EVIDENCE BASED APPROACH TO SURGERY FOR RECURRENT LARYNGEAL NEUROPATHY (RLN)
Safia Barakzai, BVSc, MSc, DESTS, Assoc. ECDVI, DECVS
The University of Edinburgh, United Kingdom

Key Points:
- There is a predominance of level 4 evidence evaluating surgical treatment of RLN.
- Many studies do not stand up well to rigorous methodological evaluation.
- Whilst randomisation of treatments and inclusion of an appropriate control population is extremely difficult in clinical equine studies, there are several other simple ways in which the evidence base could be significantly improved.

Despite a large number of papers having been published on treatments for recurrent laryngeal neuropathy, the critical evidence base to support these treatments is weak, consisting mainly of low level clinical evidence. When considering ‘levels of evidence’ (Figure 1), systematic reviews of multiple studies are considered the highest level of evidence (level 1a) followed by well conducted randomised controlled studies (RCTs, level 1b). Unfortunately, RCTs have been avoided in equine veterinary surgery because of methodological, financial and ethical constraints, and unsurprisingly, there is only one RCT published in the field of RLN surgery, which was performed in horses with experimentally induced RLN. Cohort studies represent level 2 in the clinical evidence ‘pyramid’, followed by case-control studies at level 3, case series and case reports at level 4 and expert opinion representing the lowest level of evidence at level 5.

Using ‘PubMed’ as an easy to access database to search for publications on RLN treatment in horses (1990-2012), including studies of level 4 and above and excluding cadaver studies, there appear to be only one RCT, two level 2 papers, 29 level 4 case series and 1 level 4 case report published in the peer reviewed literature. Seventeen of these describe treatment in horses with naturally occurring disease, and 12 papers examine horses that have had RLN experimentally induced. Objective methods of analysis are used in 18 papers, subjective in 8 papers and both subjective and objective analyses are used in 5 papers.

Figure 1: ‘Pyramid’ of evidence based studies. The layer size represents the volume of studies available at each level, with editorials and expert opinion representing the lowest level of evidence, and systematic reviews representing the highest level.

Assessing the methodological quality of individual studies can improve or reduce the level of evidence provided by the study. For instance, a poorly designed cohort study does
not necessarily provide better evidence than a well-designed and executed case-control study. These multiple factors that can affect the quality of a study also should be considered when designing new studies for potential publication. Methodological assessment of studies will usually include the following:

1. Was the assignment of patients to treatments randomised?
2. Is a comparison group used and is it appropriate?
3. Are case and control/comparison groups inclusion criteria clearly defined?
4. Were comparison groups similar at the start of the trial?
5. Were all patients who entered the trial accounted for?
6. Are greater than 80% included in follow up?
7. Were the outcome measures used appropriate and clearly defined?
8. Are potential confounders identified and controlled for?
9. Was a power calculation performed?
10. Is sample size likely to be sufficient to detect a clinically relevant effect?
11. Was there a clear description of the intervention and was it standardised between cases?
12. Were adverse effects reported?
13. Are the limitations of the study discussed?
14. Are the data analyses appropriate?
15. Are the conclusions of the study supported by the results?
16. Are the results generalizable to other populations of interest?

**Inclusion criteria for cases:**

**Method of diagnosis:**
Endoscopic examination at rest is the standard way in which to diagnose RLN in horses, however we know that resting endoscopy does not 100% predict laryngeal function at exercise. Thus the grades of laryngeal function at rest that are used as inclusion criteria may have an effect on the accuracy of this method of diagnosis. Comparisons of resting and exercising endoscopy have shown us that all Havemeyer grade 4 horses will have grade C function (complete collapse) during exercise. However, if, for instance, grade 3 horses are included in a laryngoplasty study without performing exercising endoscopy, a proportion of these horses (60% of grade 3.1 and 22% of grades 3.2-3.3) will likely have normal (grade A) exercising laryngeal function and it is then arguable whether a laryngoplasty was indeed indicated. Overall, Garrett et al. indicated that resting endoscopy has a sensitivity and specificity of 80 and 81% respectively for laryngeal function at exercise.

Additionally, multiple treadmill and over ground endoscopic studies have shown that a large proportion of horses (24-51%) exhibit multiple upper airway abnormalities during exercising endoscopy. The presence of undiagnosed and untreated concurrent upper airway abnormalities may of course have an effect on the efficacy of RLN surgery, particularly in terms of on-going abnormal noise and poor performance.

**Influence of other inclusion criteria:** Inclusion criteria based on horses fulfilling certain performance-related conditions will undoubtedly affect the end result achieved. Often, horses that do not have a full race history (e.g. 3 or 5 races both pre and post-operatively) are excluded from performance analysis. This method is used in order to provide a baseline for intra-horse comparison, however it is likely to exclude horses that were markedly affected with the disorder in training or which were detected endoscopically prior to entering training and thus were unable to race pre-treatment. It also excludes treatment failures which never
return to the track post-intervention. If information is not provided on the numbers of such excluded horses, it is difficult to calculate the true ‘success’ rate of published RLN treatments.

*Use of control or comparison populations:* A control population should be used when using racing performance to assess the efficacy of treatment, because this will limit unseen bias which may affect performance, such as concurrent illnesses and injuries, track conditions, managemental factors and weight carried. The ideal control population would consist of horses which have a diagnosis of RLN but which do not receive any treatment, however such cases are almost impossible to find. Several studies looking at URT disorders have used ‘comparison’ rather than control populations, which consist of horses presumed to not have RLN which are matched to cases by race or training yard\(^7,8,9,10,11\) however, the disease status of the comparison population is usually unknown, and this is therefore not ideal.

*Choice of outcome measure:* The quantification of ‘successful’ treatment of clinical cases is also a matter for debate. Whilst subjective methods such as the owners’ perception of improvement in clinical signs and performance may be potentially unreliable, the importance of the effect of treatment on owner/trainer satisfaction post-operatively should not be dismissed. 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 methods, subjective assessment usually involves a retrospective pre-/post- intervention assessment and memory decay and even a placebo effect may affect results\(^1\). For instance, Barnett et al.\(^12\) examined horses that underwent laryngoplasty post-operatively using over-ground endoscopy and found a high rate (78%) of dynamic upper respiratory abnormalities despite 93% of owners expressing that they were ‘satisfied’ or ‘very satisfied’ with the results of surgery. In a similar example for horses after DDSP surgery, McCluskie et al.\(^13\) showed that there was actually very little correlation between owners’ perception of success after surgery and whether DDSP still occurred during post-operative treadmill endoscopic examination, which is concerning!

From a scientific viewpoint, objective methods of analysis are therefore preferable; however there is no easy way to objectively measure success of upper airway surgery. Analysis of racing performance (usually using earnings, race starts or placings) pre- and post-treatment is the most common way of objectively analysing treatment success, but it is obvious that many factors other than URT abnormalities may affect a horse’s racing performance. It has recently been suggested that both earnings and number of starts should be analysed and that the effect of independent variables such as age, breed, gender and track surface should be incorporated in data analysis\(^14\).

Endoscopic evidence of arytenoid abduction is an alternative way of quantifying ‘success’. However, depending on the type of work performed by the horse, it could be argued that arytenoid stability during exercise rather than abduction grade at rest per se is the most important factor and that wide arytenoid abduction increases risk of other post-operative complications. Davidson et al.\(^15\) found that larynges with lower grades of abduction were more likely to be unstable than widely abducted arytenoids but that the presence of other disorders (right vocal fold collapse, ary-epiglottic fold collapse and DDSP) were very commonly seen in horses post laryngoplasty. These findings of concurrent ‘unrelated’ disorders have been echoed by 2 more recent studies, and have been shown in horses with all grades of post LP abduction\(^12,16\).
Take home message: In order to increase the level of evidence available to evaluate RLN treatments, researchers should make every effort to ensure that all patients that initially entered the trial are accounted for, that sample size is sufficient, inclusion criteria for cases and controls are rigorous, analysis of race records is performed using robust and repeatable methods and that a control population is evaluated if at all possible. Post-operative exercising endoscopy should be regarded as an integral part of analysis of treatment efficacy. When recommending a treatment for RLN, the level of evidence provided by published studies should be considered.

References:

