age prior to burning and community type were the same for treatment and control but no information is presented concerning diversity or initial floristic composition prior to burning: 3		
Intra treatment variation	No heterogeneity within treatment and control arms with respect to community type, location, altitude, post burn time and stand age at time of burning as far as can be ascertained: 4		
Measurement of intervention and Cointerventions	Sheep grazing with summer cattle grazing assumed to be equal across the site, monitoring times are equal in treatment and control arms (There is some soil moisture variation but it is not clear how this is balanced as regards treatment and control): 1		
Replication & parameter of abundance	Subjective parameter of abundance used (x presented): 0		
Number / Proportion of species monitored	Comprehensive species list (although *Cladonia* are lumped) representing all species on the sites: 2		
Attrition bias	No losses to follow up: 2		
Sum of Data quality	52		
Notes	Data based on table 1. It is not clear exactly how many samples are in each arm or which stands are assigned to treatment or control. No measure of variance can be obtained from the data presented which prohibits use of this study in meta-analysis.		

Study	Miller 1979		
Methods	Site comparison based on three stands >20 years since burning (control) and three burnt stands with cover estimated by eye in nine plots per stand.		
Population	Stand age at time of burning: Unknown. Community type / Initial floristic composition: H10? Very little information other than *Calluna vulgaris* production. Location: Kerloch (Kincardineshire). Altitude: 210m.		
Intervention and Co interventions	Post burn time (treatment): 11 years. Equivalent monitoring time (control): monitored at same time as treatment stands.		
Outcomes		Simpson's index	Species richness
	Treatment	NA	1
	Control	NA	1
	Only *Calluna* was monitored. *Calluna* cover increases with post burn time.		
Study design	Site Comparison: 40		
Baseline Comparison	The location, altitude and community type were similar for treatment and control but no information is presented concerning diversity or initial floristic composition prior to burning: 3		
Intra treatment variation	There is heterogeneity within treatment and control arms with respect to post burn time although location, altitude and community type were similar: 2		
Measurement of intervention and Cointerventions	Grazing by grouse and hares assumed to be equal across the site, monitoring times are equal in treatment and control arms (No information on soil moisture): 1		
Replication & parameter of abundance	Subjective parameter of abundance used (x presented): 0		
Number / Proportion of species monitored	Only one species monitored: 0		
Attrition bias	No losses to follow up: 2		
Sum of Data quality	48		
Notes	Data based on table 5. Only one species was monitored. No measure of variance can be obtained from the data presented which prohibits use of this study in meta-analysis.		