

Systematic Review

A systematic review of the efficacy of interventions for dynamic intermittent dorsal displacement of the soft palate

K. J. ALLEN*, R. M. CHRISTLEY†, M. A. BIRCHALL‡ and S. H. FRANKLIN§

Department of Clinical Veterinary Science, University of Bristol, UK

†Department of Epidemiology and Population Health, Institute of Infection and Global Health, University of Liverpool, UK

‡Faculty of Biomedical Science, University College London Ear Institute, UK

§School of Animal and Veterinary Sciences, University of Adelaide, Australia.

*Corresponding author email: kate.allen@bristol.ac.uk; Received: 10.02.11; Accepted: 09.03.11

Summary

There are numerous treatments for correction of dorsal displacement of the soft palate (DDSP). However, the efficacy of these treatments is controversial and there is little consensus on how best to treat this condition. The aims of this study were to systematically review the literature and to assess the evidence on the clinical effects of interventions for dynamic intermittent DDSP. A secondary objective was to assess whether factors relating to study quality affected reported success rates. Twenty-three studies were included, covering a wide number of interventions but also differing widely in terms of study design, sample size, method of diagnosis, outcome measure and the number lost to follow-up. The assessment of adverse effects was severely limited because of lack of reporting. The way in which success is measured appears to have a great effect on the reported results. Research synthesis has been severely limited because of the heterogeneity in the included studies. The low level of evidence makes it difficult to draw firm conclusions as to the efficacy of procedures for DDSP. Hence it is currently not possible to determine which procedure is the most appropriate. This systematic review highlights the difficulties of studying palatal dysfunction and suggests areas where improvements can be made in future studies.

Keywords: horse; systematic review; DDSP; interventions; efficacy

Introduction

Decisions for interventions are often made based upon the most recent or well known study or upon expert opinion (Sheldon 2005). A systematic review is a process of combining information from all relevant studies to understand the knowledge base better, aiming to improve the reliability and accuracy of recommendations when compared to single studies. Systematic reviews are particularly useful when there are variations in clinical practice and when there is uncertainty over potential benefits and harms of an intervention. Studies may be combined in 2 ways: meta-analysis is the statistical synthesis of results of similar studies into a single quantitative estimate of effect (Haynes 2006), whilst narrative synthesis is the process of synthesising primary studies to explore heterogeneity descriptively rather than statistically (Popay *et al.* 2006).

This is the first systematic review to be published in *Equine Veterinary Journal* and examines the efficacy of interventions for dorsal displacement of the soft palate (DDSP). In many circumstances, human systematic reviews are large volumes of work and as a result are often published in online databases such as The Cochrane Collaboration. Currently, no such database exists for equine systematic reviews. Therefore much of the information in this systematic review has been made available in the online supplementary items.

Palatal dysfunction in exercising horses comprises palatal instability (PI), which may or may not progress to DDSP. Palatal instability is described as dorsoventral billowing movements of the caudal portion of the soft palate, with flattening of the epiglottis against the dorsal surface of the soft palate (Kannegieter and Dore 1995; Tan *et al.* 2005; Lane *et al.* 2006a). In DDSP the caudal border of the soft palate displaces to a position above the epiglottis, obstructing the *rima glottidis* (Parente *et al.* 2002; Franklin *et al.* 2004; Lane *et al.* 2006a). Whilst the exact prevalence of the condition remains unknown, palatal dysfunction is the most common dynamic upper respiratory tract (URT) obstruction (Morris and Seeherman 1991; Kannegieter and Dore 1995; Martin *et al.* 2000; Lane *et al.* 2006a).

The aetiopathogenesis of DDSP remains unclear and numerous treatment options have been developed in order to address the different proposed mechanisms. However, it should be noted that, for many interventions, there is limited scientific evidence to confirm the rationale for their use. The efficacy of treatments therefore remains controversial with little consensus about how best to treat this condition.

Objectives

The aims of this paper were to review the literature systematically to assess the evidence on the efficacy and harms of interventions for dynamic intermittent DDSP and to assess whether factors relating to study quality affected the reported success rates.

Inclusion criteria for studies in this review

Types of studies

Studies of level 4 evidence and above (<http://www.cebm.net/index.aspx?o=1025>) were included. Where comparator groups were studied these included interventions that were compared with each other or with affected horses that underwent no intervention or a comparison population without the condition.

Participants

Mature horses in which naturally occurring dynamic intermittent DDSP was diagnosed (as defined in the primary source reports) were included. Foals, cases with persistent DDSP, experimentally-induced DDSP and horses undergoing concurrent interventions for other URT conditions were excluded.

Interventions

Surgical interventions: Individual and combination surgical procedures were included. Studies were excluded when the results of different surgical procedures or different combinations of surgical procedures were presented as one result and it was not possible to determine which success rates were related to which intervention. However, for several procedures, slight variations in the individual surgical techniques were permitted. For myectomy/tenectomy, variation in muscle group and method of resection was allowed. For tension palatoplasty, concurrent subepiglottal resection was allowed and for the laryngeal tie-forward procedure, concurrent sternothyroid tenectomy was allowed.

Conservative and medical interventions: Individual and combination conservative interventions were included. Studies were permitted when the results of different conservative procedures were presented as one result.

Outcome measures

The success of the intervention (in the primary source reports) included either subjective or objective outcome measures. Subjective measures include assessment by the owner/trainer of decreased respiratory noise or increased performance. Objective outcome measures include the analysis of race form pre- and post intervention and pre- and post intervention treadmill endoscopic examination. All adverse effects of the intervention reported in the trial were included.

Publication

Publications from 1990 onwards, with English language text copy were considered.

Search methods

Studies were identified from electronic databases including MEDLINE, PUBMED, ISI Web of Science, CAB abstracts, EMBASE and IVIS (for conference proceedings) in October 2008 and repeated in November 2009. The search string used was (horse OR equine) AND (dorsal displacement of the soft palate OR soft palate) AND (treatment). Bibliographies of referenced textbooks and the reference lists of all retrieved studies were hand-searched for additional relevant studies. Eighteen authors/centres were contacted to identify unpublished and ongoing studies.

All retrieved bibliographic references were managed in EndNote X reference manager software¹. The abstracts and titles of references retrieved were screened for relevance. Full paper copies of potentially relevant articles were assessed for inclusion by 2 independent reviewers (K.A./S.F.).

Methods of the review

Assessment of study quality

Study quality was assessed according to pre-established criteria (Supporting item 1).

Data extraction, statistical analysis and data synthesis

Data for each study and relevant results are presented in summary tables (Supporting item 2). The effect measures reported by the trial authors were used. Where possible, a quantitative analysis was performed and effectiveness summarised as odds ratio using 95% confidence intervals. The statistical analysis was performed using Review Manager software² using the Mantel-Haenszel method for dichotomous data.

Due to variations in study design, diagnostic method, comparator and outcome measure for each intervention a meta-analysis was not applicable. Therefore, results were combined using a narrative synthesis (Popay *et al.* 2006). Differences between studies assessing the same intervention were explored narratively by examining differences in the study design and quality, diagnostic method and outcome measure.

Results

Quantity and quality of the research available

The combined searches identified 1193 studies, of which 1117 were excluded based on the title. The remaining 76 studies were assessed and 23 included in this review (Table 1). The bibliographic details of excluded studies and reasons for exclusion are detailed (Supporting item 3).

The evidence base covers a wide number of interventions, but differs widely in terms of study design, sample size, method of diagnosis, outcome

measure and the number of cases lost to follow-up. There was an overall preponderance of studies towards surgical interventions.

Nine studies were case series and 2 were case reports, in which no comparator group was included. Five studies were described as case-control studies; however, insufficient information about the comparator population was provided so the appropriateness of the comparisons could not be fully assessed. In particular, where control horses were randomly selected, their DDSP status was unknown. Seven studies compared 2 surgical interventions; however, only 3 assessed this statistically.

Only 3 studies (plus the 2 case reports) were based on horses in which a definitive diagnosis of DDSP was made in all included horses. In a further 2 studies, a definitive diagnosis of palatal dysfunction (PI or DDSP) was achieved in all horses.

Three studies determined efficacy of the procedure using subjective outcome measures, 14 used race performance and 3 reported both subjective and race performance outcomes. Only 2 studies and the 2 case reports used exercising endoscopy as an outcome measure. One of these studies also used subjective measures. The sample sizes varied from 1–405 at the intervention stage and from 1–197 at the analysis stage. In no study was a sample size calculation performed. In many studies a large number of horses underwent the procedure and were not included in the efficacy analysis. Some studies failed to report how many horses underwent the procedure but did not meet the inclusion criteria for analysis. For the 16 studies in which this information was available, 11 studies had <80% of horses in the analysis.

Assessment of adverse effects was severely limited because of lack of reporting. It was frequently unclear whether no adverse events occurred or whether adverse events were not reported.

Intervention summaries

A summary of the evidence to support each of the following interventions is available in Supporting item 4: oral palatopharyngoplasty (Ahern procedure), oral palatoplasty by thermal cautery, laryngeal tie-forward, composite surgery- combination of laryngeal tie-forward and thermal cautery, staphylectomy, myectomy/tenectomy of *sternothyrohyoideus/omohyoideus*, composite surgery - combination surgery including *sternothyroideus* myotomy/tenectomy and staphylectomy, palatal sclerotherapy, epiglottic augmentation, medical and conservative interventions. The heterogeneity of studies means that results are often not directly comparable. It is therefore difficult from the current evidence to draw firm conclusions regarding the true efficacy of these procedures or to determine which procedures might be the most successful and least harmful for treatment of DDSP.

Effect of study quality on results

It is probable that several factors relating to study quality may influence the reported results. This systematic review permitted preliminary conclusions to be made for diagnosis, outcome measure and previous surgery. It was not possible to determine how the proportion of horses that underwent the procedure and were not included in the final analysis affected the results.

Effect of inclusion criteria: presumptive or definitive diagnosis

There was variation between studies as to whether only DDSP was assessed or whether palatal dysfunction (PI or DDSP) was assessed and how the diagnosis was made. Many studies relied on a presumptive diagnosis of DDSP based on clinical history and/or resting endoscopic findings. Definitive diagnosis was possible only through the use of exercising endoscopy. However, there was variation between studies as to whether PI was ever diagnosed, whether PI was considered presumptive of DDSP or whether PI and DDSP were grouped together and considered to be definitive of palatal dysfunction.

Franklin *et al.* (2009) found no significant difference in the success rates between horses diagnosed with DDSP and those diagnosed with PI during treadmill endoscopy. Similarly, in the study by Woodie *et al.* (2005) there was no significant difference in results between horses with a definitive diagnosis of DDSP and those with a presumptive diagnosis based on

TABLE 1: The methodological features of included studies

Study reference	Sample size at intervention stage	Sample size at analysis stage	What percentage of horses included in this review had a definitive diagnosis?	Outcome measure used	Were adverse effects reported?
Ahern (1993b)	111	100* of which 95 included in this review+	0	Subjective	No
Anderson et al. (1995)	209	149	0	Objective (race data) Subjective	No
Barakzai and Dixon (2005)	Unclear	31	23% DDSP (+6% PI)	Objective (race data)	No
Barakzai et al. (2004)	104	53	0	Objective (race data)	No
Barakzai et al. (2009a)	Unclear	78	100% DDSP	Objective (race data)	Yes
Bonenclark et al. (1999)	87	44	Unclear	Objective (race data)	No
Cheetham et al. (2008)	263	106	34% DDSP (+12% PI)	Objective (race data)	No
Dart (2006)	n/a	1	100% DDSP	Objective (treadmill endoscopy)	Yes
Duncan (1997)	Unclear	50	0	Objective (race data)	Yes
Dykgraaf et al. (2005)	96	77* of which 58 included in this review+	Unclear	Objective (race data)	No
Franklin et al. (2002)	6	6*	100% DDSP	Objective (treadmill endoscopy)	No
Franklin et al. (2009)	234	197*	61% DDSP (+39% PI)	Objective (race data)	No
Llewellyn and Petrowitz (1997)	405	41	0	Objective (race data)	Yes
Marcoux et al. (2008)	8	8*	0	Subjective	Yes
McCluskie et al. (2009)	116	42, of which 29 included in this review+	72% DDSP (+28% PI)	Objective (treadmill endoscopy) Subjective	No
Ordidge (2001)	252	187	0	Subjective	Yes
Parente et al. (2002)	92	32	100% DDSP	Objective (race data)	No
Peloso et al. (1992)	1	1	100% DDSP	Objective (treadmill endoscopy)	Yes
Picandet et al. (2005)	Unclear	51	0	Objective (race data) Subjective	Yes
Reardon et al. (2008a)	Unclear	110	0	Objective (race data)	Yes
Reardon et al. (2008b)	98	35	0	Objective (race data)	No
Smith and Embertson (2005)	102	73	Unclear	Objective (race data)	Yes
Woodie et al. (2005)	116	98, of which 20 included in this review+	100% DDSP	Objective (race data) Subjective	Yes

DDSP: dorsal displacement of the soft palate; PI: palatal instability; *: $\geq 80\%$ of horses that underwent procedure were included in analysis; +: horses that underwent concurrent surgeries for other URT obstructions were removed from this review.

history and resting endoscopy or history alone. However, in a subsequent study by the same group (Cheetham et al. 2008), horses with a definitive diagnosis of DDSP were less likely to race post operatively compared with horses that had a presumptive diagnosis. Only 66% of definitively diagnosed cases raced post operatively whereas 84% of presumptively diagnosed cases raced post operatively and the analysis performed for this review showed that having only a presumptive diagnosis significantly favoured post operative racing (Fig 1). The presumptive diagnosis category comprised horses in which treadmill endoscopy was not performed (81%)

and horses in which PI, but not DDSP, was observed during treadmill endoscopy (19%). Although firm conclusions cannot be made it is possible that success rates may be lower when only definitively diagnosed cases are reported.

Effect of outcome measure: There were wide variations in the outcome measure used and some studies used many outcomes. There is some evidence that the outcome measure may have a substantial effect on the reported results.



Fig 1: Forest plot showing success rate for the definitively diagnosed group compared with a presumptively diagnosed group. The results suggest that having a presumptive diagnosis significantly favours having a post operative start.

Return to racing (i.e. a post operative start) may not be the most appropriate indicator of a successful surgical outcome. Several authors reported the proportion of horses that return to racing and then provided a more stringent definition of success (i.e. increased earnings) (Duncan 1997; Barakzai *et al.* 2004, 2009a; Barakzai and Dixon 2005; Dykgraaf *et al.* 2005; Franklin *et al.* 2009) (Table 2). In all cases the proportion of horses considered to be successfully treated was substantially lower (up to 48% lower) than the proportion that returned to racing, showing that a post operative start may be an optimistic measure of success.

The way race performance is assessed may also substantially affect the apparent success rates (Reardon *et al.* 2008a,b; Franklin *et al.* 2009; Barakzai *et al.* 2009a). Success rates may be affected by race parameter used as well as the number of races. Reardon *et al.* (2008a) showed a significant, but very weak correlation between ratings and earnings and found significant differences between ratings and performance index and earnings and performance index. This resulted in variation in success rates (28–51% and 42–67%) when different parameters were examined over the same time period (Reardon *et al.* 2008a,b). The number of races assessed before and after an intervention also had an effect on the reported success rates (Barakzai *et al.* 2009a; Franklin *et al.* 2009). Franklin *et al.* (2009) showed that the effect of multiple variations in racing parameters was considerable. When using the same horse data, but varying between ratings and earnings, together with varying the number of races used, the apparent 'success rate' varied widely from 32–59%, 26–62% and 38–73% for 3 different interventions.

For subjective outcome measures (by trainer questionnaire) results may vary depending on the question asked and individual opinion in what constitutes success. In one study, 72% of horses were considered by the trainer to be successfully treated (Ordidge 2001). However, if success was more strictly defined to be 'cessation of gurgling noise' only 48% of the horses that were reported to make 'gurgling' noise presurgery would be considered to be successfully treated.

Only one study included both subjective assessment and objective assessment using race performance in a similar population of horses (Woodie *et al.* 2005). The reported success rates were very similar (86% for subjective methods and 82% for objective methods). In contrast subjective trainer assessment had no correlation with improvement in upper airway function assessed by repeat treadmill endoscopic examination (McCluskie *et al.* 2009). However, when repeat endoscopy is used as an outcome measure this may also lead to variation in apparent success rate because it is uncertain what level of palatal stability should constitute a successful outcome (McCluskie *et al.* 2009).

Confounding variables (previous surgery): Woodie *et al.* (2005) found no effect of prior surgery on the trainers' assessment of performance. However, a significant improvement in performance index was found in horses that had not undergone previous surgery compared with those in which there had been previous surgical interventions (Woodie *et al.* 2005). In contrast, Parente *et al.* (2002) suggested that there was a significant association between previous surgery and a positive performance outcome.

TABLE 2: Differences in results where post operative start is the measure of success compared with those using a different race performance outcome as a measure of success

	Percentage of horses that raced post operatively	The trial authors determination of success using a different race performance outcome (%)	Percentage difference
Barakzai and Dixon (2005)	94	61	33
Barakzai <i>et al.</i> (2004)	92.5	60	32.5
Barakzai <i>et al.</i> (2009a)	83	35	48
	93	53	40
Duncan (1997)	94	70	24
Dykgraaf <i>et al.</i> (2005)*	88	62	26
Franklin <i>et al.</i> (2009)*	96	48	48

*Results from combined interventions.

Discussion

This is the first systematic review in the field of equine dynamic URT disorders. Numerous difficulties were encountered and undertaking evidence-based veterinary medicine is challenging due to a 'serious lack of high quality patient centred veterinary research' (Murphy 2002).

Systematic reviews should contain studies of the highest available level of evidence. Well conducted randomised controlled trials are the preferred study design because they are least likely to be biased (Reeves *et al.* 2008) but have been avoided in equine veterinary practice because of methodological, financial and ethical constraints (Murphy 2002). Therefore, the inclusion criteria were widened to level ≥ 4 evidence to more fully consider the current evidence base. Broad inclusion criteria were used (Stroup *et al.* 2000), with the aim of analysing differences in the study designs and their relationship to the reported outcomes.

This systematic review included all relevant studies regardless of publication status, with independent assessment of study quality. All the included studies had been peer reviewed; however, the stringency of this process is likely to vary between (and within) journals and conferences. Other veterinary systematic reviews have restricted inclusion criteria to peer reviewed journals (Olivry and Mueller 2003; Nuttal and Cole 2007) as it has been suggested that studies that have not been peer reviewed may have unreliable results (Chalmers *et al.* 1987). However, publication bias (Meakins 2002) is thought to be greatest for small nonrandomised studies (Newcombe 1987; Easterbrook *et al.* 1991; Dickersin and Min 1993). Therefore guidelines from the Cochrane Collaboration (<http://www.cochrane.org>) and Centre for Reviews and Dissemination (<http://www.york.ac/inst/crd/>) suggest that reviews should aim to include all relevant studies, regardless of publication status.

Search strategies used widely in the medical field (Haynes *et al.* 1994) may not be effective for locating veterinary literature (Murphy 2002, 2003); therefore, a broad search query was used (Olivry and Mueller 2003; Aragon *et al.* 2007; Nuttal and Cole 2007). PubMed yielded relatively few results, whereas the other electronic databases yielded large numbers of irrelevant studies, confirming that no one database provides comprehensive indexing to all relevant veterinary literature (Murphy 2002).

Development of the inclusion criteria proved problematic. The initial aim was to assess mutually exclusive interventions only. However, after initial review of the database, it became clear that for many studies variation in the interventions undertaken was present. Formulating a question that strikes a justifiable balance between the ideal and the feasibility of answering the question is important (Haynes 2006). Hence, the inclusion criteria were redefined, with the aim of more fully understanding the evidence base, particularly for interventions currently being performed. Due to the limited studies assessing conservative techniques, studies were permitted when the results of different conservative procedures were presented as one result. For surgical interventions involving myectomy/tenectomy, tension palatoplasty or laryngeal tie-forward, variation in the surgical technique was allowed. Subepiglottic resection performed in conjunction with any tension palatoplasty procedure was classified as variation in technique. This was described as part of the

original technique (Ahern 1993a) but has subsequently been omitted by some surgeons. Furthermore, where studies combined the results from horses undergoing tension palatoplasty with and without subepiglottic resection, statistical analysis was performed prior to grouping and showed no significant difference in success rates (Franklin et al. 2009). It was also decided to permit laryngeal tie-forward with and without sternothyroid tenectomy as variation in technique, because both of these procedures affect rostral positioning of the larynx. Sternothyroid tenectomy was not described in the original study (Woodie et al. 2005) but was described as a modification of the technique later (Ducharme 2005). Again, a statistical analysis showed no significant difference in success rates between groups (Franklin et al. 2009). However, sternothyroid tenectomy performed in conjunction with another form of surgery (e.g. tension palatoplasty) was classified as a separate technique because it would not be possible to determine whether the results might arise due to alterations in laryngo-hyoid positioning or changes in the tension of the palatal tissues themselves. The original more stringent inclusion criteria would have resulted in exclusion of the majority of studies in which a definitive diagnosis was achieved (Barakzai and Dixon 2005; Cheetham et al. 2008; Franklin et al. 2009; Barakzai et al. 2009a; McCluskie et al. 2009). As better studies become available for all interventions, the inclusion criteria should be more strictly defined in future reviews.

Quality assessment indicates the likelihood that the results are a valid estimate of the truth (Moher et al. 1995). Differences in study quality may explain the heterogeneity in the results. Study quality assessment checklists developed for human studies were not suitable for this review; therefore, a quality assessment checklist was developed to identify the main potential limitations for each study. However, it is still unknown which of these criteria are the most important in establishing study quality. Hence, for studies that fulfil different criteria, it remains unclear which represents the better quality study.

Most studies in this field are before-and-after studies reporting pre- and post intervention data. Other studies used the same approach of describing pre- and post intervention data but have also used a comparison group, such as a different intervention (parallel group study). Several authors described studies as case-control studies whereby pre- and post intervention data for 'cases' (diagnosed with DDSP) were compared to the same data for 'control' horses (not diagnosed with DDSP). The ideal 'control' group for intervention studies is cases with DDSP that undergo no treatment (Cheetham et al. 2008; Barakzai et al. 2009a). As this is difficult to achieve, it was suggested that the 'control' group could be unaffected horses. However, in no study was any attempt made to confirm the 'controls' were DDSP negative and therefore a valid comparison. Furthermore, knowledge of the outcome status before collection of exposure information is the defining feature of a case-control study (Fosgate and Cohen 2008). It has previously been suggested that these studies should not be classified as case-control studies (Fosgate and Cohen 2008); therefore for the purposes of this review they were reclassified as cohort studies.

There are difficulties in grading the level of evidence of veterinary studies as several different systems have been published and as yet there is no clear consensus (Innes 2007). Furthermore, many of the levels of evidence guidelines developed for human evidence based medicine do not fully include or differentiate the types of study included in this review. It has been argued that both hierarchies of study design and common sense judgement be used when assessing quality of research studies (Greenhalgh 2010). Hence, although it was difficult to rank individual studies, we generally considered studies in which a comparator group was used to be a better study design than a case series.

A problem with many studies is the lack of a definitive diagnosis prior to treatment. Results of studies conducted in which horses were not confirmed to have the disorder being investigated are potentially misleading. It was not possible to fully confirm to what degree this affected the results. However, the results of one study suggested that success rates may be lower in horses with a definitive diagnosis than those with a presumptive diagnosis. Several studies have documented that respiratory noise, resting endoscopy findings or both in conjunction may be unreliable in predicting dynamic events that occur during exercise (Morris and Seeherman 1991; Kanngieter and Dore 1995; Martin et al. 2000; Tan et al. 2005; Lane et al. 2006b; Witte et al. 2011; Barakzai and Dixon 2011).

However, 2 studies did demonstrate that the specificity of DDSP during resting endoscopy was high (Lane et al. 2006b; Barakzai and Dixon 2011); therefore intervention studies based on this criterion would have a low proportion of false positive diagnoses (Barakzai and Dixon 2011). It should be noted that most of the included studies used broader inclusion criteria based on resting endoscopy and history findings, and only 2 studies required all horses to demonstrate DDSP during resting endoscopy for inclusion (Duncan 1997; Reardon et al. 2008a). The low sensitivity of DDSP at rest (Lane et al. 2006b; Barakzai and Dixon 2011) should also be considered. Studies based on this criterion would include only a small subset of cases and it is unclear whether these are representative of the wider population of horses experiencing DDSP during exercise, or whether these cases might be more severely affected. Further clarification as to whether PI and DDSP are manifestations of the same condition is also important. Ideally intervention studies should be based on horses confirmed to have the disorder being investigated.

It is also important to confirm absence of other forms of dynamic URT collapse because this may have an impact on subsequent success rate. The prevalence of complex URT collapse is high (Tan et al. 2005; Lane et al. 2006a; Barakzai and Dixon 2011), and any additional forms of URT collapse tend not to be addressed in horses that do not undergo exercising endoscopy. With palatal dysfunction this may be complicated by the potential link between palatal dysfunction and axial deviation of the aryepiglottic folds (Parente et al. 1994; Ahern 2005; Tan et al. 2005; Lane et al. 2006a).

No study assessed co-interventions. The concurrent use of conservative measures such as a tongue tie following surgical treatments for DDSP (Barakzai et al. 2009b) may influence results. Other management changes are also likely to be important when race performance is used as the outcome measure. Veterinary diagnosis of DDSP and associated advice may inform these decisions. Previously performed URT surgeries may also be a confounding variable and were encountered in several studies (Ordidge 2001; Parente et al. 2002; Barakzai and Dixon 2005; Smith and Embertson 2005; Woodie et al. 2005), although the results were contradictory as to whether previous surgery had a positive or negative effect.

The way in which success is measured appears to have the greatest effect on the reported results. Outcome measures should be valid, consistent and accurate for the condition being investigated. At present there is a degree of uncertainty of the accuracy of the outcome measures used. Outcome measures used in veterinary surgery are starting to receive more attention (Brown 2008) and it has been suggested that whether the outcome measure is subjective or objective is not as critical as whether it is valid and reliable (Brown 2008). Furthermore, it is important that investigators define what constitutes treatment success or failure *a priori* and that a limited number of such measures should be used, or steps be taken to reduce the risk of type I error where multiple outcomes are considered.

Subjective measures of success by the trainer may provide useful information. However, it is unclear whether improvements in upper airway function can be accurately detected by a trainer's assessment of changes in respiratory noise and/or performance. In contrast to more objective measures, subjective assessment usually involves a retrospective pre-/post intervention assessment, and memory decay and even a placebo effect may affect results. It is also likely that trainers' perceptions of success may vary between horses. This review shows that it is important that questions are well formulated and specific.

Although the value of using race performance data may be justified as owners aim for improved racing performance, there are several elements that may make these results unreliable. The inference is that improved racing performance occurs because of improvements in URT function. However, racing performance is multifactorial and the multifactorial nature of poor performance and the high prevalence of complex forms of URT collapse will also influence subsequent racing performance. Also, differences in results between studies may be largely introduced by the different population of racehorses referred to each centre (Beard and Waxman 2007).

This review showed that the use of post operative start as an outcome measure results in high 'success rates'. However, it is likely that this is a weak indicator of success, simply because in many cases abnormal

respiratory noise or poor performance only occurs during racing; therefore, trainers have to enter the horse in a race to determine whether the intervention was successful. Cheetham *et al.* (2010) also suggested that the decision to use 'starts' compared with earnings as an outcome measure could have a marked effect on reported success rates. This systematic review also showed that large variations in success rates were observed when the race performance measure and number of races assessed is altered and this casts serious doubt on the validity of this outcome measure. Furthermore as race performance is often converted to a binomial outcome, it is unclear to what degree the results remain clinically relevant. For example, a horse only needs to earn £1 more after the intervention than before to be grouped in the success category. In a recent study from North America it was suggested that age, breed, sex, track surface and gait should be controlled for in the study design and analysis of race performance following an intervention (Cheetham *et al.* 2010). It is unclear whether there are other factors such as handicapping that should also be accounted for.

From a veterinary perspective, the use of repeat exercising endoscopy is probably the most sensible method to determine the efficacy of an intervention. Understanding how an intervention alters the structure and function of the pharynx in naturally occurring disease is of great importance and further studies undertaking this approach should be encouraged. Even so, it remains unclear to what extent palatal function should be restored to constitute success. DDSP is an intermittent event; hence, where PI occurs, post intervention it remains unclear as to whether this truly reflects resolution of DDSP or whether DDSP might occur during subsequent runs or under different exercise conditions. It is important that the same exercise test is undertaken pre- and post intervention and, for results to be clinically relevant, the exercise test should be representative of racing. Studies which evaluate the repeatability of DDSP under the same exercise test conditions are required before this method is truly valid. Unfortunately, smaller numbers of horses are likely to be included than for subjective or race performance studies. The development of overground endoscopy may better enable these studies; however, it is critical that exercise test design is appropriate (Allen and Franklin 2010).

In some studies, there were large differences in the number of horses that underwent the procedure and the number that were subsequently analysed. The greater this difference, the greater the potential for inaccurate results due to introduction of bias. The use of racing performance results in many cases, which have undergone the intervention, being excluded from the analysis as many horses will not have completed the requisite number of starts pre- and post intervention. This method probably creates a bias towards cases in which the intervention was successful, as cases in which the treatment was unsuccessful are less likely to continue racing. Recruiting cases for endoscopic studies is also difficult and inappropriate recruitment may introduce the opposite bias into the results, due to the possibility of poorly performing horses being more likely to be presented for reassessment. It is important that in the study design methods to reduce inclusion bias are taken. Furthermore, reasons for all exclusions should be specifically described. Substantial research needs to be undertaken on which outcome measures provide the most clinically relevant information and these should then be standardised between studies.

Systematic reviews only include efficacy studies in clinical cases, therefore several research studies that provide evidence to support or refute an intervention are not discussed in this review. However, in the authors' opinions, the disparity of the conclusions between experimental studies on normal horses and efficacy studies in clinical cases needs addressing.

The ability to draw conclusions regarding the potential treatment harms was also severely restricted due to under-reporting. As well as assessing short-term complications associated with surgery, it is necessary to investigate whether procedures fail to resolve palatal dysfunction and whether procedures result in worsening of this condition or induce any additional forms of upper airway collapse.

Research synthesis in this review has been severely limited because of the heterogeneity in the included studies. The systematic review suggests that factors relating to study methodology may have an important impact on the reported efficacy of procedures and therefore

the accuracy of the results. Whilst more recent studies have attempted to overcome weaknesses of earlier studies, substantial limitations still often exist. Overall, the low level of evidence makes it difficult to draw firm conclusions as to the efficacy of procedures for DDSP. Hence it is currently not possible to determine which procedure is the most appropriate. The intent of a systematic review is not to belittle individual studies. Rather, the purpose is to establish the limits of the current evidence base, which will allow future studies to target these areas. This systematic review has highlighted the difficulties of studying palatal dysfunction. Whilst many of these may not be readily overcome, the review highlights areas where improvement can be made and underlines the need for high quality studies, rather than just more studies.

Authors' declaration of interests

No conflicts of interest have been declared.

Source of funding

No source of funding.

Manufacturers' addresses

¹Endnote, <http://www.endnote.com>.
²RevMan, <http://www.cc-ims.net/revman>.

References

Ahern, T. (1993a) Oral palatopharyngoplasty. *J. equine vet. Sci.* **13**, 185-188.

Ahern, T. (1993b) Oral palatopharyngoplasty (part 2). *J. equine vet. Sci.* **13**, 670-672.

Ahern, T.J. (2005) Primary of secondary upper airways dysfunction? *J. equine vet. Sci.* **25**, 332-333.

Allen, K.J. and Franklin, S.H. (2010) Assessment of the exercise tests used during overground endoscopy in UK Thoroughbred racehorses and how these may affect the diagnosis of dynamic upper respiratory tract obstructions. *Equine vet. J. Suppl.* **38**, 587-591.

Anderson, J.D., Tulleners, E.P., Johnston, J.K. and Reeves, M.J. (1995) *Sternothyroideus myectomy or staphylectomy for treatment of intermittent dorsal displacement of the soft palate in racehorses: 209 cases (1986-1991)*. *J. Am. vet. med. Ass.* **206**, 1909-1912.

Aragon, C.L., Hofmeister, E.H. and Budsberg, S.C. (2007) Systematic review of clinical trials of treatments for osteoarthritis in dogs. *J. Am. vet. med. Ass.* **230**, 514-521.

Barakzai, S.Z. and Dixon, P.M. (2005) Conservative treatment for Thoroughbred racehorses with intermittent dorsal displacement of the soft palate. *Vet. Rec.* **157**, 337-340.

Barakzai, S.Z. and Dixon, P.M. (2011) Correlation of resting and exercising endoscopic findings for horses with dynamic laryngeal collapse and palatal dysfunction. *Equine vet. J.* **43**, 18-23.

Barakzai, S.Z., Boden, L.A., Hillyer, M., Marlin, D. and Dixon, P. (2009a) Efficacy of thermal cauterity for intermittent dorsal displacement of the soft palate as compared to conservatively treated horses: Results from 78 treadmill diagnosed horses. *Equine vet. J.* **41**, 65-69.

Barakzai, S.Z., Finnegan, C., Dixon, P.M., Hillyer, M.H. and Boden, L.A. (2009b) Use of tongue ties in Thoroughbred racehorses in the United Kingdom, and its association with surgery for dorsal displacement of the soft palate. *Vet. Rec.* **165**, 278-281.

Barakzai, S.Z., Johnson, V.S., Baird, D.H., Bladon, B. and Lane, J.G. (2004) Assessment of the efficacy of composite surgery for the treatment of dorsal displacement of the soft palate in a group of 53 racing Thoroughbreds (1990-1996). *Equine vet. J.* **36**, 175-179.

Beard, W.L. and Waxman, S. (2007) Evidence based upper respiratory tract surgery. *Vet. Clin. N. Am.: Equine Pract.* **23**, 229-242.

Bonenclark, G., Bryant, J., Hernandez, J., Ferrell, E.P. and Colahan, P.C. (1999) *Sternothyroideus tenectomy or sternothyroideus tenectomy with staphylectomy for the treatment of soft palate displacement*. *Proc. Am. Ass. equine Practnrs.* **45**, 85-86.

Brown, D.C. (2008) Outcome measures in veterinary surgery: the good, the bad and the ugly. In: *Proceedings of the American College of Veterinary Surgeons Symposium*. Available at: <http://www.acvs.org> (accessed August 2010).

Chalmers, T.C., Levin, H., Sacks, H.S., Reitman, D., Berrier, J. and Nagalingam, R. (1987) Meta-analysis of clinical trials as a scientific discipline I: control of bias and comparison with large co-operative trials. *Stat. Med.* **6**, 315-328.

Cheetham, J., Pigott, J.H., Thorson, L.M., Mohammed, H.O. and Ducharme, N.G. (2008) Racing performance following the laryngeal tie-forward procedure: A case-controlled study. *Equine vet. J.* **40**, 501-507.

Cheetham, J., Riordan, A.S., Mohammed, H.O., McIlwraith, C.W. and Fortier, L.A. (2010) Relationship between race earnings and horse age, sex, gait, track surface and number of race starts for Thoroughbred and Standardbred racehorses in North America. *Equine vet. J.* **42**, 346-350.

Dart, A.J. (2006) Vocal fold collapse after laryngeal tie-forward correction of dorsal displacement of the soft palate. *Vet. Surg.* **35**, 584-585.

Dickersin, K. and Mir, Y.I. (1993) Publication bias: The problem that won't go away. *Ann. N. Y. Acad. Sci.* **703**, 135-146.

Duncan, D. (1997) Retrospective study of 50 Thoroughbred Racehorses subjected to radical myectomy for treatment of dorsal displacement of the soft palate. *Proc. Am. Ass. equine Practnrs.* **43**, 237-238.

Ducharme, N.G. (2005) Management of DDSP: tie forward and throat support device. *Proceedings of the 3rd World Equine Airways Symposium*. pp 107-108.

Dykgraaf, S., McIlwraith, C.W., Baker, V.A., Byrd, W.J. and Daniel, R.C. (2005) *Sternothyroideus* tenectomy combination surgery for treatment of dorsal displacement of the soft palate in 96 Thoroughbred racehorses (1996-2004). *Proc. Am. Ass. equine Practnrs.* **51**, 323-326.

Easterbrook, P.J., Berlin, J.A., Gopalan, R. and Matthews, D.R. (1991) Publication bias in clinical research. *Lancet* **337**, 867-872.

Fosgate, G.T. and Cohen, N.D. (2008) Epidemiological study design and the advancement of equine health. *Equine vet. J.* **40**, 693-700.

Franklin, S.H., Naylor, J.R. and Lane, J.G. (2002) The effect of a tongue-tie in horses with dorsal displacement of the soft palate. *Equine vet. J., Suppl.* **34**, 430-433.

Franklin, S.H., Price, C. and Burn, J.F. (2004) The displaced equine soft palate as a cause of abnormal respiratory noise during expiration. *Equine vet. J.* **36**, 590-594.

Franklin, S.H., McCluskie, L.K., Woodford, N.S., Tremaine, W.H., Lane, J.G., Bladon, B.M., Barakzai, S.Z., Dixon, P.M., Hillyer, M.H. and Allen, K.J. (2009) A comparison of the efficacy of soft palate cauterization and the laryngeal tie-forward procedure in thoroughbred racehorses with a definitive diagnosis of palatal dysfunction. In: *Proceedings of the British Equine Veterinary Association Congress*, Equine Veterinary Journal Ltd., Newmarket. p 168.

Greenhalgh, T. (2010) *How to Read a Paper: The Basics of Evidence Based Medicine*, 4th edn., Wiley Blackwell.

Haynes, B. (2006) *Clinical Epidemiology: How to Do Clinical Practice Research*, 3rd edn., Eds: R.B. Haynes, D.L. Sackett, G.H. Guyatt and P. Tugwell, Lippincott Williams and Wilkins, USA.

Haynes, R.B., Wilczynski, N., McKibbin, K.A., Walker, C. and Sinclair, J.C. (1994) Developing optimal search strategies for detecting clinically sound studies in Medline. *J. Med. Inform. Ass.* **1**, 447-458.

Innes, J.F. (2007) Outcome-based medicine in veterinary surgery: Levels of evidence. *Vet. Surg.* **36**, 610-612.

Kannegieter, N.J. and Dore, M.L. (1995) Endoscopy of the upper respiratory tract during treadmill exercise: a clinical study of 100 horses. *Aus. vet. J.* **72**, 101-107.

Lane, J.G., Bladon, B., Little, D.R., Naylor, J.R. and Franklin, S.H. (2006a) Dynamic obstructions of the equine upper respiratory tract. Part 1: Observations during high-speed treadmill endoscopy of 600 Thoroughbred racehorses. *Equine vet. J.* **38**, 393-399.

Lane, J.G., Bladon, B., Little, D.R., Naylor, J.R. and Franklin, S.H. (2006b) Dynamic obstructions of the equine upper respiratory tract. Part 2: Comparison of endoscopic findings at rest and during high-speed treadmill exercise of 600 Thoroughbred racehorses. *Equine vet. J.* **38**, 401-407.

Llewellyn, H.R. and Petrowitz, A.B. (1997) *Sternothyroideus* myotomy for treatment of dorsal displacement of the soft palate. *Proc. Am. Ass. equine Practnrs.* **43**, 239-243.

Marcoux, M., Picandet, V., Celeste, C., Macieira, S., Morisset, S., Rossier, Y., Schambourg, M. and Jean, D. (2008) Palatal sclerotherapy: A potentially useful treatment of intermittent dorsal displacement of the soft palate in juvenile Standardbred racehorses. *Can. vet. J.* **49**, 587-591.

Martin, B.B., Jr, Reef, V.B., Parente, E.J. and Sage, A.D. (2000) Causes of poor performance of horses during training, racing, or showing: 348 cases (1992-1996). *J. Am. vet. med. Ass.* **216**, 554-558.

McCluskie, L.K., Franklin, S.H., Tremaine, W.H., Woodford, N.S. and Allen, K.J. (2009) Post operative high speed treadmill endoscopy to determine the efficacy of upper respiratory tract surgery in horses with palatal dysfunction. In: *Proceedings of the British Equine Veterinary Association Congress*, Equine Veterinary Journal Ltd., Newmarket. p 168.

Meakins, J.L. (2002) Innovation in surgery: The rules of evidence. *Am. J. Surg.* **183**, 399-405.

Moher, D., Jadad, A.R., Nichol, G., Penman, M., Tugwell, P. and Walsh, S. (1995) Assessing the quality of randomised controlled trials: An annotated bibliography of scales and checklists. *Control. Clin. Trials* **16**, 62-73.

Morris, E.A. and Seeherman, H.J. (1991) Clinical evaluation of poor performance in the racehorse: The results of 275 evaluations. *Equine vet. J.* **23**, 169-174.

Murphy, S.A. (2002) Applying methodological search filters to CAB Abstracts to identify research for evidence-based veterinary medicine. *J. Med. Libr. Ass.* **90**, 406-410.

Murphy, S.A. (2003) Research methodology search filters are they effective for locating research for evidence based veterinary medicine in PubMed? *J. Med. Libr. Ass.* **91**, 484-489.

Newcombe, R.G. (1987) Towards a reduction in publication bias. *Br. med. J. (Clin. Res. Ed.)* **295**, 656-659.

Nuttal, T. and Cole, L.K. (2007) Evidence-based veterinary dermatology: A systematic review of interventions for treatments of *Pseudomonas otitis* in dogs. *Vet. Dermatol.* **18**, 69-77.

Olivry, T. and Mueller, R.S. (2003) Evidence- based veterinary dermatology: A systematic review of the pharmacotherapy of canine atopic dermatitis. *Vet. Dermatol.* **14**, 121-146.

Ordridge, R. (2001) The treatment of dorsal displacement of the soft palate by thermal cauterization: A review of 252 cases. In: *Proceedings of the Second World Equine Airways Symposium*.

Parente, E.J., Martin, B.B., Tulleners, E.P. and Ross, M.W. (1994) Upper respiratory dysfunctions in the horse during high-speed exercise. *Proc. Am. Ass. equine Practnrs.* **40**, 81-82.

Parente, E.J., Martin, B.B., Tulleners, E.P. and Ross, M.W. (2002) Dorsal displacement of the soft palate in 92 horses during high-speed treadmill examination (1993-1998). *Vet. Surg.* **31**, 507-512.

Peloso, J.G., Stick, J.A., Nickels, F.A., Lumsden, J.M. and Derksen, F.J. (1992) Epiglottic augmentation by use of polytetrafluoroethylene to correct dorsal displacement of the soft palate in a standardbred horse. *J. Am. vet. med. Ass.* **201**, 1393-1395.

Picandet, V., Marcoux, M., Jean, D., Celeste, C., Macieira, S., Morisset, S. and Schambourg, M. (2005) Palatal sclerotherapy for the treatment of intermittent dorsal displacement of the soft palate in 50 Standardbred racehorses. In: *Proceedings of the Third World Equine Airways Symposium*. p 128.

Popay, J., Roberts, H., Sowden, A., Petticrew, M., Arai, L., Rodgers, M., Britten, N., Roen, K. and Duffy, S. (2006) *Guidance on the Conduct of Narrative Synthesis in Systematic Reviews*. A product from the ESRC methods programme. Institute for Health Research, University of Lancaster, Lancaster.

Reardon, R.J., Fraser, B.S., Heller, J., Lischer, C., Parkin, T. and Bladon, B.M. (2008a) The use of race winnings, ratings and a performance index to assess the effect of thermocautery of the soft palate for treatment of horses with suspected intermittent dorsal displacement. A case-control study in 110 racing Thoroughbreds. *Equine vet. J.* **40**, 508-513.

Reardon, R.J.M., Fraser, B.S.L. and Bladon, B.M. (2008b) Laryngeal tie-forward combined with thermocautery as treatment of horses with signs associated with intermittent dorsal displacement of the soft palate: a case-control study in British racing Thoroughbreds. In: *Proceedings of the British Equine Veterinary Association Congress*, Equine Veterinary Journal Ltd., Newmarket. p 340.

Reeves, B.C., Deeks, J.J., Higgins, J.P.T. and Wells, G.A. (2008) Including non-randomized studies. In: *Cochrane Handbook for Systematic Reviews of Interventions*, Version 5, Eds: J.P.T. Higgins and S. Green, The Cochrane Collaboration. Chapter 13. Available at: <http://www.cochrane-handbook.org> (accessed March 2009)

Sheldon, T. (2005) Making evidence synthesis more useful for management and policy making. *J. Health Serv. Res. Policy* **10**, Suppl. **1**, 1-5.

Smith, J.J. and Embertson, R.M. (2005) *Sternothyroideus* myotomy, staphylectomy, and oral caudal soft palate photothermoplasty for treatment of dorsal displacement of the soft palate in 102 thoroughbred racehorses. *Vet. Surg.* **34**, 5-10.

Stroup, D.F., Berlin, J.A., Morton, S.C., Olkin, I., Williamson, G.D., Rennie, D., Moher, D., Becker, B.J., Sipe, T.A. and Thacker, S.B. (2000) Meta-analysis of observational studies in epidemiology: A proposal for reporting. *J. Am. med. Ass.* **283**, 2008-2012.

Tan, R.H., Dowling, B.A. and Dart, A.J. (2005) High-speed treadmill videoendoscopic examination of the upper respiratory tract in the horse: The results of 291 clinical cases. *Vet. J.* **170**, 243-248.

Witte, S.H.P., Witte, T.H., Harriss, F., Kelly, G. and Pollock, P. (2011) Association of owner-reported noise with findings during dynamic respiratory endoscopy in Thoroughbred racehorses. *Equine vet. J.* **43**, 9-17.

Woodie, J.B., Ducharme, N.G., Kanter, P., Hackett, R.P. and Erb, H.N. (2005) Surgical advancement of the larynx (laryngeal tie-forward) as a treatment for dorsal displacement of the soft palate in horses: A prospective study 2001-2004. *Equine vet. J.* **37**, 418-423.

Supporting information

Additional Supporting Information may be found in the online version of this article:

Supporting item 1. Methodological assessment checklist.

Supporting item 2. Summary tables and data of included studies.

Supporting item 3. Excluded studies.

Supporting item 4. Intervention summaries.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.