



Jämförelse mellan hältbedömning av häst utförd av 12 klinikveterinärer och en applikation baserad på datorseende

Comparison of 12 clinicians versus a computer vision application for lameness evaluation in horses

Julia Stenius

Självständigt arbete • 30 hp
Sveriges lantbruksuniversitet, SLU
Fakulteten för veterinärmedicin och husdjursvetenskap
Veterinärprogrammet
Uppsala 2022



Jämförelse mellan håltbedömning av häst utförd av 12 klinikveterinärer och en applikation baserad på datorseende

*Comparison of 12 clinicians versus a computer vision application
for lameness evaluation in horses*

Julia Stenius

Handledare: Elin Hernlund, Sveriges lantbruksuniversitet, Institutionen för anatomi, fysiologi och biokemi
Bitr. handledare: Anna Byström, Sveriges lantbruksuniversitet, Institutionen för anatomi, fysiologi och biokemi
Examinator: Marie Rhodin, Sveriges lantbruksuniversitet, Institutionen för anatomi, fysiologi och biokemi

Omfattning: 30 hp
Nivå och fördjupning: A2E
Kurstitel: Självständigt arbete i veterinärmedicin
Kurskod: EX0869
Program/utbildning: Veterinärprogrammet
Kursansvarig inst.: Institutionen för kliniska vetenskaper

Utgivningsort: Uppsala
Utgivningsår: 2022
Omslagsbild: Wilhelm Laska (med tillstånd)

Nyckelord: Hålt, rörelseanalys, rörelseasymmetri, håltbedömning, objektiv rörelseanalys, artificiell intelligens, datorseende, neurala nätverk, överensstämmelse

Sveriges lantbruksuniversitet
Fakulteten för veterinärmedicin och husdjursvetenskap
Institutionen för anatomi, fysiologi och biokemi

Publicering och arkivering

Godkända självständiga arbeten (examensarbeten) vid SLU publiceras elektroniskt. Som student äger du upphovsrätten till ditt arbete och behöver godkänna publiceringen. Om du kryssar i **JA**, så kommer fulltexten (pdf-filen) och metadata bli synliga och sökbara på internet. Om du kryssar i **NEJ**, kommer endast metadata och sammanfattning bli synliga och sökbara. Fulltexten kommer dock i samband med att dokumentet laddas upp arkiveras digitalt.

Om ni är fler än en person som skrivit arbetet så gäller krysset för alla författare, ni behöver alltså vara överens. Läs om SLU:s publiceringsavtal här:

<https://www.slu.se/site/bibliotek/publicera-och-analysera/registrera-och-publicera/avtal-for-publicering/>.

☒ JA, jag/vi ger härmed min/vår tillåtelse till att föreliggande arbete publiceras enligt SLU:s avtal om överlåtelse av rätt att publicera verk.

☐ NEJ, jag/vi ger inte min/vår tillåtelse att publicera fulltexten av föreliggande arbete. Arbetet laddas dock upp för arkivering och metadata och sammanfattning blir synliga och sökbara.

Sammanfattning

Ortopediska sjukdomar är hästens största medicinska problem världen över och sjukdomar i rörelseapparaten är den vanligaste anledningen till att svenska hästar behandlas av en veterinär. För framgångsrik behandling av ortopediska problem krävs i första hand en korrekt identifiering av det halta benet.

Tidigare publicerade studier har visat att subjektiv rörelseanalys utförd av veterinärer lider av flera brister. Överensstämmelsen mellan veterinärers subjektiva bedömningar av hälta är otillfredsställande, speciellt vid låggradiga och bakbenshältor, både med avseende på ben och grad. Bedömningen av diagnostiska anestesi kan påverkas av bias eller tolkas felaktigt. Objektiva metoder för att kvantifiera asymmetri har därför länge varit ett fokus för ortopedisk forskning och har idag utvecklats tillräckligt för att användas i praktiskt kliniskt arbete. Idag existerar ingen validerad skala för hältgradering i något av de objektiva systemen utan endast ett tröskelvärde som indikerar repeterbarhet mellan mätningar.

Denna studie jämförde subjektiv och objektiv rörelseanalys av 10 hästar i trav på trakt spår med avseende på primär och kompensatorisk hälta, hältgrad och hälttyp. Hästarna hade valts ut för att representera ett symmetriskt rörelsemönster samt en ökande grad av asymmetri jämnt fördelat mellan fram- och bakben. Lindrig till kraftig hälta hade verifierats med en positiv diagnostisk anes-tesi.

Den subjektiva analysen utfördes på video, av 12 veterinärer med i snitt 16,4 års erfarenhet inom hästortopedi. Den objektiva analysen utfördes av Sleip AI, en mobil applikation för markörlös optisk *motion capture*, en nyligen utvecklad AI-assisterad teknik för rörelseanalys. Applikationen använder datorseende för att identifiera hästens kroppsdelar och analyserar dess rörelsebanor med hjälp av maskininlärning. Veterinärerna besvarade även frågor om erfarenhet, säkerhet och användande av objektiv rörelseanalys i en enkät. Överensstämmelse analyserades med Fleiss' och Cohens kappa och tolkades som dålig $\kappa \leq 0,3$, acceptabel $\kappa = 0,31-0,5$, bra $\kappa = 0,51-0,8$ och mycket bra för $\kappa > 0,8$. Korrelation beräknades med Pearsons korrelationskoefficient (r) och Spearmans rangkorrelationskoefficient (r_s).

Överensstämmelsen inom veterinärgruppen skiljer sig i denna studie inte nämnvärt från tidigare studier. För primärhalt ben var samstämmigheten acceptabel ($\kappa = 0,45$), men för kompensatorisk halt ben ($\kappa = 0,26$) och gradering av primärhälta ($\kappa = 0,24$) var den dålig. Överensstämmelsen mellan veterinärgruppen och Sleip AI var bra ($\kappa = 0,63/0,75$). Veterinärer som var överens med Sleip AI med avseende på primärhalt ben var också mer samstämmiga med resten av veterinärgruppen. De var dessutom säkrare i sin bedömning och mer konsekventa i sin hältgradering. Varken tid i praxis, antal fall/år eller användande av objektiv rörelseanalys påverkade den enskilde veterinärens överensstämmelse med gruppen eller Sleip AI.

På grund av för liten datamängd samt dålig subjektiv samstämmighet kan denna studie inte presentera en fungerande modell för sambandet mellan subjektiv och objektiv gradering. För att i framtida studier utarbeta en fungerande modell bör varje bedömare gradera fler hästar, gärna samma häst flera gånger exempelvis före/efter diagnosisk anes-tesi.

Nyckelord: Hälta, rörelseanalys, rörelseasymmetri, hältbedömning, objektiv rörelseanalys, artificiell intelligens, datorseende, neurala nätverk, överensstämmelse

Abstract

Orthopedic disease is the most common equine health problem all over the world and injury to the locomotor system is the most common group of diagnoses in the Swedish horse population. To successfully treat orthopedic conditions, one must first correctly identify the lame leg.

Previous studies have highlighted several drawbacks of subjective gait analysis. Inter-observer agreement is unsatisfying regarding which leg is lame and grading of lameness. Agreement is especially poor for mild lameness and hindlimb lameness. Diagnostic analgesia could potentially be affected by expectation bias and also misinterpreted, producing false negatives and false positives. Objective gait assessment has therefore been a focus for orthopedic research for a long time but has now developed far enough to be used in clinical practice. There's no validated grading scale in any of the systems for quantitative gait analysis, only thresholds which indicate repeatability.

This study compared subjective and objective gait analysis of 10 horses in trot on a straight line regarding primary and compensatory lameness, grading and type of lameness. The horses were chosen to represent a symmetric gait pattern and an increasing grade of lameness, equally distributed on fore- and hind limbs. Mild to severe lameness was verified by a positive diagnostic analgesia.

Subjective analysis was performed on videos by 12 vets with a mean experience of 16.4 years in equine orthopedic practice. Objective analysis was performed by Sleip AI, a mobile application for marker less optical motion capture, a recently developed AI-assisted technology for equine gait analysis. The application uses computer vision to identify multiple body parts of the horse, track and analyze their trajectory with machine learning. The vets also answered a questionnaire regarding experience, confidence, and daily use of objective gait analysis. Inter-observer agreement was estimated by calculation of Fleiss' and Cohen's kappa (κ) and was considered poor for $\kappa \leq 0.3$, acceptable $\kappa = 0.31\text{--}0.5$, good $\kappa = 0.51\text{--}0.8$ and excellent $\kappa > 0.8$. Correlation was estimated with Pearson's correlation coefficient (r) and Spearman's rank correlation coefficient (r_s).

Agreement for primary lameness (which limb) was acceptable ($\kappa = 0.45$) but poor for compensatory lameness ($\kappa = 0.26$) and for grading of the primary lameness ($\kappa = 0.24$). Agreement between the vet group and Sleip AI was good ($\kappa = 0.63/0.75$). Individual vets in good agreement with Sleip AI regarding primary lameness was also more likely to agree with the vet group. They were also more confident in their judgement and more consistent in their grading. Neither time in praxis, caseload or use of objective gait analysis affected the individual vet's agreement with the group or Sleip AI.

Due to a small data set and poor subjective agreement, this study is unable to present a reliable model for correlation between subjective and objective lameness grading. To develop a reliable model each assessor needs to grade more horses, same horse multiple times, preferably before/after diagnostic analgesia.

Keywords: Lameness, gait analysis, gait asymmetry, lameness evaluation, objective gait analysis, agreement, inter-rater

Innehållsförteckning

Förkortningar	9
1. Inledning	11
2. Litteraturöversikt	13
2.1. Hälta	13
2.1.1. Vanligt och viktigt	13
2.1.2. Vad är hälta?	14
2.1.3. Hur går en hältutredning till?	14
2.1.4. Att analysera hästens rörelsemönster	15
2.1.5. Att gradera hälta subjektivt	16
2.1.6. Svårigheter med subjektiv rörelseundersökning	17
2.2. Objektiv rörelseanalys	19
2.2.1. Att mäta hälta är ingen nyhet	19
2.2.2. Framtiden för kvantitativ hältbedömning	19
2.2.3. Var mäter man rörelse?	20
2.2.4. Tekniker för objektiv hältanalys	21
2.2.5. Kinetik	21
2.2.6. Kinematik	21
2.2.7. Gradering i objektiva system	23
2.2.8. Jämförelse mellan subjektiv och objektiv rörelseanalys	24
2.2.9. Artificiell intelligens tolkar objektiva rörelsedata	25
2.2.10. Hälta versus naturlig lateralitet	26
3. Material och metod	27
3.1.1. Sleip AI	27
3.1.2. Hästarna	28
3.1.3. Deltagarna	31
3.1.4. Undersökningen	31
3.1.5. Statistiska beräkningar	32
3.1.6. Jävsdeklaration	33
4. Resultat	34
4.1. Deltagarna	34
4.2. Primärhälta	34

4.2.1.	Överensstämmelse ben	35
4.2.2.	Hältgradering	39
4.2.3.	Hälttyp	43
4.3.	Kompensatorisk hälta	43
4.4.	Deltagarnas egenskaper	44
5.	Diskussion.....	45
5.1.1.	Deltagarna	45
5.1.2.	Hästarna	45
5.1.3.	Undersökningen.....	47
5.1.4.	Video	47
5.1.5.	Ljud	48
5.1.6.	Statistik	49
5.1.7.	Primärhälta.....	49
5.1.8.	Kompensatorisk hälta	52
5.1.9.	Hältgradering	52
5.1.10.	Användande av objektiv rörelseanalys	56
5.1.11.	Konklusion	58
	Referenser.....	59
	Tack	68
	Populärvetenskaplig sammanfattning	69
	Bilaga 1.....	71
Häst 1		72
Häst 2		73
Häst 3		74
Häst 4		75
Häst 5		76
Häst 6		77
Häst 7		78
Häst 8		79
Häst 9		80
Häst 10		81

Förkortningar

AAEP	American Association of Equine Practitioners
AI	Artificiell intelligens
B	Belastningshålla
F	Frånskjutshålla
HDmax	<i>Head difference maximum</i> (skillnad i huvudets högsta punkt)
HDmin	<i>Head difference minimum</i> (skillnad i huvudets lägsta punkt)
HB	Höger bakben
HF	Höger framben
IMUs	<i>Inertial measurement units</i>
MaxDiff	Maximumdifferens
MinDiff	Minimumdifferens
ML	Maskininlärning
NRS	<i>Numeric rating scale</i> (numerisk graderingsskala)
PDmax	<i>Pelvis difference maximum</i> (skillnad i bäckenets högsta punkt)
PDmin	<i>Pelvis difference minimum</i> (skillnad i bäckenets lägsta punkt)
RUD	<i>Range up difference</i> (skillnad i den vertikala uppåtrörelsen)
VB	Vänster bakben
VF	Vänster framben
VRS	<i>Verbal rating scale</i> (verbal graderingsskala)

1. Inledning

Ortopediska sjukdomar är hästens största medicinska problem världen över (Priester 1970; Clausen *et al.* 1990; Kaneene *et al.* 1997; Penell *et al.* 2005) och sju av de tio vanligaste diagnoserna i Sverige är relaterade till problem i rörelseapparaten (Agria Djurförsäkring 2021). Huvudorsaken till att hästar har nedsatt prestation, uteblir från tävling och träning och slutligen avlivas är ortopediska problem (Morris & Seeherman 1991; Egenvall *et al.* 2006, 2013). Hästveterinärer lägger merparten av sin arbetstid på hältutredningar (Loomans *et al.* 2007).

För framgångsrik behandling av en hälta krävs först att någon upptäcker att hästen är halt, vilket har visat sig svårt för ägare, ryttare och tränare (Mullard *et al.* 2017). Vidare måste veterinären korrekt identifiera det halta benet och för bästa behandling även ställa en så specifik diagnos som möjligt (Buchner 2013).

Tyvärr är hältbedömning svårt även för veterinärerna. Vid låggradiga hältor speciellt härrörande från bakbenen är samstämmigheten i veterinärernas visuella bedömningar bristfällig (Fuller *et al.* 2006; Keegan *et al.* 2010, 2013; Hammarberg *et al.* 2016). Diagnostiska anestesier har visats påverkas av bias (Arkell *et al.* 2006) eller misstolkas vilket genererar både falskt positiva och negativa resultat (Schumacher *et al.* 2014). Dessa nedslående men viktiga insikter kan förklaras av mänskliga faktorer av både fysiologisk och psykologisk art (Parkes *et al.* 2009; Starke & May 2017).

För att undvika den inneboende subjektiviteten i visuell hältbedömning har objektiva metoder att kvantifiera hälta länge varit fokus för ortopedisk forskning och idag används dessa metoder även i kliniken (Serra Bragança *et al.* 2018). Överensstämmelsen mellan objektiv och subjektiv rörelseanalys varierar mellan dålig (Keegan *et al.* 2013), acceptabel (Keegan *et al.* 2013; Bragança *et al.* 2020), och måttlig (Keegan *et al.* 2013), men ingen av dessa studier har inkluderat fler än fyra subjektiva bedömare. Idag existerar ingen validerad skala för kvantitativ hältgradering utan endast ett tröskelvärde som indikerar repeterbarhet i rörelseasymmetri (Keegan *et al.* 2010).

Syftet med denna studie är att jämföra subjektiv rörelseanalys utförd av 12 hästveterinärer med ett markörlöst optiskt *motion capture*-system (Sleip AI). Sleip AI

är en nyligen utvecklad mobilapplikation, som med hjälp av datorseende och neurala nätverk analyserar symmetrin i en hästs rörelse.

Studien kommer att jämföra bedömning av primärhalt ben, hältgradering samt hälttyp med syfte att undersöka samstämmighet såväl mellan veterinärerna som mellan Sleip AI och veterinärerna. Ett försök att ta fram en modell för hur subjektiv gradering korresponderar med Sleip AI kommer utföras.

2. Litteraturöversikt

2.1. Hälta

2.1.1. Vanligt och viktigt

En häst värderas och används idag i stor utsträckning baserat på dess rörelseapparats kapacitet och funktion, trots att dess primära användningsområden under 1900-talet drastiskt har skiftat från jordbruk, armé och transport till sport och hobby är det fortfarande hästens rörelseförmåga som är dess viktigaste egenskap (Serra Bragança *et al.* 2018).

Störningar i rörelseapparaten, vilket ses som hälta, är idag svenska hästars största medicinska problem och sju av de tio vanligaste diagnoserna är relaterade till problem i rörelseapparaten (Agria Djurförsäkring 2021). Även historiskt sett är hälta och störningar i rörelseapparaten den vanligaste sjukdomsdiagnosen såväl som orsak till avlivning av hästar i och utanför Sverige (Priester 1970; Clausen *et al.* 1990; Kaneene *et al.* 1997; Penell *et al.* 2005; Egenvall *et al.* 2006).

I en longitudinell studie som kartlade varför hopphästar på elitnivå uteblev från träning och tävling, enligt konceptet ”days – lost to training”, var huvudorsaken ortopediska problem och 77 % av de förlorade dagarna kunde tillskrivas akuta och icke-akuta skador i rörelseapparaten (Egenvall *et al.* 2013). Hästar med tidigare kända ortopediska problem hade även en fördubblad incidensrat ratio för att förlora aktiva dagar på grund av ortopedisk skada. I en studie av aktivt tävlande trav- och galopphästar med nedsatt prestation, kunde man härleda prestationsnedsättningen helt eller delvis till muskuloskeletala problem i 204 fall av totalt 275 (Morris & Seeherman 1991).

Trots att hälta är vanligt är låggradiga asymmetrier ofta svåra att upptäcka för ägaren, ryttaren, tränare och mindre erfarna veterinärer (Mullard *et al.* 2017). Det är inte ovanligt att aktiva sporthästar, som av sina ägare anses ohalta, i själva verket är halta (Greve & Dyson 2014). I en studie med 506 sporthästar i normalt arbete där ägaren ansåg hästen vara frisk, konstaterades 47 % av hästarna halta eller ha smärtrelaterade rörelsestörningar. I en annan studie av samma författare genomgick 57 stycken aktiva sporthästar en fullständig hältutredning med trav på rakt spår, longering på hårt och mjukt underlag, ridprov samt böjprov och endast 25 % bedömdes som ohalta (Dyson & Greve 2016). Liknande resultat beskrevs i ytterligare en studie där objektiv rörelseanalys användes för att utvärdera friska hästar i trav på rakt spår (Rhodin *et al.* 2017). Författarna visade att 72.5 % av hästarna hade ett rörelsemönster som låg över kliniskt använda gränsvärden för hälta.

2.1.2. Vad är hälta?

Definitionen av hälta har fram till introduktionen av objektiv rörelseanalys varit ganska enkel och inte förändrats nämnvärt på över ett århundrade, då Liautard (1888) beskrev att hälta är en manifestation av en föregående skada i en eller flera delar av rörelseapparaten och inte en ursprunglig ondska eller en sjukdom i sig. En mer nutida beskrivning är att hälta endast är en klinisk bild – en manifestation av tecknen på inflammation, så som smärta, eller en mekanisk defekt – vilket resulterar i en rörelsestörning, som uttrycks i en linkande gång (Ross 2011a). Studierna som med objektiva mätmetoder visat att stora delar av hästpopulationen har ett asymmetriskt rörelsemönster, har startat en livlig debatt om definitionen av hälta (Weeren *et al.* 2017; Adair *et al.* 2018, 2019; Bathe *et al.* 2018). Sammanfattningsvis engagerar frågan huruvida en *linkande gång* är ett tecken på smärta och patologi eller om ett asymmetriskt rörelsemönster kan vara en normalvariation, vilket förhoppningsvis leder till mer forskning inom området.

2.1.3. Hur går en hältutredning till?

En hältutredning syftar till att i första hand lokalisera det skadade området och därefter ställa en så specifik diagnos som möjligt och slutligen erbjuda en behandling baserat på denna diagnos (Buchner 2013). Hältutredningar är det som hästveterinärer lägger merparten av sin arbetstid på (Loomans *et al.* 2007).

Normala komponenter i en hältutredning är anamnestagning, visuell bedömning av konformation, symmetri och hållning, palpation av rörelseapparaten,

rörelseundersökning i skritt och trav på rakt och böjt spår på olika underlag, provokationstester så som böjprov och även diagnostiska analgesier och bilddiagnostik (Ross 2011a). Undersökningen följs ofta av behandling och upprättande av en rehabiliteringsplan och därefter ett eller flera återbesök. Vissa anser att ridprov kan vara avgörande i en hältundersökning och beskrev i en studie att 10–12,3 % av hästarna vid ridprov uppvisade hälta, vilken ej kunde ses för ögat vid andra delar av rörelseundersökningen (Dyson 2016; Dyson & Greve 2016). Andra beskriver dynamisk utvärdering på höghastighetsrullband som en viktig komponent vid subtila diagnoser hos prestationshästar (Morris & Seeherman 1991). Sammanfattningsvis är det viktigt att utredningen utförs systematiskt och steg för steg, på ett sätt som veterinären är väl förtrogen med (Ross 2011a). Naturligtvis måste flexibilitet finnas då flera faktorer, som exempelvis hästens hanterbarhet, misstanke om fraktur och djurägarens ekonomi kan förändra eller begränsa undersökningen.

2.1.4. Att analysera hästens rörelsemönster

Trav är bästa gångarten för att utvärdera om och från vilket ben en häst är halt. Traven är en tvåtaktig gångart där det diagonala benparet förflyttas och bär vikt simultant (Ross 2011b). I normal hastighet är traven symmetrisk och båda halvorna av stegcykeln är identiska. Stegcykeln delas upp i en belastningsfas, där höger framben och vänster bakben respektive vänster framben och höger bakben belastas. Därefter följer en frånskjutsfas, då den horisontella kraften från hoven mot marken riktas bakåt. Cykeln avslutas med en pendlingsfas, vilken i sitt slutskede inkluderar svävfasen, då alla benen är i luften (Barrey 1999).

Under stegcykeln rör sig hästens huvud respektive bäcken vertikalt i en bifasisk sinusliknande kurva, där de två högsta punkterna är i slutet av belastningsfasen eller under svävfasen och de lägsta punkterna mitt under belastningsfas för respektive diagonalt benpar (Buchner *et al.* 1996). Vid en hälta anpassar hästen rörelsemönstret för att avlasta berört ben (Weishaupt *et al.* 2010). Detta åstadkommer den genom att omfördela vikt till det andra benet i diagonalen samt det kontralaterala benet. Resultatet blir en reduktion av amplitud i huvudets respektive bäckenets vertikala rörelse för det halta benet (Rhodin *et al.* 2017).

Under det halta benets belastningsfas reduceras den vertikala amplituden för huvud/bäcken i förhållande till det friska benet (Keegan 2007). Huvud/bäcken sänks alltså inte lika mycket när detta ben belastas och får således en högre minimumposition i förhållande till när det friska benet belastas. Skillnaden mellan de två sidornas minimumpositioner ger en minimumdifferens (MinDiff) för antingen huvud (HDmin) eller bäcken (PDmin), vilket alltså indikerar en belastningshälta för fram- respektive bakben. Rent visuellt upplevs och beskrivs denna rörelse ofta som *down on sound* (ned på friskt) eller att hästen nickar ner på det ohalta benet.

Likaledes skjuter det halta benet inte huvud/bäcken lika högt som det friska benet, vilket gör att maxpositionen blir lägre (Buchner *et al.* 1996; Kramer *et al.* 2004). Skillnaden mellan höger och vänster fram- respektive bakben blir en maximumdifferens (MaxDiff) för antingen huvud (HDmax) eller bäcken (PDmax), vilket indikerar en frånsjutshälta.

Kompensatorisk hälta

När man försöker analysera en hästs rörelse måste man vara medveten om att den har flera strategier för att avlasta det smärtande benet, vilket kan göra att den ser ut att vara halt på ett ben som egentligen är friskt, en så kallad kompensatorisk hälta (Kelmer *et al.* 2005). Per definition är dessa hältor inte smärtutlösta och ska försvinna vid positiv bedövning av den primära hältan (Maliye *et al.* 2015). Flera studier visar att en ipsilateral hälta, exempelvis höger framben (HF) och höger

bakben (HB), har sitt ursprung i bakbenet (May & Wyn-Jones 1987; Keegan 2007; Rhodin *et al.* 2013). Något mer komplicerat verkar det vara vid en primär frambenshätta, där man vid inducerad hälla, har beskrivit kompensatoriska mönster för båda bakbenen (Buchner *et al.* 1996; Weishaupt *et al.* 2006; Keegan 2007). Detta förklaras delvis av att man faktiskt mätt olika saker (Kelmer *et al.* 2005). Bedömer man vertikal uppåtrörelse för exempelvis *sacrum* minskar den för det diagonala bakbenet, vilket tolkas som en frånskjutshälla (PDmax). Om man istället mäter belastning (Weishaupt *et al.* 2006), ses en ökning på det diagonala bakbenet och således en minskning för det samsidiga bakbenet, vilket tolkas som en belastningshälla. PDmin indikerar då en samsidig belastningsbakhälla. Generellt är den kompensatoriska huvudnickningen vid en primär bakbenshälla mer uttalad än bäckenets kompensatoriska rörelser vid en frambenshälla (Rhodin *et al.* 2013).

2.1.5. Att gradera hälla subjektivt

Att gradera hälltan gör att veterinären mer exakt kan dokumentera, jämföra över tid eller före respektive efter en bedövning eller behandling samt kommunicera med kollegor eller djurägare (Dyson 2011). Idag används två olika typer av skalor; verbal ratingskala (VRS) och numerisk ratingskala (NRS) (Hewetson *et al.* 2006). I den verbala skalan används adjektiv för att beskriva hästens rörelser vid olika nivåer av asymmetri. Allvarlighetsgraden är rangordnad från mild till kraftig och graderad i enlighet med detta. I en numerisk skala graderas hälltan vanligtvis från 0–5 eller 0–10, där 0 betyder att hästen är ohalt och 5 eller 10 indikerar att hästen är så halt som den någonsin kan bli.

Den enda skalan som anses officiell är utvecklad av American Association of Equine Practitioners (AAEP) (Anon 1999). Den är en VRS och graderar hälla 0–5. Både Ross (2011b) och Stashak (1987) har publicerat egna VRS med gradering 0–5 respektive 0–4, där de inkluderar en mer detaljrik beskrivning av varje hälltgrad och där den senare även beaktar ljudet vid hovisättningen.

Det har varit vanligt att använda en NRS med gradering 0–10 i Storbritannien (Wyn-Jones 1988) men i strävan efter en universal skala (Ross 2011b), verkar denna inte användas i så stor utsträckning längre, i alla fall inte hela vidden av skalan (Dyson 2011). Dyson själv använder en numerisk skala med gradering 0–8 (0 = ohalt; 2 = lindrig; 4 = måttlig; 6 = kraftig; 8 = icke vitktbärande), vilken hon applicerar i skritt, trav, rakt ut, under longering och vid ridprov. Hon adderar även en beskrivning av rörelsemönstret och tycker att systemet ger en större flexibilitet än AAEP-skalan, men är tydligare än skalan 0–10.

I en besiktningsmanual utgiven av Sveriges Veterinärförbund instrueras veterinärer att gradera en hälta 0–5 (Roethlisberger Holm 2012), men exakt enligt vilken skala är ej definierat.

2.1.6. Svårigheter med subjektiv rörelseundersökning

Det är allmänt erkänt att erfarna veterinärer kan upptäcka och diagnosticera hältor vid subjektiv bedömning men flera studier har pekat på brister i olika aspekter av dessa bedömningar (Dyson 2014).

Dålig överensstämmelse

Flera studier har visat att både studenter, erfarna klinikveterinärer och ortopediska experter har en god överensstämmelse *intra-observer*, det vill säga att de identifierar och graderar en hälta konsekvent när de exponeras för den vid olika tillfällen (Arkell *et al.* 2006; Fuller *et al.* 2006). När man istället jämför bedömarnas resultat med varandra, framförallt för låggradiga hältor, speciellt härrörande från bakbenen, är de inte överens om på vilket ben hästen är halt, vilket ses som låg *inter-observer agreement* (Fuller *et al.* 2006; Keegan *et al.* 2010, 2013; Hammarberg *et al.* 2016).

I en studie bedömde 2–5 kliniker med i snitt 18,7 års ortopedisk erfarenhet 131 hästar (Keegan *et al.* 2010). När de skulle avgöra huruvida ett ben var halt eller inte efter att enbart ha sett hästen i trav på rakt spår var den generella överensstämmelsen acceptabel ($\kappa = 0,44$). För hästar med en hälta $>1,5$ grader enligt AAEP var överensstämmelsen mycket bra ($\kappa = 0,86$) men för hältor graderade $<1,5$ var den dålig ($\kappa = 0,23$), vilket i praktiken betyder att veterinärerna är överens huruvida ett specifikt ben var halt i lite mindre än 2 av 3 fall (61 %). För framben var överensstämmelsen bra ($\kappa = 0,51$) och för bakben acceptabel ($\kappa = 0,36$). I samma studie var samstämmigheten fortsatt acceptabel ($\kappa = 0,37$) då klinikveterinärerna skulle avgöra om hästen var halt och om så var fallet, vilket ben de skulle börja utreda efter trav på rakt spår, longering och böjprov.

Liknande resultat ses med en studie som inbegrep 20 hästar, bedömda av tre erfarna klinikveterinärer (Fuller *et al.* 2006). Veterinärerna fick se hästarna rakt ut i skritt och trav samt longerade och graderade därefter hältor oavsett ben, på en skala 0–10. Överensstämmelsen var acceptabel ($\kappa = 0,41$).

I en annan studie bedömde 86 veterinärer videofilmer av 23 lindrigt halt och ohalter hästar under longering (Hammarberg *et al.* 2016). I frågan vilket ben de skulle börja utreda var överensstämmelsen för alla hästar precis acceptabel ($\kappa = 0,31$), vilket även gällde för framben ($\kappa = 0,33$) men för bakben var den dålig ($\kappa = 0,11$). När de skulle särskilja halt hästar från ohalter var samstämmigheten ännu sämre ($\kappa = 0,08$).

Konsensus verkar inte påverkas nämnvärt av huruvida man bedömer en verklig eller en animerad häst (Starke & Oosterlinck 2019), om hältan är inducerad eller naturlig (Hammarberg *et al.* 2016), eller om det görs på video eller i verkligheten (Fuller *et al.* 2006). Hur hästen bedöms – enbart i trav på rakt spår (Keegan *et al.* 2010), longerad (Hammarberg *et al.* 2016) eller i en fullständig hältundersökning (Keegan *et al.* 2010) – verkar inte heller ha någon egentlig betydelse för resultaten.

Forskare har försökt förklara dessa nedslående resultat på olika sätt. Dyson (2011, 2014) härrör den dåliga överensstämmelsen till otillräcklig träning i vad man ska titta efter, brist på övning i att applicera gradering av hältor samt variansen i graderingsschema, inkonsekvent dokumentering och slutligen den inneboende subjektiviteten i dessa bedömningar.

Synens begränsning

Oförmågan att rent visuellt detektera subtil häla kan även förklaras av fysiologiska faktorer i det mänskliga synsinnet (Keegan *et al.* 2010). Ögat ser med en tidsupplösning på 10–15 rutor/sekund, vilket uppskattats vara under eller precis på gränsen för att kunna upptäcka asymmetri i rörelserna som avslöjar en häla i trav i hastigheten 4 m/s.

Vidare konstaterar Parkes *et al.* (2009) att förmågan att uppfatta subtila asymmetrier i rörliga objekt hos den mänskliga synen är begränsad. Både för att ögat har svårt att känna igen små förändringar i acceleration (Haarmeier & Thier 2006) och för att det tar längre tid att uppfatta en subtil asymmetri än en uppenbar sådan (Huang *et al.* 2004). Tröskeln för att upptäcka rörelseasymmetri är en avvikelse på ungefär 25 % för framben och 15 % för bakben (Starke & Oosterlinck 2019). I studien av Parkes *et al.* (2009) tittade en grupp erfarna kliniker och en grupp oerfarna studenter på två objekt, representerande *tuber coxae* (vars rörelse kan användas för att detektera bakbenshälta). Den erfarna gruppen var signifikant bättre på att identifiera asymmetri baserad på bäckenets rörelsemönster hos en halt häst, men vid ett artificiellt rörelsemönster sågs ingen skillnad mellan grupperna. Resultatet tyder dock på att man kan träna sig bättre på att se bakbenshältor.

Bias och tolkningssvårigheter vid diagnostisk bedövning

Vidare har det visats svårt att undvika bias, då bedömarens föregående vetskap om att hästen fått en diagnostisk ledningsanestesi gör att den bedömer hästen som mindre halt (Arkell *et al.* 2006). Skillnaden i bedömd hältgrad före och efter bedövning ökade med 0,4 av 10 när bedömarens visste att en nervblockad administrerats och fick se filmerna i kronologisk ordning. Det förekommer även att utslaget

av diagnostiska anestasier misstolkas vilket genererar både falskt positiva och falskt negativa resultat (Schumacher *et al.* 2014).

Sammanfattningsvis skulle man kunna säga att en subjektiv bedömning av rörelse är snabb, enkel och billig men begränsas av låg validitet, sensitivitet och specificitet (Dolatabadi *et al.* 2017).

2.2. Objektiv rörelseanalys

2.2.1. Att mäta hälta är ingen nyhet

Tidigare nämnda problem vid subjektiv rörelseanalys komplicerar hältutredningar och utgör *confounding factors* vid kliniska beslut, vilket i sin tur hämmar ortopedisk forskning vad gäller utvärdering av diagnostiska procedurer, behandlingar och rehabiliteringsprotokoll (Serra Bragança *et al.* 2018).

Objektiva metoder har längre varit fokus för ortopedisk forskning. Ändå sen man uppfann foto, film och video har man använt dessa för att bättre förstå hästens gångarter (van Weeren 2013). Redan för 50 år sen publicerade Fredricson & Drevemo (1971) en studie där man använde teknik som tidigare används i flygplansforskning. Genom att gjuta specialanpassade glasfiberskor med fyra röda bollar, vilkas rörelse kunde filmas och analyseras, lade man en grund för detaljerad objektiv rörelseanalys.

Studier har visat att objektiva system är känsligare än visuell bedömning och alltså bättre på att detektera en mild hälta (McCracken *et al.* 2012). Bedömningarna är dessutom repeterbara (Keegan *et al.* 2011), vilket är avgörande vid tolkning av upprepade mätningar vid exempelvis diagnostiska anestasier och återbesök. Det har även visats att kvantitativ rörelseanalys fungerar bra för att utvärdera diagnostiska anestasier både kliniskt (Maliye *et al.* 2013, 2015; Maliye & Marshall 2016) och experimentellt (Tóth *et al.* 2014; Hoerdemann *et al.* 2017).

2.2.2. Framtiden för kvantitativ hältbedömning

En reviewartikel konkluderar att det nu finns tillräckligt med evidens för att stödja användning av objektiv rörelseanalys i forskning och klinisk verksamhet (Serra Bragança *et al.* 2018). Den kinematiska analystekniken har utvecklats tillräckligt vad gäller praktisk användbarhet och pålitlighet och därav anses det motiverat att inkludera objektiv rörelseanalys i det dagliga arbetet.

I ett konsensusdokument jämför fyra framstående forskargrupper inom kvantitativ rörelseanalys introduktionen av objektiv bedömning med intåget av avancerad bilddiagnostik och förutser att det likaledes kommer att bli ett uppskattat och outhärligt verktyg för att fatta välgrundade och evidensbaserade kliniska beslut (Weeren *et al.* 2017). De poängterar dock är det helt avgörande att tekniken används som ytterligare ett diagnostiskt verktyg och inte som ett ofelbart facit i det kliniska arbetet.

Man kan förvänta sig att objektiv rörelseanalys kommer att bli standardförfarande för långtidsmonitorering av friska och halta hästar, då variationen över tid är liten för varje individ (Hardeman *et al.* 2019), samt för inhämtande av exakta och entydiga rörelsedata (Serra Bragança *et al.* 2018).

2.2.3. Var mäter man rörelse?

För att objektivt detektera och kvantifiera hälta finns en uppsjö potentiella mätvärden, exempelvis rörelse, acceleration och hastighet (Keegan 2007). För att göra de kvantitativa mätningarna så intuitiva och kliniskt relevanta som möjligt bör mätvärden och referenspunkter som liknar de som används för visuell subjektiv bedömning användas. Därför har man de senaste tjugo senaste åren fokuserat mest på den vertikala rörelsen hos olika kroppsdelar (Serra Bragança *et al.* 2018).

För frambenshälta används idag främst skillnaden mellan huvudets max- och minposition (HDmax/min) för höger och vänster ben. Även mankens vertikala rörelse och huvudets vertikala acceleration är bra indikatorer för frambenshälta (Buchner *et al.* 1996) men har idag i stort sätt övergivits som enskilda mått, då de kan vara svåra att uppfatta med ögat samt skiljer sig från vad man tittar på i en vanlig visuell bedömning och således blir mindre intuitiva och svårare att förstå (Serra Bragança *et al.* 2018).

För bakbenshälta används skillnaden i *tubera sacrales* max- och minposition för höger och vänster ben, vilket ger PDmax och PDmin (Buchner *et al.* 1996). Uppgående rörelse av *tuber coxae*, så kallad *pelvic hike*, är det mest känsliga mätvärdet för utvärdering av diagnostisk analgesi av bakben, men totalt rörelseomfång för *tuber coxae* eller PDmin för *tubera sacrale* fungerar i princip lika bra (Pfau *et al.* 2014). För att maximera chansen att diagnosticera milda hältor, speciellt vid subjektiv bedömning, rekommenderas att veterinären utvärderar både vertikal amplitudskillnad för *tubera sacrale* och *tuber coxae*, då hästars individuella rörelsemönster kan göra respektive mätvärde mer eller mindre känsligt (Starke *et al.* 2015).

Man har nyligen föreslagit att specifika kombinationer av mätvärden, så kallade asymmetrimönster, är bättre på att identifiera hälta än enstaka parametrar (Harde-
man *et al.* 2021). Exempelvis kunde en kombination av *Range up diff* (RUD) huvud
+ Min diff manke – RUD bäcken, ett så kallat *forelimb lame pattern*, skilja mellan
hästar med en grav och en lindrig anmärkning på besiktning.

2.2.4. Tekniker för objektiv hältanalys

Inom fältet objektiv hältanalys idag används två tekniker för att kvantifiera hästens
rörelsemönster, antingen via krafter (kinetik) eller rörelse (kinematik) (Serra
Bragança *et al.* 2018).

2.2.5. Kinetik

I kinetiska studier analyseras krafter skapade av hästens rörelseapparat. Stationära
kraftplattor var en av de första beskrivna teknikerna (Morris & Seeherman 1987;
Aviad 1988; Merckens & Schamhardt 1988) och anses fortfarande vara gold
standard för att mäta belastningsasymmetrier inom kinetiken. Praktiska begräns-
ningar i form av komplexitet i datainsamling och analys gör att denna teknik idag
endast används inom forskningsfältet (Serra Bragança *et al.* 2018).

2.2.6. Kinematik

I kinematiska studier studerar man kroppsdelars rörelse under förflyttning.
Kroppsdelarnas rörelse beskrivs i termer av förflyttning/hastighet/acceleration i
relation till tid eller position i förhållande till varandra eller i ett koordinatsystem
(Serra Bragança *et al.* 2018).

Optisk motion capture – markörbaserad

Idag är tredimensionell markörbaserad *motion capture* gold standard för kinematisk
analys (Serra Bragança *et al.* 2018). Tekniken kan mycket precist analysera rörelse
med hjälp av infraröda höghastighetskameror, vilka simultant registrerar positionen
av reflektiva markeringar som placerats på hästen. Oqus/Qhorse (Qualisys AB),
Vantage (Vicon) och Motion Analysis (Motion Analysis Systems) de mest använda
systemen. Mängden kameror samt kalibrering och service av dessa gör dessa sys-
tem dyra och stationära.

Optisk motion capture – markörlös

Under de senaste åren har man även utvecklat markörlös optisk *motion capture*, där
man använder datorseende för att identifiera kroppsdelar och sen analyserar dess

rörelsebanor med hjälp av maskininlärning (ML), vilket är en typ av artificiell intelligens (AI) (Takeda *et al.* 2021). Datorseende är ett tvärvetenskapligt fält som handlar om hur datorer kan avkoda och förstå bilder och video (Wikipedia 2021). Fältet strävar efter att på digital väg registrera, automatisera och förbättra vad det mänskliga synsystemet naturligt kan åstadkomma. ML är populärt inom human rörelseanalys eftersom det ger en snabb, exakt och robust klassificering av biomekaniska data i förhållande till tid (Khera & Kumar 2020).

I en experimentell studie i kvantitativ human rörelseanalys konstaterade Takeda *et al.* (2021) att AI-assisterad markörlös optisk *motion capture* skulle kunna vara en adekvat ersättare till den markörbaserade tekniken, vilken idag är den mest använda. Den markörlösa tekniken kunde, utan att kompromissa med exakthet, analysera patientens rörelse i deras hem eller där de utför sin rehabilitering, vilket skulle kunna reducera kostnader och göra mätningarna enklare.

I en annan studie tränade Wang *et al.* (2021) ett neuralt nätverk att på video identifiera en hästs olika kroppsdelar och därefter analysera om hästen är halt baserat på huvudets och bäckenets vertikala rörelse. Hästarna var filmade med mobilkamera och man använde inga sensorer eller markeringar på hästarna. Man utvecklade även en applikation för att testa om tekniken var praktiskt användbar för att detektera hälta hos hästar. Resultatet visar att tekniken kunde avgöra om en häst är halt eller inte med 91 % säkerhet, men graden av hälta rapporterades inte.

Applikationen som användes i denna studie (Sleip AI) använder markörlös optisk *motion capture*. Mer information om Sleip AI under *Material och metod*.

Inertial measurement units (IMUs)

Ungefär samtidigt som markörbaserad optisk *motion capture* började användas i större skala, lanserades en teknik med tröghetssensorer, så kallade inertial measurement units (IMUs), monterade på hästens kropp (Serra Bragança *et al.* 2018). Sensorerna registrerar acceleration, vilken dubbelintegreras för att generera en uppskattning av kroppsdelens position och möjliggör rörelseanalys utan större krav på den omgivande miljön. Några av de vanligaste systemen är Lameness Locator, Equigait och EquiMoves och de används i forskning, klinisk verksamhet och undervisning världen över. Denna teknik är mindre exakt än markörbaserad optisk *motion capture* vad gäller kroppsdelars position men mer precis och pålitlig vad gäller acceleration. Systemen är lätta att använda, mobila och förhållandevis billiga, vilket gör att det har potential för vardagligt kliniskt arbete

2.2.7. Gradering i objektiva system

För att tolka resultaten från de objektiva systemen har man tagit fram tröskelvärden för vad som anses klassificera en häst som halt eller ohalt baserat på repeterbarhet (Keegan *et al.* 2012, 2013), eller symmetrisk respektive asymmetrisk, som vissa hellre benämner det (Weeren *et al.* 2017). För IMU-systemet Lameness Locator är gränsen för frambenshälta ± 6 mm för HDmin/HDmax och för bakbenshälta ± 3 mm för PDmin/PDmax (McCracken *et al.* 2012). Det är värt att notera att detta inte rör sig om faktiska millimetrar, och att varje häst för detta IMU-system får ett mätvärde som relateras till dess eget vertikala rörelseomfång. Utöver detta tröskelvärde finns idag ingen annan uniform objektiv graderingsskala. Flera studier har dock rapporterat att en stor andel av den friska hästpopulationen ligger över dessa gränsvärden (Rhodin *et al.* 2017; Hardeman *et al.* 2019; Persson-Sjodin *et al.* 2019; Müller-Quirin *et al.* 2020) och Hardeman *et al.* (2021) visade nyligen att hästar med hälta av 1–2 grader låg långt över dessa värden och redovisade medelvärden för HDmin på 22,2 mm, HDmax 14,4 mm, PDmin 17,9 mm och PDmax 18,6 mm, vilket också delvis kan förklaras av att Hardeman har mätt med ett annat system (QHorse).

På grund av den stora variabiliteten mellan hästar, menar somliga forskare att det troligen aldrig kommer att vara möjligt att ta fram universella tröskelvärden (Hardeman *et al.* 2019). Andra ser en möjlighet att kunna jämföra analyser från olika system, då man i två studier som jämfört dels två system med IMUs, dels markörbaserad optisk *motion capture* med IMUs, kommer fram till att dessa tröskelvärden korresponderar väl med ”Widths of limits of agreement” mellan olika system (Warner *et al.* 2010; Pfau *et al.* 2016). En annan studie föreslår att man bör korrigera mätvärdena för det vertikala rörelseomfånget för varje individ, så kallad *range of motion* (ROM) för att på bästa sätt kunna jämföra mellan olika system och på sikt kunna utveckla referensvärden som är användbara i olika system (Hardeman *et al.* 2021). I en, för dagen opublicerad, jämförande studie av markörbaserad (QHorse) och markörlös optisk *motion capture* (Sleip AI) sågs endast små skillnader mellan systemen (Hernlund *et al.*, opubl. accepterat abstract, 2022). 20 hästar mättes i trav på rakt spår och då standardavvikelsen i varje försök var 9–25 mm för den markörbaserade tekniken och 9–30 mm för den markörlösa ansåg författarna att den låga nivån av mättningsfel samt användarvänligheten gjorde markörlös optisk *motion capture* till ett lovande verktyg för att mäta kliniskt relevanta asymmetrier.

Tolkning av objektiv rörelseanalys

De olika systemen har individuella sätt att presentera rörelseanalysen; Lameness Locator har en funktion (AIDE) som tolkar mätningen (Di Tulio *et al.* 2017; Keegan 2018). Först graderas varje ben (0,5–4) efter hur långt i millimeter över tröskel-

värdet dess asymmetri går. Denna siffra multipliceras sen med bevisstyrkan (1–3) för mätningen, vilken definierats av variabiliteten i mätningens inspelade steg. Produkten blir en gradering av asymmetri (ohalt, lindrig, måttlig, kraftig) med en angiven säkerhet (svag, måttlig, stark) för respektive ben.

Sleip AI, som användes i denna studie, är baserad på markörlös optisk *motion capture* (Nyström, A., Sleip AI, pers. medd., 2021). Applikationen använder *range of motion* (ROM) för huvud och bäcken som referens till asymmetrigraderingen (mycket lindrig, lindrig, måttlig, kraftig) och valideras (låg, medel, hög) individuellt för fram- och bakben, baserat på antal steg och dess variabilitet.

2.2.8. Jämförelse mellan subjektiv och objektiv rörelseanalys

Under senaste decenniet har ett antal studier jämfört subjektiv och objektiv rörelseanalys. McCracken *et al.* (2012) visade att IMUs kunde identifiera en mer låggradig håltä än konsensus för tre erfarna hästveterinärer. I studien utrustades 15 hästar med skor, där en skruv kunde dras åt för att öka sultrycket. Hästarna bedömdes upprepade gånger i trav på rakt spår. Sultrycket ökade gradvis för varje bedömning och hästen räknades som subjektivt bedömd när konsensus rädde mellan veterinärerna. I 58,3 % av fallen identifierade IMU-systemet rätt ben innan veterinärerna, i 8,3 % av fallen var ordningen omvänd och i 33,3 % av fallen identifierades rätt ben samtidigt av IMUs och veterinärer.

I en jämförande studie av subjektiv och objektiv rörelseanalys lät författarna tre erfarna klinikveterinärer och ett system med IMUs bedöma 106 hästar (Keegan *et al.* 2013). Klinikerna genomförde en fullständig håltutredning exklusive diagnostiska anestasier och värden från IMUs registrerades i trav på rakt spår. Överensstämmelsen var acceptabel ($\kappa = 0,20$ – $0,40$) till måttlig ($\kappa = 0,41$ – $0,60$) för framben och dålig ($\kappa = 0,00$ – $0,20$) för bakben.

En studie undersökte överensstämmelse mellan subjektiv och objektiv bedömning för *fit-to-compete* definierad av Fédération Equestre Internationale (FEI) (Bragança *et al.* 2020). Författarna jämförde en markörbaserad optisk *motion capture* med subjektiv bedömning, utförd av tre erfarna veterinära delegater från FEI och en veterinär med specialistutbildning, av 12 hästar i trav på rakt spår. Veterinärernas samstämmighet var acceptabel ($\kappa = 0,40$). Om man använde veterinärernas konsensus som referens var sensitiviteten 83,3 % och specificiteten 66,7 % för den kvantitativa metoden, vilket gjorde att man ansåg att objektiv rörelseanalys kan användas för att understödja beslut om en häst är *fit-to-compete*.

En retrospektiv studie jämförde diagnos efter fullständig håltundersökning med IMU-mätvärden från trav på rakt spår för 1224 hästar (Reed *et al.* 2020). För de

flesta hästar som bedömdes ohalta eller fick en diagnos som förklarade halthan, var resultaten från den första objektiva mätningen på rakt spår identiskt med de definitiva slutsatserna när man beaktade resultat från alla undersökningar, exempelvis palpation, rörelseanalys, diagnostiska bedövningar och bilddiagnostik (röntgen, ultraljud, scintigrafi, CT, MR). Författarna poängterar att det dock fanns undantag från detta. Man ska notera att denna studie inte har jämfört objektiv och subjektiv bedömning, utan de utredande veterinärerna har använt objektiva data i sina utredningar. Vidare beskrev författarna att de flesta kombinerade fram- och bakbensasymmetrier kunde tolkas med kunskap om kompensatoriska rörelsemönster, och således användas för att lokalisera den primära halthan.

2.2.9. Artificiell intelligens tolkar objektiva rörelsedata

I och med att det är lätt att generera stora mängder data vid objektiva mätningar, samt att specifika kombinationer av flera mätvärden kan generera ett mer precist resultat (Hardeman *et al.* 2021), är robust men flexibel hantering och tolkning av data viktigt (Khera & Kumar 2020).

Artificiell intelligens (AI) kan efterlikna hjärnans förmåga att dra slutsatser, planera, lösa problem och inhämta ny kunskap (Nationalencyklopedin 2021) och kan idag hantera data med hög komplexitet (Dolatabadi *et al.* 2017) och utveckla modeller som automatiskt lär sig från databaser, gör exakta förutsägelser och agerar intelligent (Kutilek & Hozman 2011). Maskininlärning (ML), ett fält inom AI, används frekvent i diagnostisering, mönsterigenkänning, bildhantering, klassificering, prediktionsanalys och monitorering, vilket gör att det passar bra för rörelseanalys (Khera & Kumar 2020). I en reviewstudie som undersökt vilken roll ML kan ha i fältet human rörelseanalys fann man att exaktheten för asymmetridetektion är 85–95 % (Khera & Kumar 2020). ML har använts inom humanmedicin för att diagnosticera rörelsestörningar och planera rehabilitering (Takeda *et al.* 2021) samt förutspå fallrisk hos äldre (Begg *et al.* 2005). För hästuppopulationen är utvecklingen pågående. De tekniker som finns publicerade gör en prediktion för halthan utifrån extraherade punkter (Wang *et al.* 2021) eller från en helkroppsmodell av hästen (Li *et al.* 2021) och har visat varierande kapacitet att klassificera halthan från video, sannolikt på grund av olika grader av halthan hos de hästar man har använt i sina tester.

ML fungerar så att man väljer en (eller flera) tekniker (modeller) för att träna en klassificeringsuppgift med hjälp av ett träningsdataset (Khera & Kumar 2020). Därefter utvärderas modellen på ett tidigare osett dataset. När man är nöjd med resultatet avslutas processen, vilken annars fortsätter och modellens parametrar justeras tills önskad exakthet uppnås. ML använder olika huvudtekniker för att lära sig hantera data; övervakad, oövervakad samt förstärkt inlärning och inom dessa

nyttjas algoritmer som support vector machine (SVM), neurala nätverk (NNs), random forest (RF), hidden Markov model (HMM), ensemble learning, k-nearest neighbour (kNN) och beslutsträd (DTs).

2.2.10. Hälta versus naturlig lateralitet

Den utvecklade användningen av objektiv rörelseanalys har aktualiserat frågan om hästar har en naturlig lateralitet då flera studier visat att hästar som anses friska av sina ägare och presterar bra ändå rör sig asymmetriskt.

I en studie där objektiv rörelseanalys användes för att utvärdera friska hästar hade 72.5 % av hästarna ett rörelsemönster som låg bortom tröskelvärdena för hälta i trav på rakt spår (Rhodin *et al.* 2017). I denna studie drog man dock inte några slutsatser huruvida dessa hältor var smärtorsakade eller ej.

I en crossoverstudie med syfte att skilja smärtutlöst asymmetri från naturlig lateralitet behandlades hästar med meloxicam (Persson-Sjodin *et al.* 2019). Hästarna som ingick i studien ansågs ohalt av sina ägare men uppvisade med objektiv analys ett asymmetriskt rörelsemönster. Fyra dagars behandling med meloxicam gav igen förändring i hästarnas rörelsemönster och författarna konkluderar att antingen är symmetrin en naturlig variation eller orsakas den av smärta/dysfunktion som ej påverkas av behandling med meloxicam.

Dyson (2014) anser i en personlig krönika att andelen hästar med ett asymmetriskt rörelsemönster, som inte är orsakat av smärta, är ganska liten. Hon medger att det visserligen finns enstaka hästar som i normalt arbete uppvisar en regelbunden asymmetri och för att anse denna vara orsakad av biologisk lateralitet ska riktningförändringar eller ryttarens sittben inte påverka hästens rörelsemönster. Hästarna ska inte heller visa tecken på obehag, utföra ett gott arbete, vara arbetsvilliga och förbli lika asymmetriska över tid oavsett ökad arbetsbelastning.

3. Material och metod

3.1.1. Sleip AI

Sleip AI är en mobilapplikation som mäter rörelseasymmetri hos hästar med markörlös optisk *motion capture*. Dess tre neurala nätverk har tränats att känna igen och fånga rörelsen hos ett flertal olika kroppsdelar på en häst. Träningen har skett med videomaterial av hästar från både publika och egna källor, vilka har genererat data, som i sin tur genererat en större mängd data, för att ytterligare träna nätverken. Detta kallas data-augmentering och används rutinmässigt för att träna neurala nätverk.

För att mäta en häst, filmas den först inuti applikationen varefter filmen laddas upp till en central server som analyserar videon. Sleip AI använder vertikal *range of motion* (ROM) för att detektera asymmetri. Skillnaden i huvudets (HDmax/HDmin) och bäckenets (PDmax/PDmin) högsta och lägsta punkt, graderas efter storlek på avvikelse jämfört med motstående ben (mycket lindrig, lindrig, måttlig, kraftig), se tröskelvärden i tabell 3. Mätningen valideras (låg, medel, hög) individuellt för fram- och bakben, baserat på antal steg och dess variabilitet. För att visualisera asymmetristorleken för bak- och framben på ett direkt jämförbart sätt, omvandlas rådata till ett asymmetrivärde (0–2) genom att för framben/huvud multipliceras med 2,5 och för bakben/bäcken med 4. Olika tal används för omvandlingen eftersom storleken på asymmetri som skapas vid hälta är mindre för bäcken i jämförelse med huvudet.

Tabell 3. Tröskelvärden i ROM och asymmetrivärde samt färgkodning för Sleip AIs fem olika nivåer av asymmetri.

Hältgradering Sleip AI				
Gradering	Färg	Asymmetrivärde (0–2)	Skillnad ROM (%)	
			Framben	Bakben
Ohalt	Grön	0–0,19	<8	<5
Mkt lindrig	Grå	0,20–0,49	8–20	5–12,5
Lindrig	Gul	0,50–0,99	20–40	12,5–25
Måttlig	Orange	1,00–1,49	40–60	25–37,5
Kraftig	Röd	>1,50	60<	37,5<

Den färdiga analysen skickas i normalfallet tillbaka till användarens telefon efter två till tre minuter, se figur 1. Analysen innehåller en översikt som presenterar resultatet i ord, som en illustration och som vektorer, tillsammans med en analys av

tillförlitligheten hos mätningen. Resultatet presenteras även i detalj, dels i en interaktiv graf, dels som cirkeldiagram där varje individuellt steg visualiseras synkront med att videon spelas upp. Varje steg har då omvandlats till en vektor, och av dessa beräknas ett vektormedelvärde. Vidare innehåller analysen den bearbetade videoversionen med zoom och lägre uppspelningshastighet, vilken visas samtidigt som varje steg ritas upp som bifasiska sinuskurvor synkroniserat med filmen. Slutligen finns även originalvideon i analysen.



Figur 1. Bilder på applikationen Sleip AI. Från vänster ses i ordning en vy med sammanfattat resultat samt analysssäkerhet, detaljerad analys, steganalys med en vektor för varje stegcykel för fram- och bakben, video med simultant uppritade sinuskurvor. Bildillustrationer med tillstånd från Sleip AI.

3.1.2. Hästarna

Totalt medverkade 10 hästar i studien, fem svenska halvblod, ett belgiskt halvblod, en Pura Rasa Española, en Lusitano, en poloponny och en varmblodig travare varav 5 valacker och 5 ston. Åldern på hästarna var 6–21 år, med en medelålder på 11,8 år. Åtta av dessa besökte Hästkliniken vid Universitetsdjursjukhuset (UDS) för hältutredning, en häst deltog i ett annat forskningsprojekt vid Sveriges lantbruksuniversitet (SLU) och en häst mättes i sin hemmiljö i kontinuerlig rörelsekontroll. Ingen av hästarna mättes med Sleip AI specifikt för denna studie. Alla djurägare har medgivit att hästen får ingå i forskning om hältidiagnostik vid SLU. Hästarna benämns i resten av studien dels efter vilken hältgrad dess primärhälsa haft i mätning med Sleip AI (se tabell 1) samt efter vilket visningsnummer de haft i undersökningen med veterinärerna.

Av de två hästar som, enligt Sleip AI, är helt ohalta (se tabell 1) bedömdes den ena ohalt vid en rörelsekontroll med anledning av en äldre muskelskada och den andra har ej bedömts av veterinär utan mättes i kontinuerlig rörelsekontroll på hemma-

plan. Fem av sex hästar med, enligt Sleip AI, lindrig–kraftig hälta hade vid besöket administrerats en diagnostisk anestesi som minskade den primära hälтан med >50 % enligt journaltexten, se tabell 1. Om behandlande veterinär skrivit att hälтан är släckt, har detta tolkats som 100 %. Alla dessa hästar mättes före bedövningen administrerades. Häst nummer sju fick ingen diagnos vid besöket från vilket mätning använts, men diagnosticerades med en infekterad corp al i strålfåran vid uppföljande besök 14 dagar senare. Av de två hästar som, enligt Sleip AI, är mycket lindrigt asymmetriska bedöms den ena som ohalt vid en uppföljning av en svullen kotsenskida höger bak och den andra är mätt efter att en diagnostisk anestesi (abaxial VB) administrerats, vilken reducerat både primärhälтан och den kompensatoriska hälтан med 70 % (från måttlig till mycket lindrig).

Tabell 1. Hästar som ingått i denna studie. Sorterade efter ökande hältgrad enligt Sleip AI. Primärhälta enligt Sleip AI. B = Övervägande belastning, F = Övervägande frånsjut. ^LVideo med ljud. *Administrerad innan använd mätning. **Information om avläsningstid ej tillgänglig. ***14 dagar senare hittas kraftigt infekterad corp. al. i medial strålfåra VB.

Studiens hästar							
Häst	Sleip AI			Journalinformation			
	Primärhälta			Bedövning			Diagnos
	Ben	Grad	Typ	Lokalisation	Avläsning	Effekt	
9	Ohalt						Ohalt
5	Ohalt						Ohalt
2	HF	Mkt lindrig	B				Ohalt
3 ^L	VB	Mkt lindrig	B	Abaxial VB*	55 min	75%	Ospecificerad
8	HF	Lindrig	B	Ringblock HF	40 min	100%	Ospecificerad
10 ^L	HB	Lindrig	F	Proximalt gaffelband HB	– **	100%	HB, desmit medial gaffelbandsgren. Misstanke involvering proximalt gaffelband.
6	HB	Lindrig	F	Gonus HB, 3 ledavdeln.	45 min	100%	HB, gonit.
1 ^L	HF	Måttlig	B	Abaxial HF	– **	50%	HF, strålbensbursit.
7 ^L	VB	Måttlig	F	Gonus VB	75 min	0%	Ospecificerad***
4	VF	Kraftig	B	Abaxial VF	25 min	50%	VF, fraktur medial hovbensgren.

Hästarna har mätts i trav vid hand, på betongunderlag i löpargången på UDS eller i ridhus med sandunderlag. Alla hästar travade fram och tillbaka två gånger, vilket ger en total sträcka om minst 80 meter. För denna studie har inga fysiska markörer använts, men eftersom alla hästar utom en även ingick i en annan studie, vilken använde markörbaserad optisk *motion capture*, hade reflex-markörer monterats på

nio av tio hästar. Mätningarna är utförda med en iPhone XS samt en iPhone 12 Pro fäst i ett stativ.

Kravet för att ingå i studien var att hästen skulle vara filmad rakt bak- respektive framifrån. Inspelningen skulle vara av god teknisk kvalitet, ha minst medelgod säkerhet och registrerat minst 12 steg för fram-/ respektive bakben. Vidare skulle aktuellt asymmetrivärde för huvud/bäcken vara minst lika stort eller större än standardvariationen för samma asymmetri för alla halta hästar. Se alla hästars Sleip AI-analyser i Bilaga 1.

Tio hästar, som uppfyllde alla kriterier, valdes ut för att representera en maximal bredd av asymmetrigrad, från ohalt (ingen asymmetri) till kraftig asymmetri, jämnt fördelade på fram- och bakben, se tabell 2. Vid valtillfället fanns ingen häst som uppmätt kraftig asymmetri för bakben och således adderades i stället ytterligare en häst med lindrig asymmetri. För att fastställa vilket ben som ska anses primärhalt, har respons vid diagnostisk anestesi använts. Vid asymmetri på två ben har respons vid diagnostisk anestesi samt definition av kompensatorisk hälta, som nämns tidigare i denna studie använts för att avgöra vilket ben som har en primär respektive kompensatorisk asymmetri. Fyra hästar hade en kompensatorisk hälta.

Visningsordning för bedömning av hästarna lottades, och var densamma för alla bedömare. För varje häst användes en video i normal uppspelningshastighet och en som processats av Sleip AI. Båda var komprimerade i quicktime-format i 3840×2160 pixlar respektive 2316×2472 pixlar. Den Sleip AI-processade videon hade lägre uppspelningshastighet och zoomade in på hästen varefter den travade längre bort från kameran, vilket gjorde att hästen hade en konstant storlek i videon.

Tabell 2. Sammanställning av data från alla hästars mätningar i Sleip AIs sorterade efter asymmetrigrad. För asymmetrivärden indikerar ett negativt tal att värdet kommer från väster ben, ett positivt tal indikerar att värdet kommer från höger ben. H = Häst. Fet stil = primärhåla.
*Mätningssäkerhet anges 0–100, där 100 är maximal säkerhet.

Mätdata Sleip AI										
H	Asymmetrivärde				Tillförlitlighet					
					Antal steg		Variation		Säkerhet*	
	HDmax	HDmin	PDmax	PDmin	Fram	Bak	Fram	Bak	Fram	Bak
9	-0,10±0,5	-0,12±0,3	-0,10±0,1	0,07±0,1	17	13	Hög	Låg	45	62
5	-0,17±0,1	-0,12±0,1	0,12±0,1	-0,08±0,1	43	26	Låg	Låg	100	96
2	-0,22±0,3	0,49±0,4	0,09±0,1	0,08±0,1	27	13	Hög	Låg	65	66
3	0,06±0,2	-0,49±0,2	0,06±0,1	0,45±0,2	35	26	Medel	Medel	89	87
8	0,34±0,2	0,51±0,2	0,02±0,1	0,00±0,1	29	20	Medel	Låg	79	84
10	0,11±0,2	0,17±0,2	0,51±0,1	0,01±0,1	43	24	Medel	Låg	88	93
6	0,53±0,2	0,77±0,3	0,64±0,1	0,18±0,1	21	13	Medel	Låg	58	62
1	0,29±0,3	1,26±0,3	-0,16±0,1	0,17±0,1	24	13	Medel	Låg	64	61
7	0,06±0,8	-1,23±0,7	-1,13±0,2	-0,48±0,2	24	15	Medel	Låg	53	56
4	-0,45±0,3	-1,54±0,3	0,36±0,1	-0,39±0,1	19	12	Medel	Låg	52	55

3.1.3. Deltagarna

Målpopulationen för undersökningen i denna studie var hästpraktiserande veterinärer i Sverige, med huvudinriktning ortopedi. Inbjudan till att delta i studien skickades via mail till 19 olika hästkliniker.

3.1.4. Undersökningen

Bedömningen av hästarna genomfördes antingen helt på distans eller vid fysiska möten på veterinärernas arbetsplatser. I distansbedömningarna fick veterinärerna ett email med instruktioner för undersökningen, en länk till filmerna på en låst Google drive samt ett svarsdokument och ett deltagarmedgivande som de återsände ifyllda. I de fysiska mötena gavs instruktioner muntligen, filmerna visades på en MacBook Pro, 15” med grafikkort Intel Iris Pro 1536 MB och deltagarna fick själva fylla i ett svarsformulär samt signera deltagarmedgivande. Veterinärerna kunde titta obegränsat antal gånger på båda filmbearbetningarna samt byta fritt mellan dem. De fick även ange vilken av filmtyperna de tyckte var bäst att bedöma på.

Bedömningen inleddes med en enkät som kartlade deltagarens kön, ålder, tid i praxis, huvudsakliga veterinärmedicinska inriktning och patientgrupp samt användning av objektiv rörelseanalys i det kliniska arbetet. För varje häst beskrev veterinärerna primärhåltan, med avseende på vilket ben (HF, VF, HB, VB), håltgrad (0–5) och hålttyp (frånskjut, belastning, annan). De fick även ange om det fanns en kompensatorisk hålt (ben, grad, typ) samt med vilken säkerhet (låg = 1, medel = 2, hög = 3) de bedömde hästen. I det fall två ben angivits som primärhålt (7/120), har det ben som samstämmer med Sleip AIs bedömning använts för beräkning av primär hålt och det andra för kompensatorisk hålt. Gradering har angivits i halvtal och oberoende vilket ben som angivits. I de fall gradering angivits som ett intervall (13/120), har ett medelvärde använts för beräkning vid heltalsintervall (4/120) och för halvtalsintervall (9/120) har avrundning skett uppåt.

Det fanns även utrymme att lämna en kommentar för varje häst. Slutligen fanns fem öppna frågor angående hästortopedi, vilka besvarades i mån av tid.

1. Vad tittar på du när du bedömer en hästs rörelse?
2. Vid vilket moment i dina håltutredningar upptäcker du oftast håltan som du sedan behandlar?
3. Hur ofta har du återbesök efter en håltutredning? Vad gör du vid återbesöken?
4. Om du använder objektiva mätmetoder, hur tycker du att det påverkar din förmåga att bedöma hålt?
5. Vad tror du skulle kunna minska förekomsten av håltor i Sverige idag?

3.1.5. Statistiska beräkningar

Veterinärernas konsensus som grupp avseende primär och kompensatorisk hålt, både vad gäller ben och grad, beräknades med Fleiss' och Cohens kappa (κ) (RStudio 2021). Varje individuell veterinärs samstämmighet med Sleip AI avseende primär hålt beräknades med Cohens kappa. Överensstämmelsen tolkades för båda κ -värden som dålig $\kappa \leq 0,3$, acceptabel $\kappa = 0,31–0,5$, bra $\kappa = 0,51–0,8$ och mycket bra för $\kappa > 0,8$ i enlighet med Hammarberg et al. (2016). Korrelation mellan veterinärens överensstämmelse med Sleip AI (κ) och median för veterinärens överensstämmelse med gruppen (κ) beräknades med Pearsons korrelationskoefficient (r) (Social Science Statistics 2021a) och Spearmans rangkorrelationskoefficient (r_s) (Social Science Statistics 2021b), efter att normalfördelning konstaterats med Lilliefors test (Arsham 2015).

Vidare visualiserades korrelationen mellan varje veterinärs gradbedömning (0–5) och Sleip AIs asymmetrivärden (0–2) i ett spridningsdiagram med en trendlinje. Vid beräkningar har det största asymmetrivärdet för det primärhaltabenet använts.

För att undersöka om det fanns några samband mellan veterinärernas uppgivna säkerhet, år i praxis, antal fall per år, användande av objektiv rörelseanalys, överensstämmelse med Sleip AI, konsekvens i hältgradering, ålder och kön gjordes beräkningar med Spearmans rangkorrelationskoefficient (r_s) (Social Science Statistics 2021b).

För beräkningarna användes RStudio samt två onlinebaserade statistikkalkylatorer (Social Science Statistics 2021a and 2021b; Arsham 2015). Diagrammen är skapade i Microsoft Excel.

3.1.6. Jävsdeklaration

Elin Hernlund, huvudhandledare för detta examensarbete, är en av grundarna av Sleip AI. Sleip AI ägs idag av grundarna och ett flertal minoritetsägare varav en är SLU Holding AB. Julia Stenius, författare till denna studie är minoritetsägarna och har bidragit i designutvecklingen av Sleip AI. Ingen finansiering av denna studie har utgått från Sleip AI.

4. Resultat

4.1. Deltagarna

Tolv veterinärer från 10 av de 19 tillfrågade klinikerna deltog i studien. Som huvudsaklig veterinärmedicinsk inriktning uppgav en veterinär *sporthästmedicin* och övriga 11 *ortopedi*. Deltagarna bestod av 4 män och 8 kvinnor, med snittålder 44,6 (30–69) år och i genomsnitt 16,4 (2,5–37) år i praxis. Veterinärerna handlade i snitt 6,7 fall/dag eller 1535 fall/år (400–3000 fall/år) och var jämnt fördelade på små (1–2 veterinärer), medelstora (3–10 veterinärer) och stora kliniker (>10 veterinärer) i länen Halland, Kronoberg, Skåne, Södermanland, Västra Götaland, Uppland och Östergötland. Sex veterinärer svarade att de aldrig använde objektiv rörelseanalys, tre stycken mycket sällan och tre dagligen. Fem veterinärer gjorde sina bedömningar vid ett fysiskt besök och sju digitalt. Veterinärerna uppskattade den totala säkerheten i sin bedömning av alla hästar till medelgod, median 20 (15–30).

På grund av tekniska problem kunde en veterinär endast se videon som processats av Sleip AI (med in-zoomning och långsammare hastighet) för alla hästar förutom nummer 3. Dessa resultat skiljer sig inte signifikant från resten och har inkluderats i studien.

4.2. Primärhåla

Alla veterinärer angav minst ett primärhalt ben för varje häst, alternativt angav den som ohalt, se tabell 4. Svaren för alla hästar fördelas i fallande ordning enligt: HB 38,3 %, VB 21,6 %, VF 18,3 %, HF 13,3 %, ohalt 8,3 % och kan grupperas i bakben 60 %, framben 31,6 % och ohalt 8,3 %. Enligt Sleip AI är primärhåla fördelad: HF 30 %, HB 20 %, VB 20 %, ohalt 20 %, VF 10 % och kan grupperas i framben 40 %, bakben 40 % och ohalt 20 %.

Tabell 4. Primärhålta. Veterinärernas svar angående vilket ben samt typ av hålt. B = belastning, F = frånskjut. ^LOriginalvideo med ljud.

Primärhålta										
Häst	Sleip AI			Veterinärer						
	Ben	Grad	Typ	Ben					Typ	
				HF	VF	HB	VB	Ohalt	Majoritet	Belastning Frånskjut
9	Ohalt					3	3	6	Ohalt	1 1
5	Ohalt				4	3	1	4	Ohalt/VF	3 2
2	HF	Mkt lindrig	B		1	9	2		HB	5 2
3 ^L	VB	Mkt lindrig	B	1	4	1	6		VB	7 2
8	HF	Lindrig	B	4	1	7			HB	8 1
10 ^L	HB	Lindrig	F			12			HB	8 3
6	HB	Lindrig	F	1	1	10			HB	9 2
1 ^L	HF	Måttlig	B	10	1	1			HF	10 1
7 ^L	VB	Måttlig	F				12		VB	9 3
4	VF	Kraftig	B		10		2		VF	8 2
Totalt antal svar				16	22	46	26	10		68 19

4.2.1. Överensstämmelse ben

Inom gruppen

Generell överensstämmelse inom veterinärgruppen, för alla hästar sammantaget, var acceptabel och beräknades med Fleiss' kappa till $\kappa = 0,45$ (se tabell 5) och som en median av Cohens kappa till $\kappa = 0,48$. För de ohalta samt mycket lindrigt halta hästarna ($n = 4$) var överensstämmelsen dålig ($\kappa = 0,14$) men för de lindrigt till kraftigt halta hästarna ($n = 6$) var överensstämmelsen bra ($\kappa = 0,64$). För de frambenshalta hästarna ($n = 4$), vilket inkluderar en mycket lindrig hålt, en lindrig, en måttlig och en kraftig, är överensstämmelsen acceptabel ($\kappa = 0,42$). För bakbenshalta hästar ($n = 4$), vilket inkluderar en mycket lindrig hålt, två lindriga och en måttlig, var överensstämmelsen bra ($\kappa = 0,60$). För de halta hästar vars videos hade ljud ($n = 4$), vilka inkluderade en mycket lindrig hålt, en lindrig hålt och två måttliga, var överensstämmelsen bra ($\kappa = 0,65$) medan överensstämmelsen för videos utan ljud, som inkluderande en mycket lindrig hålt, två lindriga och en

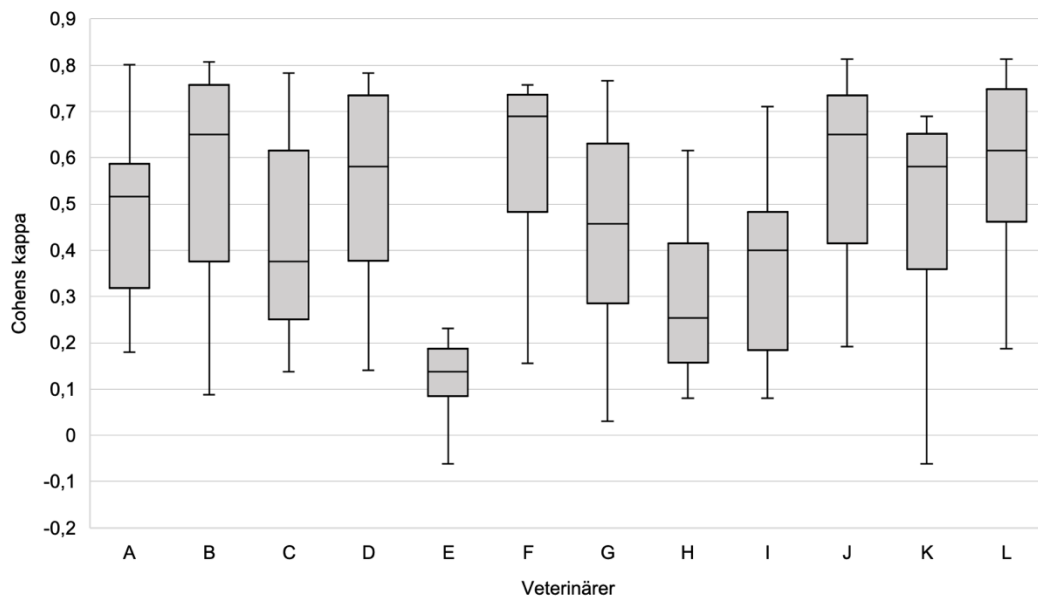
kraftig, var precis inom ramen för acceptabel ($\kappa = 0,33$). Alla ovanstående beräkningar är signifikanta ($p < 0,05$).

Tabell 5. Primärhalt ben. Överensstämmelse inom veterinärgruppen. Beräknat med Fleiss' kappa.

Primärhåla. Överensstämmelse inom veterinärgruppen.				
Hästar	Antal	Fleiss κ	p	Tolkning
Alla	10	0,45	0	Acceptabel
Ohalta–Mycket lindrigt halta	4	0,14	<0,05	Dålig
Lindrigt–Kraftigt halta	6	0,64	0	Bra
Frambenshalta	4	0,42	0	Acceptabel
Bakbenshalta	4	0,60	0	Bra
Halta, video med ljud	4	0,65	0	Bra
Halta, video utan ljud	4	0,33	<0,05	Acceptabel

Mellan veterinär och grupp

Varje individuell veterinärs överensstämmelse med resten av gruppen beräknades parvis med Cohens kappa, se figur 2. Resultatet varierade mellan ingen överensstämmelse (veterinär E och K $\kappa = -0,06$) och mycket bra (veterinär C och D $\kappa = 0,78$). Medianvärdet för varje veterinär varierade mellan $\kappa = 0,14$ och $\kappa = 0,69$.



Figur 2. Primärhalt ben. Låddiagram för varje veterinärs överensstämmelse med alla andra veterinärer, beräknat parvis med Cohens kappa.

Mellan grupp och Sleip AI

Överensstämmelsen mellan veterinärgruppens majoritetssvar, det vill säga det svar som flest veterinärer har angivit, och Sleip AI var bra ($\kappa = 0,75$ respektive $\kappa = 0,63$), se tabell 6. För en häst var majoritetsbedömningen delad mellan alternativet ohalt och VF, därför utfördes två olika beräkningar av Cohens kappa. För frambenshalta hästar ($n = 4$) var gruppmajoritetens överensstämmelse med Sleip AI precis inom acceptabel nivå ($\kappa = 0,33$) men inte signifikant ($p = 0,10$). För bakbenshalta hästar ($n = 4$) var gruppmajoritetens överensstämmelse med Sleip AI perfekt ($\kappa = 1$). I gruppen frambenshalta ingick en mycket lindrig halta, en lindrig, en måttlig och en kraftig, varav endast en video hade ljud. I gruppen bakbenshalta hästar ingick en mycket lindrig halta, två lindriga och en måttlig, varav tre videofilmer hade ljud.

Tabell 6. Primärhalt ben. Överensstämmelse mellan veterinärgruppens majoritetsval och Sleip AI. Beräknat med Cohens kappa. *Majoriteten var delad (4 mot 4) mellan två svarsalternativ för en häst, därför beräknas två kappavärden.

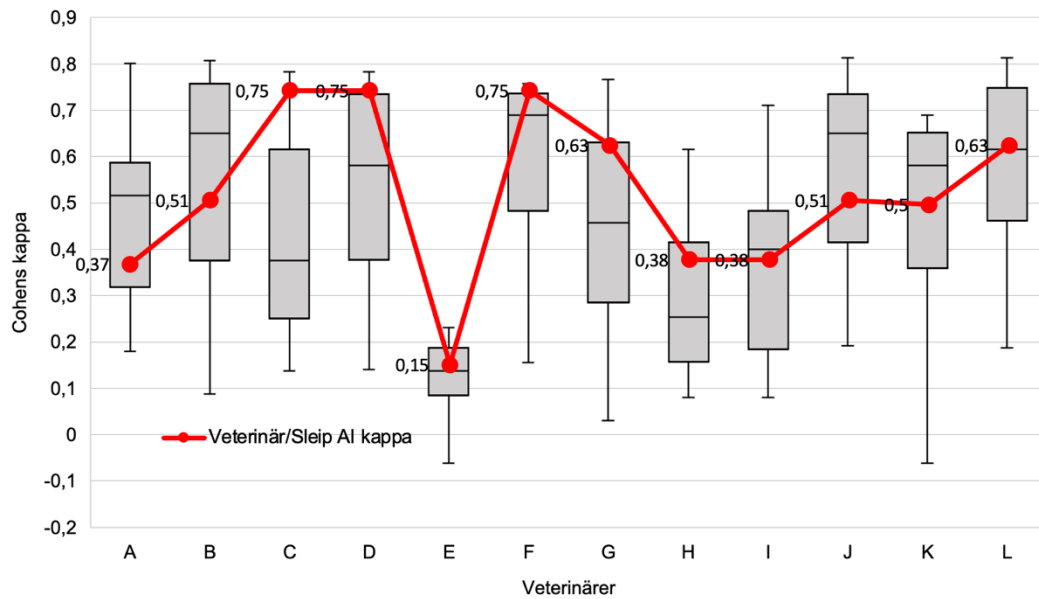
Primärhalt. Överensstämmelse mellan veterinärgrupp och Sleip AI.				
Hästar	Antal	Cohen κ	p	Tolkning
Alla	10	0,63/0,75*	0	Bra
Ohalta–Mkt lindrigt halta	4	0,39/0,64*	<0,05	Acceptabel/Bra
Lindrigt–kraftigt halta	6	0,77	<0,05	Bra
Frambenshalta	4	0,33	0,10	Acceptabel
Bakbenshalta	4	1,00	<0,05	Perfekt
Halta, video med ljud	4	1,00	<0,05	Perfekt
Halta, video utan ljud	4	0,33	0,10	Acceptabel

Mellan veterinär och Sleip AI

Varje veterinärs överensstämmelse med Sleip AI beräknades med Cohens kappa, se figur 3 (röd linje) samt tabell 7. Överensstämmelsen varierade mellan dålig ($\kappa = 0,15$) och betydande ($\kappa = 0,75$). Korrelationen mellan varje veterinärs överensstämmelse med Sleip AI och medianvärdet för varje veterinärs överensstämmelse med gruppen beräknades i två linjära regressionsmodeller till måttligt positiv (Mukaka 2012) och med Pearsons korrelationskoefficient ($r = 0,62$, $p < 0,05$, $r^2 = 0,38$) signifikant och med Spearmans rangkorrelationskoefficient icke signifikant ($r_s = 0,5$, $p = 0,099$).

Tabell 7. Primärhalt ben. Varje individuell veterinärs överensstämmelse med veterinärgruppen beräknad parvis med Cohens kappa samt varje veterinärs överensstämmelse med Sleip AI beräknad med Cohens kappa.

Primärhåla. Överensstämmelse mellan enskild veterinär och veterinärgruppen samt Sleip AI.				
Veterinär	Överensstämmelse med veterinärgrupp		Överensstämmelse med Sleip AI	
	Median κ	Tolkning	κ	Tolkning
A	0,52	Bra	0,37	Acceptabel
B	0,65	Bra	0,51	Bra
C	0,38	Acceptabel	0,75	Bra
D	0,58	Bra	0,75	Bra
E	0,14	Dålig	0,15	Dålig
F	0,69	Bra	0,75	Bra
G	0,46	Acceptabel	0,63	Bra
H	0,25	Dålig	0,38	Acceptabel
I	0,40	Acceptabel	0,38	Acceptabel
J	0,65	Bra	0,51	Bra
K	0,58	Bra	0,50	Acceptabel
L	0,62	Bra	0,63	Bra
Median	0,55	Acceptabel	0,51	Bra



Figur 3. Primärhalt ben. Linjen visar varje veterinärs överensstämmelse med Sleip AI, beräknad med Cohens kappa. Låddiagrammet visar varje veterinärs överensstämmelse med alla andra veterinärer, beräknad parvis med Cohens kappa.

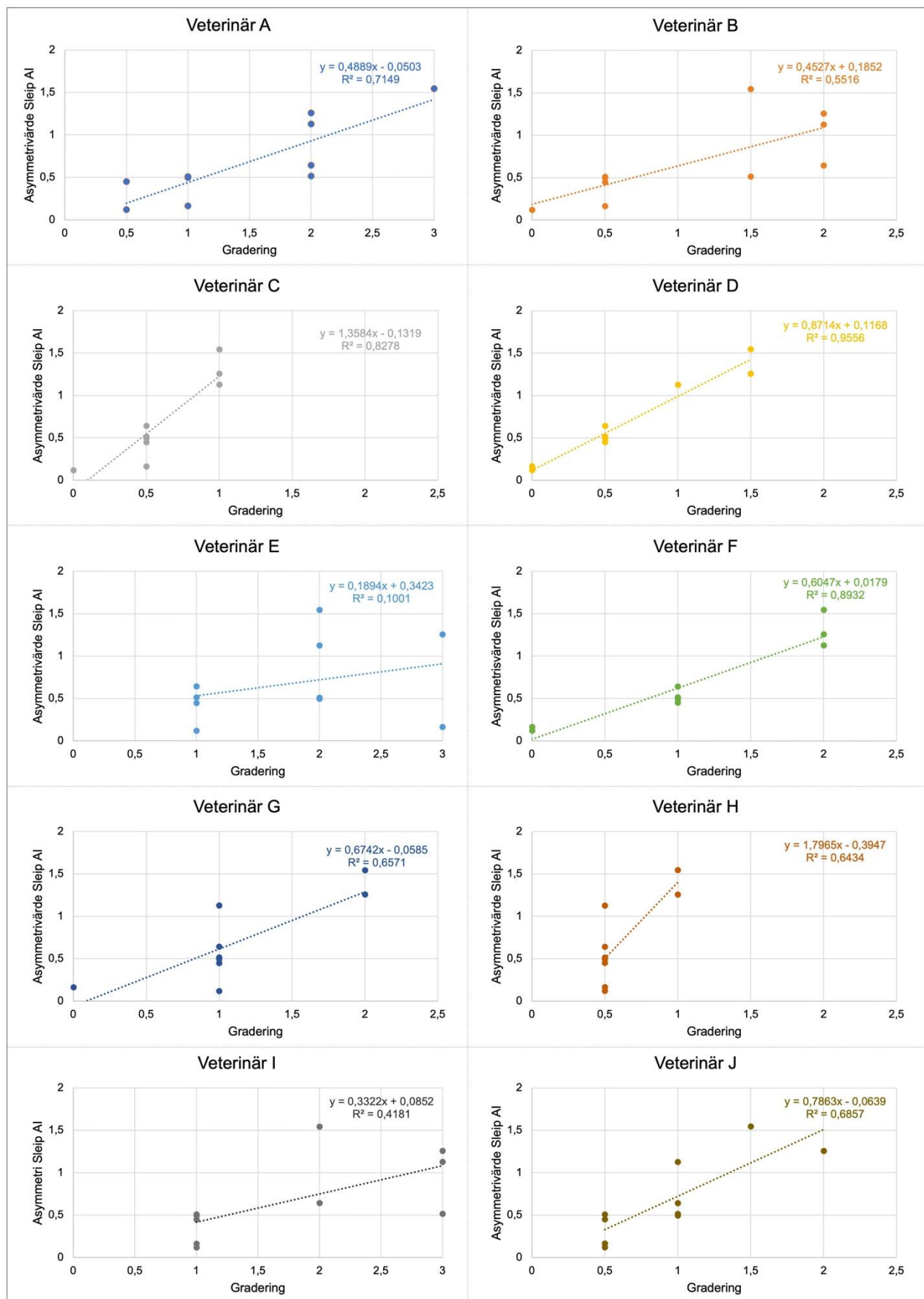
4.2.2. Hältgradering

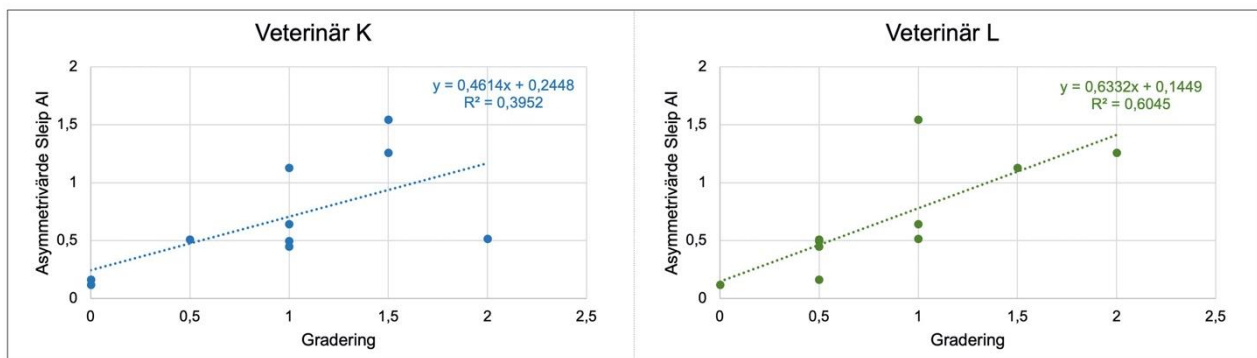
Alla veterinärer graderade varje hästs primärhalt ben, se tabell 8. Överensstämmelsen inom gruppen var dålig ($\kappa = 0,09$ $p < 0,05$) och har beräknats oberoende av vilket ben veterinären angav som primärhalt. Om veterinärernas gradering kategoriserades enligt följande: Ohalt, 0,5–1 grad, 1,5–2 grader, 2,5–3 grader, var överensstämmelsen nära acceptabel ($\kappa = 0,24$, $p = 0$).

Tabell 8. Gradering primärhåla. Antal svar för varje halvgrad. *Asymmetrivärde (0–2) inklusive standardavvikelse från det primärhåla benet. Om det finns både frånskjut- och belastnings-asymmetrier på samma ben, används här endast det största av dessa värden. Dessa värden är beräknade från rådata och justerade så att fram- och bakbensasymmetrier blir numeriskt jämförbara. Detaljerad uträkning ses under "Sleip AI" i Material och metod. ¹Originalvideo med ljud.

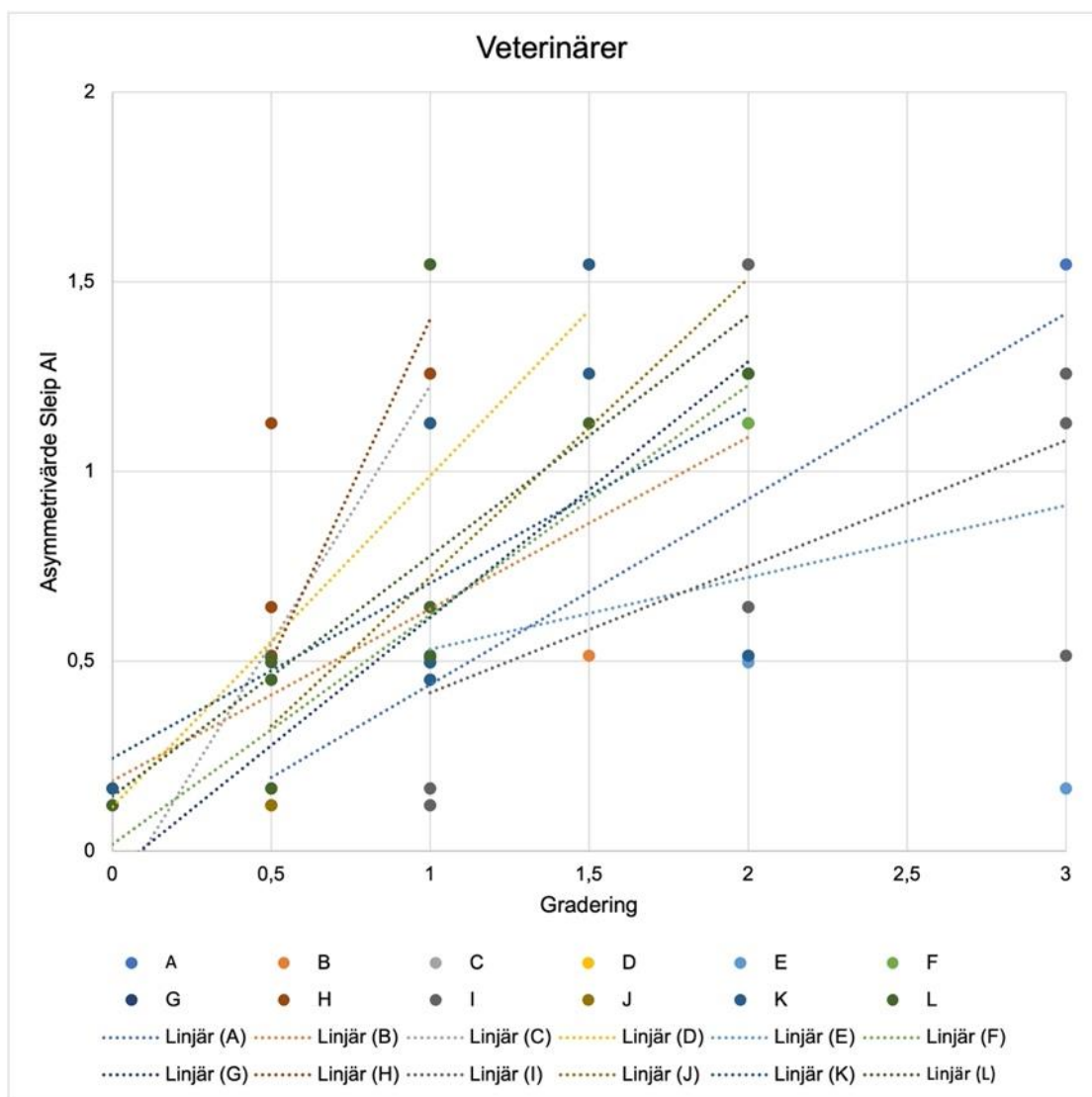
Gradering primärhåla										
Häst	Sleip AI			Veterinärer						
	Ben	Grad	Asymmetrivärde*	Grad						
				0	0,5	1	1,5	2	3	Median
9	Ohalt	Ohalt	0,12 ± 0,3	6	3	3				0,25
5	Ohalt	Ohalt	0,17 ± 0,1	4	5	2			1	0,5
2	HF	Mkt lindrig	0,49 ± 0,4		5	6		1		1
3 ^L	VB	Mkt lindrig	0,45 ± 0,2		7	5				0,5
8	HF	Lindrig	0,51 ± 0,2		7	4		1		0,5
10 ^L	HB	Lindrig	0,51 ± 0,1		3	5	1	2	2	1
6	HB	Lindrig	0,64 ± 0,1		3	6	0	3		1
1 ^L	HF	Måttlig	1,26 ± 0,3			2	2	6	2	2
7 ^L	VB	Måttlig	1,13 ± 0,2		1	5	1	4	1	1,25
4	VF	Kraftig	1,55 ± 0,3			3	4	4	1	1,5
Totalt				10	34	41	8	21	6	

Varje veterinärs håltgradering har redovisats i ett enskilt spridningsdiagram, se figur 4–16, samt i ett gemensamt diagram, se figur 17. I de enskilda diagrammen visas även en trendlinje. För den gäller räta linjens ekvation ($y = kx + m$), där y = Sleip AIs asymmetrivärde, x = veterinärens gradering, k = lutningen på linjen och m = skärningspunkten av y-axeln, vilket motsvarar förväntat Sleip AI asymmetrivärde vid håltgrad 0. I dessa diagram har även ett r^2 -värde beräknats, vilket indikerar hur bra modellen förklarar datasetet, där ett högre värde för r^2 betyder bättre förklaring. Värdet för k varierar mellan 0,19 och 1,80 (median 0,62). Värdet för m varierar mellan 0,02 och 0,40 (median 0,10). Värdet för r^2 varierar mellan 0,10 och 0,95 (median 0,65). Inga statistiska test för signifikans har utförts för trendlinjer och r^2 . Korrelation mellan den enskilde veterinärens förklaringsgrad för håltgradering (r^2) och överensstämmelse med Sleip AI (κ), beräknades till måttligt positivt ($r_s = 0,65$, $p < 0,05$) med Spearmans rangkorrelationskoefficient. En högre förklaringsgrad, det vill säga en mer konsekvent håltgradering, var alltså associerat med en bättre överensstämmelse med S.





Figur 4–16. Gradering primärhåltä. Individuella spridningsdiagram med trendlinje. Trendlinjen indikerar hur veterinärens gradering korrelerar med asymmetrivärden i Sleip AI. Ju högre tal för r^2 desto bättre förklarar modellen resultaten. För dessa modeller har inget test för signifikans utförts.



Figur 17. Gradering primärhåltä. Spridningsdiagram med en trendlinje för varje veterinär.

4.2.3. Hälttyp

Alla veterinärer förutom en angav vilken typ av hälta som hästen visade för minst en häst. Svaren fördelades mellan belastnings- (78 %), frånskjuts- (22 %) och överrullningshälta (1 %), se tabell 4. Enligt veterinärerna visade en häst (9) lika mycket belastningshälta (50 %) som frånskjutshälta (50 %), för denna häst hade dock endast två veterinärer angivit en hålttyp. För alla andra hästar var andelen svar som sa belastningshälta övervägande (60–91 %). Enligt Sleip AI visade fem av de halta hästarna huvudsakligen belastningshälta och tre stycken frånskjutshälta.

4.3. Kompensatorisk hälta

Alla veterinärer förutom en noterade minst en häst med kompensatorisk hälta, se tabell 9. I undersökningar är frågan ställd så att om veterinären bedömer att hästen har en kompensatorisk hälta, noteras påverkat ben, grad och typ. Om bedömare inte ser en kompensatorisk hälta, besvaras inte frågan, vilket noteras som ohalt i tabell 9. I snitt bedömdes 4,4 hästar (2–7) ha en kompensatorisk hälta. För de tre hästar (3, 6, 7) som hade en ipsilateral mycket lindrig till måttlig kompensatorisk hälta hade majoriteten av veterinärerna noterat samma ben som Sleip AI. För den häst (4) som hade en mycket lindrigt bilateral kompensatorisk bakbenshälta samt de sex hästarna utan kompensatorisk hälta har majoriteten av veterinärerna inte noterat någon kompensatorisk hälta. Överensstämmelsen både inom gruppen och i jämförelse med Sleip AI var dålig gränsande mot acceptabel (Fleiss/Cohen $\kappa = 0,26$, $p=0$). Då flera deltagare ej angivit håltgrad för kompensatorisk hälta utfördes inga statistiska beräkningar avseende detta.

Tabell 9. Kompensatorisk hälta. Veterinärernas svar angående vilket ben. Om veterinären ej besvarat frågan anses den ej ha noterat någon kompensatorisk hälta, vilket, har noterats som "ohalt" i tabellen. ^LOriginalvideo med ljud.

Kompensatorisk hälta										
Häst	Sleip AI			Veterinärer						
	Ben	Grad	Typ	Ben	HF	VF	HB	VB	Ohalt	Majoritet
9	Ohalt								12	Ohalt
5	Ohalt			1			3		8	Ohalt
2	Ohalt			1					11	Ohalt
3 ^L	VF	Mkt lindrig	B			6	1	2	3	VF
8	Ohalt			4					8	Ohalt
10 ^L	Ohalt			2					10	Ohalt
6	HF	Lindrig	F	10					2	HF
1 ^L	Ohalt		B	1			1	2	8	Ohalt
7 ^L	VF	Måttlig	F			8			4	VF
4	HB/VB	Mkt lindrig	F/B	1	1	3		1	6	Ohalt
Totalt antal svar				20	15	8		5	72	

4.4. Deltagarnas egenskaper

Alla deltagare angav kön, ålder, tid i praxis, fall per år, användande av objektiv rörelseanalys samt säkerhet för varje bedömd häst.

En signifikant korrelation sågs mellan hur väl den individuella veterinärens svar angående primärhalt ben överensstämde med Sleip AI och den totala säkerheten i bedömningen ($r_s = 0,60$, $p < 0,05$), där högre säkerhet hade en association med bättre överensstämmelse med Sleip AI. Signifikant korrelation sågs även mellan antal år i praxis och användandet av objektiv rörelseanalys ($r_s = -0,65$, $p < 0,05$), där färre arbetade år ökade sannolikheten att veterinären använde objektiv rörelseanalys i sitt dagliga arbete.

Utöver detta sågs inga andra signifikanta samband mellan säkerhet, år i praxis, antal fall per år, användande av objektiv rörelseanalys, överensstämmelse med Sleip AI, ålder eller kön. Huruvida veterinären hade utfört bedömningarna i ett fysiskt möte eller digitalt hade heller ingen signifikant påverkan på resultaten.

5. Diskussion

5.1.1. Deltagarna

Inbjudan att delta i denna studie skickades till 19 svenska hästkliniker i såväl personliga som generella email, vilket genererade 12 studiedeltagande veterinärer med huvudsaklig inriktning hästortopedi. Eftersom man ej med säkerhet kan säga hur många som nåtts av inbjudan är det svårt att bedöma svarsfrekvensen. Av de veterinärer som svarade på inbjudan, men avböjde medverkan, var den vanligaste anledningen tidsbrist. En veterinär avböjde på grund av sjukdom. En veterinär avböjde eftersom hen var skeptisk till studieupplägget och ifrågasatte relevansen av att bedöma hästar på video. Hen menade att studien inte hade någon verklighetsförankring, då den endast inkluderade bedömning av en enda rörelsesekvens, trav på rakt spår, och saknade andra komponenter i en hältutredning så som exempelvis palpation och böjprov. Hen ansåg också att rörelseanalysen var en förhållandevis oviktig komponent i en hältutredning. Flera deltagare i studien hade liknande invändningar, vilket diskuteras under avsnittet *Video* nedan.

Könsfördelningen bland deltagarna (67 % kvinnor) motsvarar hur det ser ut på landets djursjukhus (76 % kvinnor) och i privatpraktik (70 % kvinnor) (Wlosinska 2019). Specifikt för hästveterinärer finns ingen offentligt tillgänglig statistik för kön, ålder, tid i praxis eller verksamhetsområde (Nyman, S., Hästsektionen Sveriges veterinärförbund, pers. medd., 2021) men deltagarnas spann i ålder (30–69 år, medelvärde 44,6 år), tid i praxis (2,5–37 år, medelvärde 16,4 år), representation av sju län som tillsammans inhyser 52 % av Sveriges hästar (Heldt *et al.* 2016) samt den jämna fördelningen på små, medelstora och stora kliniker får anses nöjaktigt representera populationen hästveterinärer med huvudsaklig inriktning hästortopedi i Sverige idag.

5.1.2. Hästarna

Eftersom hästarna valdes ut för att representera en jämn fördelning av fram- och bakbenschältor av ökande grad, kan de inte anses slumpmässigt utvalda. Vidare är sex av tio av halvblodstyp och endast en ponny finns representerad, vilket gör att gruppen ej nöjaktigt kan anses representera hela den svenska hästpopulationen.

Inklusionskravet för hästarna var en Sleip AI-mätning av minst medelhög säkerhet samt att de lindrigt–kraftigt halta hästarnas primärhältor skulle vara verifierade av en, minst 50 % positiv, diagnostisk bedövning. Idealet hade varit att en positiv diagnostisk bedövning hade varit inklusionskriterium även för de två mycket lindrigt

halta hästarna men i det tillgängliga materialet fanns endast en häst (3) med mycket lindrig hälta verifierad av en diagnostisk anestesi. Den avviker dock från de andra hästarna eftersom mätningen som används är gjord efter att bedövningen administrerats. Troligen fanns få hästar i denna kategori, eftersom graden av vertikal asymmetri, <20 % för framben och <12,5 % för bakben, har visats utmanande ögat att detektera (Starke & Oosterlinck 2019), vilket gör att de vid visuell rörelseanalys ofta bedöms som ohalta, eller inom normalvariation. Detta medför således att, det i den kliniska vardagen, inte heller är brukligt att administrera diagnostisk bedövning vid denna typ av låggradiga asymmetrier.

Den andra hästen (2) med mycket lindrig hälta, besökte kliniken då den en månad tidigare haft en galla i kotsenskidan samt i fält visat en ”liten” hälta VB. Den ansågs vid besöket initialt ha en markering VB, men efter longering, då den i studien använda mätningen gjordes, bedömdes den som ohalt. Dock sågs en mycket lindrig reaktion vid lågt böjprov HB, vilken ej bedömdes signifikant att gå vidare med. Böjprovet utfördes efter mätningen med Sleip AI. I denna studie utmärker sig hästen som den enda där ingen deltagare ser primärasymmetri på samma ben som Sleip AI gör, vilket är HF. En övervägande majoritet (75 %) av deltagarna har svarat att hästen är halt på HB, vilket får författaren att fundera på vad det är som deltagarna har sett, som gör att de svarat samma ben som sen hade en – om än mycket lindrig – böjreaktion? I Sleip AIs cirkelplot (Bilaga 1, häst 2) ser man att det stora flertalet stegvektorer för frambenen är relativt jämt fördelade mellan belastning HF och frånskjut VF. Dock är det större värden för vektorerna som hamnar på HF och således blir det totala asymmetrivärdet också placerat där. För bakbenen är alla stegvektorerna inom gränsen för symmetri och bakbenen tolkas därmed som symmetriska, men alla de små avvikelser som finns mellan höger och vänster bak indikerar trots allt det högra bakbenet, men med mycket små utslag. Det finns andra rörelsekomponenter som signifikant kan förändras vid bakbens-hälta, så som rotationen av bäckenet sedd bakifrån (Hernlund *et al.*, opubl. accepterat abstract, 2022) vilket kan ha påverkat vad veterinärerna uppfattar. Sammanfattningsvis är asymmetrin i Sleip AI mycket lindrig, hästen är 11 år, saknar skadehistorik och används endast till lättare arbete i skogen och på ridbana, vilket gör att beslutet att påbörja igångsättning verkar vara sunt. Ett påstående som även stöds av att hästen ej återkommit till kliniken tre månader efter besöket.

Häst nummer 2 är inte idealisk för denna studie och borde kanske inte ha ingått eftersom dess asymmetri inte är verifierad med en bedövning. Den belyser dock en av de viktiga frågorna inom rörelseanalys; Hur vet man vad som är friskt, patologiskt eller inom normalvariation? Denna häst går hem, diagnosticerad som frisk. Den har dock en böjprovsreaktion HB och 9 av tolv veterinärer som anser att den är 0,5–2 grader halt på samma ben. Dock säger objektiva mätningar att den är nästintill helt symmetrisk på bakbenen men asymmetrisk på HF. Frambensasym-

metrin kan dock inte anses helt klassisk eller entydig för en frambenshälta då den indikerar belastnings- och frånskjutsasymmetri på olika framben (se Bilaga 1, häst 2). Flera osäkerhetsfaktorer gör det svårt att resonera vidare; man utförde inga böjprov för frambenen och inga diagnostiska anestasier administrerades. Här instämmer författaren med Egan *et al.* (2019) som beskriver att det behövs mer forskning för att bättre förstå det komplexa förhållandet smärta–hälta samt att veta hur man ska skilja mellan funktionell asymmetri och smärtutlöst rörelsestörning. Befintliga tröskelvärden inom objektiv rörelseanalys behöver utvecklas till generella referensintervall, men arbetet är omfattande och komplicerat. Ett steg på vägen är kontinuerliga mätningar av samma individ över tid.

5.1.3. Undersökningen

Av praktiska såväl som smittskyddsmässiga skäl valde sju av 12 veterinärer (58 %) att genomföra undersökningen helt i digital form. För dessa deltagare var det för författaren omöjligt att kontrollera att filmerna spelades upp korrekt samt om instruktionerna var tydliga nog. Vidare fanns viss oro att författaren omedvetet skulle råka påverka deltagarnas bedömningar i de fysiska mötena. Dock ses ingen skillnad i resultat mellan digital och fysisk grupp och de har således behandlats som likvärdiga.

5.1.4. Video

Kvaliteten på videofilmerna får över lag anses godtagbar då alla veterinärer kunnat bedöma alla hästar. Alla veterinärer tyckte dock att det var bäst att bedöma hästarna på originalvideon. En veterinär kommenterade att Sleip AI-videon hade sämre kvalitet och en annan att originalvideon hade bättre skärpa. Flera stycken svarade att de inte var vana att bedöma hälta i slow-motion och därför föredrog originalvideon. En deltagare tittade först på originalvideon, sen på Sleip AI-videon och sen återigen på originalvideon. Med anledning av den tidigare nämnda temporala begränsningen i den mänskliga synen förväntade sig författaren att fler veterinärer skulle föredra Sleip AI-videon men, som en deltagare nämnde, kanske man måste vänja sig för att uppskatta den sänkta uppspelningshastigheten.

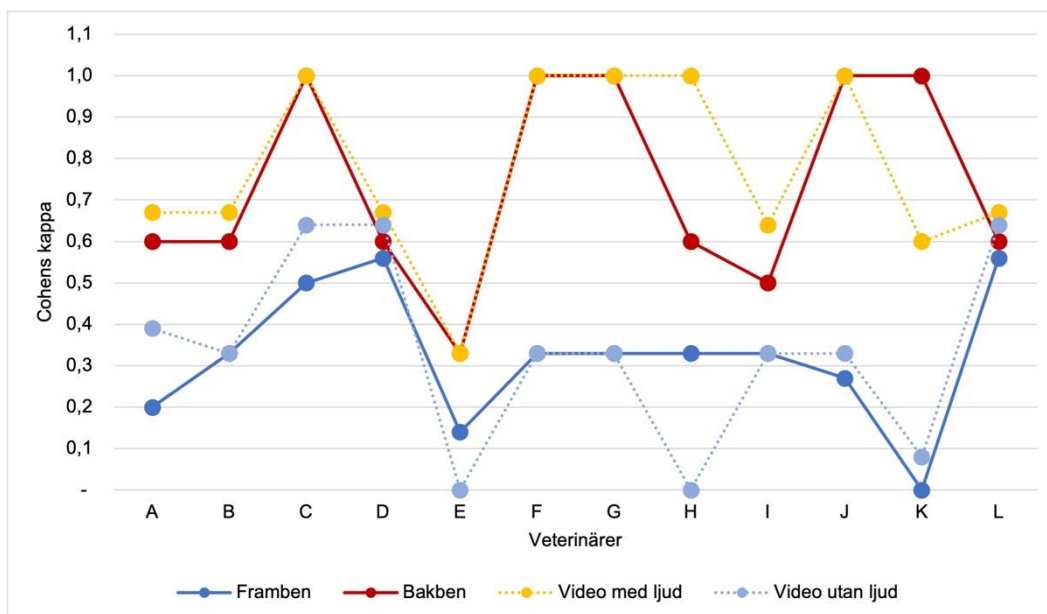
Den kliniska relevansen i att bedöma hälta hos hästar på video, endast travande på rakt spår på ett enda underlag, har ifrågasatts av såväl deltagare i denna studie, som deltagare (Hammarberg 2010) och författare av andra studier (Keegan *et al.* 1998; Hewetson *et al.* 2006). Att inte kunna se hästen från sidan, på böjt spår eller höra ljudet vid hovisättningen gör att resultat från sådana bedömningar, enligt Hewetson *et al.* (2006) ska tolkas med försiktighet när man extrapolerar dem till en klinisk vardag. Fuller *et al.* (2006) konstaterade dock att en hältbedömning utförd live hade god överensstämmelse med videobedömning av samma häst, under förutsättning

att det var samma person som gjorde bedömningen emedan Keegan *et al.* (2010) konkluderade att överensstämmelsen veterinärer emellan var sämre vid videobedömning än vid livebedömning. Vidare kunde Keegan *et al.* (2010), i en studie med 131 hästar, konstatera att överensstämmelsen mellan 2–5 erfarna hästortopeders inte förbättrades av att genomföra en fullständig hältutredning i förhållande till att endast bedöma hästen i trav rakt ut. Fördelarna med – och även anledningen till att videobedömning användes i denna studie – är att olika bedömare, utspridda i landet, kan se samma häst trava upp oändligt antal gånger. För denna studie hade ett realtidförfarande resulterat i avsevärt färre deltagare, men för framtiden är det en mycket intressant utvecklingsmöjlighet. Relevansen i att endast använda trav på rakt spår, menar författaren ändå är betydande då momentet i stort sett utgör basen för alla rörelsekontroller, så väl vid hältutredning, som vid tävlings- och försäljningsbesiktning. En av deltagarna i denna studie upptäcker oftast hältan som sen behandlas på rakt spår, medan flertalet anser att rakt och böjt spår är lika viktiga. Ross (2011b) beskriver att hältan som ses vid trav vid hand vanligtvis är densamma som får hopphästen att vägra men att det naturligtvis finns undantag, vilket konfirmeras av Dyson & Greve (2016). Slutligen anser författaren själv att om veterinärerna inte är överens i den mest grundläggande av rörelsebedömningar är chansen liten att de är det efter mer avancerade undersökningar.

5.1.5. Ljud

En deltagare kommenterar att ljudet behövs för att kunna göra en bra bedömning på video. Kommentaren skapar osäkerhet huruvida deltagaren hörde ljudet på de videofilmer som hade ljud. Deltagarens resultat skiljer sig inte signifikant från andra deltagares, och har inkluderats i studien.

Ljudet vid hovisättningen verkar i denna studie ha betydelse för rörelsebedömningen, då överensstämmelsen för primärhalt ben var högre för videofilmer med ljud än utan, inom veterinärgruppen, mellan veterinärgruppen och Sleip AI samt mellan enskild veterinär och Sleip AI. Dessa slutsatser ska dock tolkas med försiktighet då endast fyra hästar ingick i varje grupp och vissa hästar helt enkelt kan ha varit lättare att bedöma än andra. Resultaten samstämmer även i stort med när man grupperar hästar med primär frambenshälta samt primär bakbenshälta, vilket kan förklaras av att endast två hästar byter grupp med varandra, se figur 18. Förvånande är dock att överensstämmelsen är högre för bakben än framben, vilket skiljer sig från tidigare studier (Fuller *et al.* 2006; Keegan *et al.* 2010, 2013; Hammarberg *et al.* 2016). Vad som ger dessa resultat behöver studeras vidare, kanske i en jämförelse av bedömning av videofilmer av samma hästar med och utan ljud.



Figur 18. Primärhålt. Överensstämmelse mellan enskild veterinär och Sleip AI för hästar grupperade efter frambens- och bakbenschålta samt videofilmer med och utan ljud. I grupperna bakbenschålta och video med ljud är tre hästar identiska. I grupperna frambenschålta och video utan ljud är tre hästar identiska.

5.1.6. Statistik

Kappa ska korrigera för att deltagare av slumpen kan välja samma svar, men det råder delade meningar om denna korrigering görs korrekt (Hammarberg 2010). Vissa hävdar också att kappa ger en falskt låg överensstämmelse i jämförelse med verkligheten. Det diskuteras att kappa ej ska jämföras mellan studier och populationer, men trots detta är det den mest använda metoden i jämförande studier idag och valet föll på kappa även i denna studie.

Det ska noteras att alla grupperingar av hästar innehåller endast 3–6 individer och dessa resultat blir således mindre tillförlitliga trots $p < 0,05$.

5.1.7. Primärhålt

Överensstämmelse inom veterinärgruppen

I veterinärgruppen är överensstämmelsen för vilket ben som är primärhalt för alla hästar acceptabel ($\kappa = 0,45$) och i jämförelse med andra liknande studier god. Resultatet kan jämföras med en studie av Keegan *et al.* (2010), där 2–5 kliniker med i snitt 18,7 års ortopedisk erfarenhet bedömde 131 hästar. När de skulle bedöma huruvida ett ben var halt eller inte efter trav rakt ut var överensstämmelsen acceptabel ($\kappa = 0,44$). Resultatet är nära detsamma som i denna studie, men skillnaden är att deltagarna i denna studie endast fått välja ett ben och för denna metod

i studien av Keegan *et al.* haft möjligheten att peka ut flera halta ben, vilket torde öka överensstämmelsen. I samma studie lät Keegan *et al.* även veterinärerna avgöra om hästen är halt och om så var fallet, vilket ben de skulle börja utreda efter trav rakt ut, longering och böjprov. Då var överensstämmelsen acceptabel ($\kappa = 0,37$) och det är egentligen den frågeställningen som mest liknar denna studies. I studien av Hammarberg *et al.* (2016) bedömde 86 veterinärer 23 lindrigt halta hästar under longering. I frågan vilket ben de skulle börja utreda var överensstämmelsen för alla hästar var precis acceptabel ($\kappa = 0,31$).

I gruppen lindrigt till kraftigt halta hästar är överensstämmelsen god ($\kappa = 0,64$). Dock inte riktigt lika bra som för Keegan *et al.* (2010), där överensstämmelsen var mycket bra för hästar graderade $>1,5$ grader enligt AAEP ($\kappa = 0,86$). Skulle man i denna studie skapa en grupp för hästar graderade $>1,5$ blir det endast tre hästar (1, 7, 4) i den, både om man använder sig av medianen från alla veterinärers gradering och den icke-statistiska modellen för Sleip AI asymmetrivärde och håltgrad. Dessa individer är de tre måttligt–kraftigt halta hästarna och överensstämmelsen för endast dessa blir marginellt högre ($\kappa = 0,70$).

För gruppen ohalter till mycket lindrigt halta hästar är överensstämmelsen dålig ($\kappa = 0,14$). Fem deltagare, två som deltagit i studien fysiskt och tre digitalt, har inte noterat en enda ohalt häst, vilket gör att författaren ifrågasätter studieinstruktionerna. Samtidigt är det i linje med studien av Starke & Oosterlinck (2019), där erfarna veterinärer har en tendens att i större utsträckning än mindre erfarna, se en bakbenshåla, trots att de (animerade) hästarna rör sig helt symmetriskt. Författaren av denna studie funderar trots det om arbetstiteln för studien kan ha påverkat veterinärerna: *Jämförelse av hur **håla** bedöms av applikation baserad på datorseende och av kliniker?* Kan första frågan om primärhalt ben, ha fått dem att tro att alla hästar ska ha en håla och gjort dem mindre benägna att definiera en häst som ohalt? Kan det i en studiesituation finnas en känsla av skam i att riskera inte upptäcka en håla? Dessa faktorer, utreds inte i denna studie men, kan möjligen påverka med vilken inställning en deltagare tar sig an uppgiften. Resultatet stämmer dock väl överens med både Keegan *et al.* (2010), där överensstämmelsen var dålig ($\kappa = 0,23$) för hästar graderade $<1,5$ grader enligt AAEP och Hammarberg *et al.* (2016), som rapporterade dålig överensstämmelse för ohalter ($\kappa = 0,08$). Man kan fråga sig om dessa studier möjligen lider av liknande bias-problem.

Det är intressant att inte en enda deltagare har bedömt en asymmetrisk häst som ohalt. Alla deltagare har skiljt de ohalter från de mycket lindrigt halta och noterat de mycket lindrigt halta hästarna som just halta. Parkes *et al.* (2009) beskriver ju att den mänskliga synen fungerar dåligt för asymmetrier $<25\%$ och Starke & Oosterlinck (2019) konstaterar att det krävs $>25\%$ och $>15\%$ avvikelse för att korrekt identifiera frambens- respektive bakbensasymmetrier. I linje med det är veteri-

närerna alltså inte överens för de fyra hästar som är ohalter–mycket lindrigt halta samt har endast gjort samma bedömning som Sleip AI i 35,4 % fallen. Att jämföra med 79,1 % för resten av hästarna. Man verkar alltså kunna notera den låggradiga asymmetrin, men ha svårt att göra en korrekt identifiering.

För frambenshalta är överensstämmelsen acceptabel ($\kappa = 0,42$) och bakbenshalta bra ($\kappa = 0,60$). Dessa hästgrupper överlappar dock grupperingen för video med och utan ljud, endast undantaget två hästar som byter grupp. Detta faktum kan möjligen bidra till att proportionerna i överensstämmelse avviker från både Keegan *et al.* (2010); bra för framben ($\kappa = 0,51$) och acceptabel för bakben ($\kappa = 0,36$), Keegan *et al.* (2013); acceptabel för både fram- och bakben ($\kappa = 0,37$ respektive 0,31) och Hammarberg *et al.* (2016); acceptabel för framben ($\kappa = 0,33$) och dålig för bakben ($\kappa = 0,11$).

Trots att hästarna i denna studie är utvalda för att representera lika många fram- som bakbenshältor ser veterinärerna nästan dubbelt så många primära bakbenshältor (60 %) som frambenshältor (31,6 %). En veterinär ser endast bakbenshältor. Denna snedfördelning skulle kunna vara en förklaring den högre samstämmigheten för bakbenshältor. Författaren frågar sig om den nedslående överensstämmelsen för just bakbenshältor i de senaste årens forskning, kan ha ökat medvetenheten och påverkat deltagarna att verkligen anstränga sig för att se just asymmetrier för bakbenen.

Överensstämmelse mellan veterinärgruppen och Sleip AI

För 70/80 % av hästarna (beroende på om häst 5 inkluderas eller ej) valde veterinärgruppens majoritet och Sleip AI samma ben som primärhalt. Överensstämmelsen var således bra, gränsande till mycket bra ($\kappa = 0,63/0,75$), vilket är bättre än inom veterinärgruppen. För de ohalter och mycket lindrigt halta är överensstämmelsen också avsevärt bättre ($\kappa = 0,39/0,64$) än inom veterinärgruppen. I övriga grupper ses ett mönster liknande det inom veterinärgruppen där överensstämmelsen är större för bakben än framben, samt för video med ljud än utan. Överensstämmelsen i denna studie är generellt sett högre än vad Keegan *et al.* (2013) beskrev, när de jämförde subjektiv och objektiv bedömning och lät tre erfarna veterinärer och IMUs bedöma 106 hästar. Veterinärerna genomförde en fullständig hältutredning exklusive diagnostiska anestesier och värden från IMUs registrerades på rakt spår. Överensstämmelsen var acceptabel ($\kappa = 0,20\text{--}0,40$) till måttlig ($\kappa = 0,41\text{--}0,60$) för framben och dålig ($\kappa = 0,00\text{--}0,20$) för bakben. De andra studierna (Bragança *et al.* 2020; Reed *et al.* 2020) som jämför objektiv och subjektiv bedömning använder sig inte av överensstämmelse som mått och är därför svåra att jämföra med denna.

Överensstämmelse mellan den enskilde veterinären och gruppen samt Sleip AI

För de enskilda veterinärerna var överensstämmelsen med Sleip AI bra, men varierade kraftigt ($\kappa = 0,15\text{--}0,75$, median $\kappa = 0,51$). Likaså överensstämmelsen med resten av veterinärgruppen var bra med stor variation ($\kappa = 0,14\text{--}0,65$, median $\kappa = 0,55$). Till författarens vetskap finns inga jämförbara publicerade studier. En magisteruppsats av Sandström (2009) rapporterar bra överensstämmelse ($\kappa = 0,70$) mellan en veterinär och Lameness Locator i en studie av 14 hästar. Veterinären har dock bedömt hästarna efter en fullständig hältutredning inklusive diagnostiska anestesier.

I denna studie ses en måttligt positiv korrelation mellan hur bra en veterinär överensstämmer med gruppen och med Sleip AI. Att resultatet blir icke-signifikant i beräkningen av Spearman's rangkorrelationskoefficient ($p = 0,099$) är förväntat då icke-parametriska tester generellt sett är mer konservativa vad gäller signifikans, speciellt för mindre dataset som detta.

5.1.8. Kompensatorisk hälta

Innebörden av uttrycken *kompensatorisk hälta* och *hålttyp* var inte självklara för alla veterinärer i de fysiska mötena, utan förklarades eller beskrevs mer detaljerat ett flertal gånger. Detta gör att man kan anta att samma osäkerhet råder hos resten av Sveriges hästveterinärer. Dessa uttryck behöver alltså bekräftas mer för att kallas allmängiltiga. Detta faktum skapar också en fundering hos författaren; om man inte känner till uttrycket kompensatorisk hälta, känner man då till dess mekanismer? Hur ofta tolkas en höger bakbenshalt häst som höger frambenshalt när en veterinär inte har kunskap om det mest vedertagna kompensatoriska rörelsemönstret, där bakbenshälta orsakat en kompensatorisk ipsilateral frambenshalt? I denna studie har deltagarna exempelvis bedömt en häst (3) med bedövnings-verifierad primär hälta VB och kompensatorisk VF, som primärhalt VF i fyra fall, varav en av veterinärerna har uppgivit att den har en kompensatorisk hälta VB, vilket troligen inte skulle vara slutsatsen om man hade kunskap om det vanligaste kompensationsmönstret vid bakbenshälta.

5.1.9. Håltgradering

Trots att veterinärerna antas vara överens i fyra håltgradskategorier blir överensstämmelsen dålig ($\kappa = 0,24$), vilket i både förfarande och resultat liknar ett examensarbete av Hammarberg (2010). Detta kan jämföras med en studie som inbegrep 20 hästar, bedömda av tre erfarna veterinärer av Fuller *et al.* (2006). Veterinärerna fick se hästarna rakt ut i skritt och trav samt longerade och graderade

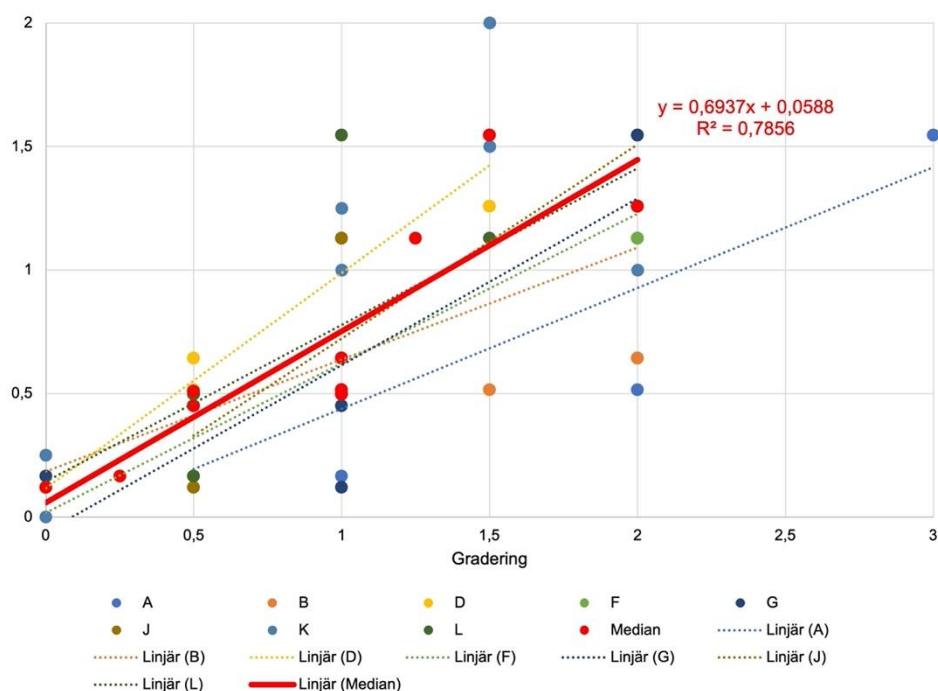
därefter hälta oavsett ben, på en skala 0–10. Överensstämmelsen var acceptabel ($\kappa = 0,41$).

Keegan *et al.* (2013) rapporterade att veterinärernas subjektiva gradering var signifikant positivt associerad med värden från IMUs vid bedömning av 106 hästar. Författarna använder sig av en medeldifferens mellan höger och vänster ben, uppdelat för bak- och framben för den subjektiva graderingen. Som objektiva värden använder de Max-/MinDiff separat men även adderade till en vektorsumma. För framben var överensstämmelse med subjektiv bedömning och förklaringsgrad högst för HDmin ($\kappa = 0,43$, $r^2 = 0,51$). För bakben var överensstämmelsen störst för PDmax ($\kappa = 0,26$) men summa av PDmax och PDmin hade något högre förklaringsgrad ($r^2 = 0,39$). Hardeman *et al.* (2021) undersökte hur 17 olika objektiva mätvärden korrelerade med subjektiv bedömning av huruvida en häst var halt eller endast uppvisade en mindre anmärkning vid försäljningsbesiktning. Författarna konstaterar att MinDiff och *range up difference* (RUD) för huvudet kunde skilja frambenshalta hästar från de som fick en mindre anmärkning vid besiktning. För bakbenshalta kunde MinDiff, MaxDiff, RUD för bäckenet samt *hip hike difference swing* (HHDswing) och *hip hike difference stance* (HHDstance) göra samma sak. Vidare kunde två så kallade hältnummer (*lameness patterns*) skilja både fram- och bakbenshalta från besiktningshästar med små anmärkningar. I mönstret för frambenshalta ingick RUD huvud + MinDiff manke – RUD bäcken och i mönstret för bakbenshalta ingick RUD bäcken + RUD huvud – MinDiff manke.

I denna studie valde författaren att endast jämföra det största asymmetrivärdet (HDmax, HDmin, PDmax, PDmin) för det primärhalta benet med den subjektiva graderingen, trots att det skulle kunna vara så att man vid visuell bedömning snarare uppfattar hela räckvidden mellan minimi- och maximivärde, vilket kanske bäst speglas av en summa av värdet för Min och MaxDiff.

Ett försök att utveckla en statistiskt robust linjär modell för veterinärernas graderingar och Sleip AIs asymmetrivärde gjordes. Som objektiva data provades användning av det största asymmetrivärdet, en summa av HDmin + HDmax eller PDmin + PDmax samt en summa av alla hästens asymmetrivärden. På grund av att svaren inte är tillräckligt samstämmiga och datasetet är för litet, kunde varken en modell för fram- eller bakben fås att fungera. För att i framtiden ta fram en sådan modell behöver varje veterinär gradera fler hästar och gärna samma häst flera gånger för olika grad av hälta, exempelvis före/efter diagnostisk anestesi, vilket gör att man kan korrigera för effekten av häst.

För att ändå visualisera en grov uppskattning av vilket Sleip AI-asymmetrivärde som för denna studiepopulation korresponderar med veterinärernas gradering skapades en preliminär modell där veterinärerna med de fyra mest avvikande k -värdena exkluderades. Författaren valde att exkludera veterinärerna med flackast lutande trendlinje (E och I), det vill säga de två lägsta k -värdena, samt veterinärerna med brantast lutande trendlinje (C och H), det vill säga de två högsta k -värdena, i ett försök att skapa mindre spridning i graderingsbedömningarna. Trots detta kan man, i figur 19, se att det fortfarande är stor spridning i veterinärernas bedömningar. Exempelvis korrelerar ett Sleip AI-asymmetrivärde på 1,00 för en veterinär (D) med 1 grads hälta och för en annan veterinär (A) med nära 2,5 graders hälta.



Figur 19. Gradering primärhåla. Spridningsdiagram med trendlinje för veterinär A, B, D, F, G, J, K, Ls gradering av primärhåla samt en trendlinje för medianvärdet från dessa åtta veterinärers graderingar för varje häst. De fyra veterinärerna (C, E, H, I) med studiens två högsta och två lägsta k -värden exkluderades från denna beräkning.

Trots den relativt stora spridningen, plottades medianvärdet från de kvarvarande åtta veterinärernas graderingar för varje häst i ett spridningsdiagram (röda punkter) med en trendlinje (röd linje) med ekvationen $y = 0,6937x + 0,0588$, $r^2 = 0,79$ (se figur 19). En gradering (x) för varje häst beräknades med denna ekvation, baserat på y = varje hästs Sleip AI-asymmetrivärde (se tabell 10). Dessa beräknade Sleip AI-graderingar jämfördes sen med veterinärernas individuella graderingar i tabell 11 med avseende på skillnad i totalvärde för gradering samt överensstämmelse för gradering av alla hästar.

Tabell 10. Beräknad hältgrad för varje häst. Hältgraden beräknad utifrån trendlinjerekvationen ($y = 0,6937x + 0,0588$, $r^2 = 0,79$) i figur 19, inkluderande medianvärde från veterinär A, B, D, F, G, J, K, Ls gradering av primärhåla för varje häst, x = hältgrad, y = asymmetrivärde. *A, B, D, F, G, J, K, L. **Beräknad från veterinär A, B, D, F, G, J, K, Ls gradering av primärhåla.

Beräknad hältgrad för studiens hästar						
Häst	Sleip AI			Veterinärer*	Beräknad grad	
	Ben	Grad	Värde*	Grad		
				Median**	$y = 0,6937x + 0,0588$	Avrundat
9	Ohalt	Ohalt	$0,12 \pm 0,3$	0	0,0882	0
5	Ohalt	Ohalt	$0,17 \pm 0,1$	0,25	0,1530	0
2	HF	Mkt lindrig	$0,49 \pm 0,4$	1	0,6318	0,5
3 ^L	VB	Mkt lindrig	$0,45 \pm 0,2$	0,5	0,6551	0,5
8	HF	Lindrig	$0,51 \pm 0,2$	0,5	0,6494	0,5
10 ^L	HB	Lindrig	$0,51 \pm 0,1$	1	0,6575	0,5
6	HB	Lindrig	$0,64 \pm 0,1$	1	0,8518	1
1 ^L	HF	Måttlig	$1,26 \pm 0,3$	2	1,7284	1,5
7 ^L	VB	Måttlig	$1,13 \pm 0,2$	1,25	1,5412	1,5
4	VF	Kraftig	$1,55 \pm 0,3$	1,5	2,1433	2
Totalt				9	9,0997	8

Tabell 11. Preliminär jämförelse av veterinärernas totala gradering av primärhåltan och beräknad total avrundad gradering av primärhåltan enligt tabell 10. *För att ta fram värde för gradering baserat på Sleip AIs asymmetrivärden har en linjär modell ($y = 0,6937x + 0,0588$) baserad på veterinärernas (A, B, D, F, G, J, K, L) mediangradering av primärhåltan använts. För modellen har gradering från veterinärer med de två högsta, samt de två lägsta κ -värdena exkluderats. Totalvärde för avrundad beräknad gradering enligt Sleip AI = 8. **Kappa jämför varje individuell veterinärs gradering med beräknad avrundad gradering enligt tabell 10. Kappa är beräknad enligt Cohen.

Gradering primärhåltan. Jämförelse mellan enskild veterinär och beräknad Sleip AI*				
Veterinär	Gradering	Avvikelse total beräknad gradering*	Överensstämmelse beräknad gradering**	Tolkning
	Total		κ	
A	15	+7	-0,06	Dålig
B	11	+3	0,20	Dålig
C	6	-2	0,30	Acceptabel
D	6,5	-1,5	0,58	Bra
E	18	+10	0,13	Dålig
F	11	+3	0,32	Acceptabel
G	11	+3	0,21	Dålig
H	6	-2	0,09	Dålig
I	18	+10	0,03	Dålig
J	9,5	+1,5	0,09	Dålig
K	9,5	+1,5	0,40	Acceptabel
L	8,5	+0,5	0,47	Acceptabel
Median	10,25	+2,25	0,21	Dålig

5.1.10. Användande av objektiv rörelseanalys

Objektiv rörelseanalys är ett relativt nytt verktyg i den kliniska verksamheten och denna studie visar inte på någon skillnad mellan veterinärer som använder det och inte. Man skulle alltså kunna argumentera för att användande objektiv rörelseanalys varken förbättrar eller försämrar förmågan att upptäcka hältor. Man måste dock ta i beaktning att tiden i praxis skiljer avsevärt mellan gruppen som *alltid* använder objektiv rörelseanalys (medelvärde 3 år) och för de som *aldrig* gör det (medelvärde 22,2 år). Veterinärerna med i snitt 3 år i praxis har på mindre än en sjundedel av tiden uppnått genomsnittligt likvärdig överensstämmelse med Sleip AI (*alltid* $\kappa = 0,51$ *aldrig* $\kappa = 0,57$) och veterinärgruppen (*alltid* $\kappa = 0,48$ *aldrig* $\kappa = 0,44$) som veterinärerna med i snitt 22,2 år i praxis. Jämförelserna ska dock endast ses som tendenser då dessa grupper är små och ojämna i storlek.

Författaren spekulerar i att den omedelbara feedbacken dagligt användande av objektiv rörelseanalys ger, möjligen skulle kunna resultera i en snabbare inläring. Bland forskarna råder delade meningar huruvida längre erfarenhet eller specialistutbildning gör veterinärer bättre på att bedöma hältor. Hammarberg *et al.* (2016) konkluderar att överensstämmelsen var dålig ($\kappa = 0,25$) för mindre erfarna hästortopeder (<5 år i praxis) och acceptabel ($\kappa = 0,38$) för de med längre erfarenhet emedan Starke & Oosterlinck (2019) beskriver att år i praxis eller antal fall per månad inte hade någon signifikant effekt på förmågan att korrekt avgöra om en häst var halt eller ohalt.

För de veterinärer som är överens med Sleip kan man fråga sig om det överhuvudtaget finns någon anledning att använda objektiv rörelseanalys. De veterinärer som i sin vardag använder objektiv rörelseanalys vid hältutredning uppger i denna studie att det förbättrar förmågan att hitta och bekräfta låggradig hälta, gör att de snabbare fokuserar på ett ben, visuellt aktualiserar hur hästen rörde sig vid förra besöket samt att det hjälper till att jämföra hästens rörelse före och efter bedövning. Författaren tänker att det även torde stärka veterinären i sin bedömning och i framtiden förenkla journalföring samt uppföljning tillsammans med djurägare, tränare och rehabiliteringspersonal. I framtiden, när de objektiva systemen är mer samstämmiga och etablerade skulle data kunna användas vid remiss och som objektiv fakta vid en försäljning och införsäkring.

För de veterinärer som inte är överens med Sleip AI i fråga om primärhalt ben kan man konstatera att de inte heller är särskilt överens med de andra i veterinärgruppen. Oavsett om dessa individer är konsekventa i sina egna bedömningar men avviker från gruppen och Sleip AI, vilket författaren tolkar som hög precision men låg noggrannhet (*accuracy*) och i denna studie kanske skulle kunna översättas till hög förklaringsgrad (r^2) men ett värde långt ifrån median för k , eller om de har en större spridning i sina bedömningar, det vill säga låg precision, således lågt r^2 , skulle en objektiv analysmetod kunna fungera som ett diagnostiskt hjälpmedel, en digital *second opinion* om man arbetar utan kollegor och i bästa fall en kalibrering av den subjektiva analysen.

Till syvende och sist är frågan ska man nu helt förlita sig på en asymmetriapplikation? Här är författaren enig med van Weeren *et al.* (2017) som jämför introducerandet av objektiv rörelseanalys med användandet av avancerad bilddiagnostik och menar att de kan kategoriseras på samma sätt; som mycket uppskattade och oumbärliga verktyg för kliniskt beslutsfattande. Således är och förblir veterinärens syntes av information från den kliniska undersökningen samt adekvata diagnostiska verktyg det en diagnos ska baseras på.

Eftersom Sleip AI är mobilbaserad och därför mer lättillgänglig än andra objektiva system kanske man kan betänka systemet som en vanlig termometer. Hästägare uppmanas att genom upprepade mätningar ta reda på hästens normaltemperatur. Liksom det skulle vara svårt att göra endast medelst handen mot hästens hud, har många svårt att analysera sin hästs rörelsemönster enbart visuellt. Här finns en plats för ett, för djurägare, lättanvänt verktyg som, liksom en termometer, kan signalera att något ligger utanför hästens normala värden. Detta skulle i förlängningen kunna få fler hästar till veterinär tidigare eller skicka en signal till ryttare eller kusk att minska på träningsmängden eller förändra upplägget.

5.1.11. Konklusion

Slutsatser som författaren drar utifrån denna studie;

1. Eftersom överensstämmelsen mellan veterinärer endast är acceptabel för primärhalt ben samt dålig för kompensatoriskt halt ben samt gradering av hälta, är det orimligt att subjektiv rörelseanalys används som gold standard för att utvärdera andra bedömningsmetoder eller kliniska försök. Det kan inte heller anses vara en robust metod vid återbesök till en annan veterinär eller vid remiss till annan klinik.
2. Överensstämmelsen mellan veterinärer och Sleip AI är bra avseende identifiering av det halta benet.
3. Veterinärer som är överens med Sleip AI med avseende på primärhalt ben är också mer överens med resten av veterinärgruppen. De är dessutom mer konsekventa i sin hältgradering.
4. Veterinärer som är säkra i sin bedömning är mer överens med Sleip AI.
5. Veterinärer som hör ljudet vid hovisättningen är mer överens om vilket ben som är primärhalt.

Denna studie är en jämförelse av subjektiv och objektiv hältbedömning utförd av veterinärer och Sleip AI. Studien utgår från att Sleip AI har gjort en korrekt bedömning då alla lindriga till kraftiga hältor utom en har verifierats med diagnostisk anestesi. För att i framtida studier utarbeta en fungerande modell för sambandet mellan subjektiv och objektiv hältgradering bör varje bedömare gradera fler hästar, gärna samma häst flera gånger exempelvis före/efter en diagnosisk anestesi.

Referenser

- Adair, S., Baus, M., Belknap, J., Bell, R., Boero, M., Bussy, C., Cardenas, F., Casey, T., Castro, J., Davis, W., Erskine, M., Farr, R., Fischer, T., Forbes, B., Ford, T., Genovese, R., Gottschalk, R., Hoge, M., Honnas, C., Hunter, G., Joyce, J., Kaneps, A., Keegan, K., Kramer, J., Lischer, C., Marshall, J., Oosterlinck, M., Radue, P., Redding, R., Reed, S.K., Rick, M., Santschi, E., Schoonover, M., Schramme, M., Schumacher, J., Stephenson, R., Thaler, R., Vedding Neilsen, J. & Wilson, D.A. (2018). Response to Letter to the Editor: Do we have to redefine lameness in the era of quantitative gait analysis. *Equine Veterinary Journal*, 50 (3), 415–417. <https://doi.org/10.1111/evj.12820>
- Adair, S., Baus, M., Bell, R., Boero, M., Bussy, C., Cardenas, F., Casey, T., Castro, J., Davis, W., Erskine, M., Farr, R., Fischer, A., Forbes, B., Ford, T., Genovese, R., Gottschalk, R., Hoge, M., Honnas, C., Hunter, G., Joyce, J., Kaneps, A., Keegan, K., Kramer, J., Labens, R., Lischer, C., Marshall, J., Oosterlinck, M., Radue, P., Redding, R., Reed, S., Rick, M., Santschi, E., Schoonover, M., Schramme, M., Schumacher, J., Stephenson, R., Thaler, R., Nielsen, J.V. & Wilson, D. (2019). Letter to the Editor: A response to ‘What is lameness and what (or who) is the gold standard to detect it?’. *Equine Veterinary Journal*, 51 (2), 270–272. <https://doi.org/10.1111/evj.13043>
- Agria Djurförsäkring (2021). *Hältorna fortsätt vanligaste diagnos*. <https://news.cision.com/se/agria-djurforsakring/r/haltorna-fortsatt-vanligaste-diagnos,c3414094> [2021-09-16]
- Anon (1999). *Guide to Veterinary Services for Horses Shows*. American Association of Equine Practitioners, Lexington.
- Arkell, M., Archer, R.M., Guitian, F.J. & May, S.A. (2006). Evidence of bias affecting the interpretation of the results of local anaesthetic nerve blocks when assessing lameness in horses. *Veterinary Record*, 159 (11), 346–348. <https://doi.org/10.1136/vr.159.11.346>
- Arsham, P.H. (2015). *E-labs Learning Objects. Goodness-of-fit Test: Lilliefors Test for Normality*. JavaScript. <http://home.ubalt.edu/ntsbarsh/Business-stat/otherapplets/Normality.htm> [2021-11-25]
- Aviad, A.D. (1988). The use of the standing force plate as a quantitative measure of equine lameness. *Journal of Equine Veterinary Science*, 8 (6), 460–462. [https://doi.org/10.1016/S0737-0806\(88\)80095-9](https://doi.org/10.1016/S0737-0806(88)80095-9)
- Barrey, E. (1999). Methods, applications and limitations of gait analysis in horses. *The Veterinary Journal*, 157 (1), 7–22. <https://doi.org/10.1053/tvjl.1998.0297>

- Bathe, A.P., Judy, C.E. & Dyson, S. (2018). Letter to the Editor: Do we have to redefine lameness in the era of quantitative gait analysis? *Equine Veterinary Journal*, 50 (2), 273–273. <https://doi.org/10.1111/evj.12791>
- Begg, R.K., Palaniswami, M. & Owen, B. (2005). Support vector machines for automated gait classification. *IEEE Transactions on Biomedical Engineering*, 52 (5), 828–838. <https://doi.org/10.1109/TBME.2005.845241>
- Bragança, F.M.S., Brommer, H., van den Belt, A.J.M., Maree, J.T.M., van Weeren, P.R. & van Oldruitenborgh-Oosterbaan, M.M.S. (2020). Subjective and objective evaluations of horses for fit-to-compete or unfit-to-compete judgement. *The Veterinary Journal*, 257, 105454. <https://doi.org/10.1016/j.tvjl.2020.105454>
- Buchner, H.H.F. (2013). Gait adaptation in lameness. *Equine Locomotion*. 2. ed. Philadelphia: Elsevier, 175
- Buchner, H.H.F., Savelberg, H.H.C.M., Schamhardt, H.C. & Barneveld, A. (1996). Head and trunk movement adaptations in horses with experimentally induced fore- or hindlimb lameness. *Equine Veterinary Journal*, 28 (1), 71–76. <https://doi.org/10.1111/j.2042-3306.1996.tb01592.x>
- Clausen, M., Preisinger, R. & Kalm, E. (1990). Analysis of disease in German warm blooded breeding horses. *Züchtungskunde*, 62, 167–178
- Di Tulio, K., Trolinger-Meadows, D. & Castro, J.R. (2017). *Equine Lameness & Equinosis Q–Diagnostic Tool for Objective Gait Assessment*. University of Tennessee College of Veterinary Medicine.
- Dolatabadi, E., Taati, B. & Mihailidis, A. (2017). An automated classification of pathological gait using unobtrusive sensing technology. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 25 (12), 2336–2346. <https://doi.org/10.1109/TNSRE.2017.2736939>
- Dyson, S. (2011). Can lameness be graded reliably? *Equine Veterinary Journal*, 43 (4), 379–382. <https://doi.org/10.1111/j.2042-3306.2011.00391.x>
- Dyson, S. (2014). Recognition of lameness: Man versus machine. *The Veterinary Journal*, 201 (3), 245–248. <https://doi.org/10.1016/j.tvjl.2014.05.018>
- Dyson, S. (2016). Evaluation of poor performance in competition horses: A musculoskeletal perspective. Part 1: Clinical assessment. *Equine Veterinary Education*, 28 (5), 284–293. <https://doi.org/10.1111/eve.12426>
- Dyson, S. & Greve, L. (2016). Subjective gait assessment of 57 sports horses in normal work: a comparison of the response to flexion tests, movement in hand, on the lunge, and ridden. *Journal of Equine Veterinary Science*, 38, 1–7. <https://doi.org/10.1016/j.jevs.2015.12.012>
- Egan, S., Brama, P. & McGrath, D. (2019). Research trends in equine movement analysis, future opportunities and potential barriers in the digital age: A scoping review from 1978 to 2018. *Equine Veterinary Journal*, 51 (6), 813–824. <https://doi.org/10.1111/evj.13076>
- Egenvall, A., Penell, J.C., Bonnett, B.N., Olson, P. & Pringle, J. (2006). Mortality of Swedish horses with complete life insurance between 1997 and 2000: variations with

- sex, age, breed and diagnosis. *Veterinary Record*, 158 (12), 397–406.
<https://doi.org/10.1136/vr.158.12.397>
- Egenvall, A., Tranquille, C.A., Lönnell, A.C., Bitschnau, C., Oomen, A., Hernlund, E., Montavon, S., Franko, M.A., Murray, R.C., Weishaupt, M.A., Weeren, van R. & Roepstorff, L. (2013). Days-lost to training and competition in relation to workload in 263 elite show-jumping horses in four European countries. *Preventive Veterinary Medicine*, 112 (3–4), 387–400. <https://doi.org/10.1016/j.prevetmed.2013.09.013>
- Fredricson, I. & Drevemo, S. (1971). A new method of investigating equine locomotion. *Equine Veterinary Journal*, 3 (4), 137–140. <https://doi.org/10.1111/j.2042-3306.1971.tb04456.x>
- Fuller, C.J., Bladon, B.M., Driver, A.J. & Barr, A.R.S. (2006). The intra- and inter-assessor reliability of measurement of functional outcome by lameness scoring in horses. *The Veterinary Journal*, 171 (2), 281–286.
<https://doi.org/10.1016/j.tvjl.2004.10.012>
- Greve, L. & Dyson, S.J. (2014). The interrelationship of lameness, saddle slip and back shape in the general sports horse population. *Equine Veterinary Journal*, 46 (6), 687–694. <https://doi.org/10.1111/evj.12222>
- Haarmeier, T. & Thier, P. (2006). Detection of speed changes during pursuit eye movements. *Experimental Brain Research*, 170 (3), 345–357.
<https://doi.org/10.1007/s00221-005-0216-6>
- Hammarberg, M. (2010). *Erfarna hästpraktiserande veterinärers subjektiva bedömning av hältor hos hästar som longeras i trav*. (Examensarbete 2010:82). Sveriges lantbruksuniversitet. Veterinärprogrammet.
<http://urn.kb.se/resolve?urn=urn:nbn:se:slu:epsilon-6-417>
- Hammarberg, M., Egenvall, A., Pfau, T. & Rhodin, M. (2016). Rater agreement of visual lameness assessment in horses during lungeing. *Equine Veterinary Journal*, 48 (1), 78–82. <https://doi.org/10.1111/evj.12385>
- Hardeman, A.M., Egenvall, A., Serra Bragança, F.M., Koene, M.H.W., Swagemakers, J., Roepstorff, L., Weeren, R. & Byström, A. (2021). Movement asymmetries in horses presented for prepurchase or lameness examination. *Equine Veterinary Journal*, evj.13453. <https://doi.org/10.1111/evj.13453>
- Hardeman, A.M., Serra Bragança, F.M., Swagemakers, J.H., Weeren, P.R. & Roepstorff, L. (2019). Variation in gait parameters used for objective lameness assessment in sound horses at the trot on the straight line and the lunge. *Equine Veterinary Journal*, 51 (6), 831–839. <https://doi.org/10.1111/evj.13075>
- Heldt, T., Macuchova, Z., Alnyme, O. & Andersson, H. (2018). *Samhällsekonomiska effekter av hästnäringen - Skattningar baserat på en B.I. – modell av hästnäringen för 2016*. (2018:04). Dalarna University: Ed. Fleyeh, H., Working papers in transport, tourism, information technology and microdata, analysis, 2018_04. Access via Hästnäringens Nationella Stiftelse:
<https://hastnaringen.se/app/uploads/2018/11/samhallsekonomiska-effekter-av-hastnaringen.pdf>

- Hernlund, E., Byström, A., Nyström, A., Serra Bragança, F., Rhodin, M., Engell, M.T. & Almlöf, M. (opublicerat abstrakt, acceperat till konferens ICEEP 2022). Comparison of a computer vision application versus optical motion capture for lameness quantification.
- Hernlund, E., Engell, M.T., Persson-Sjodin, E., Rhodin, M., Haubro Andersen, P., Serra Bragança, F.M., Dolvik, N.I. & Byström, A. (2021). Changes in pelvic axial rotation following diagnostic analgesia in hindlimb lame horses. *Proceedings of International Conference on Equine Exercise Physiology*, Uppsala, 2021.
- Hewetson, M., Christley, R.M., Hunt, I.D. & Voute, L.C. (2006). Investigations of the reliability of observational gait analysis for the assessment of lameness in horses. *Veterinary Record*, 158 (25), 852–858. <https://doi.org/10.1136/vr.158.25.852>
- Hoerdemann, M., Smith, R.L. & Hosgood, G. (2017). Duration of action of mepivacaine and lidocaine in equine palmar digital perineural blocks in an experimental lameness model. *Veterinary Surgery*, 46 (7), 986–993. <https://doi.org/10.1111/vsu.12689>
- Huang, L., Pashler, H. & Junge, J.A. (2004). Are there capacity limitations in symmetry perception? *Psychonomic Bulletin & Review*, 11 (5), 862–869. <https://doi.org/10.3758/BF03196713>
- Kaneene, J.B., Ross, W.A. & Miller, R. (1997). The Michigan equine monitoring system. II. Frequencies and impact of selected health problems. *Preventive Veterinary Medicine*, 29 (4), 277–292. [https://doi.org/10.1016/S0167-5877\(96\)01080-X](https://doi.org/10.1016/S0167-5877(96)01080-X)
- Keegan, K., Wilson, D.A., Wilson, D.J., Smith, B., Gughan, E.M., Pleasant, R.S., Lillich, J.D., Kramer, J., Howard, R.D., Bacon-Miller, C., Davis, E.G., May, K.A., Cheramie, H.S., Valentino, W.L. & van Harreveld, P.D. (1998). Evaluation of mild lameness in horses trotting on a treadmill by clinicians and interns or residents and correlation of their assessments with kinematic gait analysis. *American Journal of Veterinary Research*, 59 (11), 1370–1377
- Keegan, K.G. (2007). Evidence-based lameness detection and quantification. *Veterinary Clinics of North America: Equine Practice*, 23 (2), 403–423. <https://doi.org/10.1016/j.cveq.2007.04.008>
- Keegan, K.G. (2018). Breaking down the AIDE algorithms. *Eye on Objectivity*. 2018 (8), 5-15. <https://www.flipsnack.com/equinosis/eoo-august-2018.html>
- Keegan, K.G., Dent, E.V., Wilson, D.A., Janicek, J., Kramer, J., Lacarrubba, A., Walsh, D.M., Cassells, M.W., Esther, T.M., Schiltz, P., Frees, K.E., Wilhite, C.L., Clark, J.M., Pollitt, C.C., Shaw, R. & Norris, T. (2010). Repeatability of subjective evaluation of lameness in horses: Repeatability of subjective evaluation of lameness in horses. *Equine Veterinary Journal*, 42 (2), 92–97. <https://doi.org/10.2746/042516409X479568>
- Keegan, K.G., Kramer, J., Yonezawa, Y., Maki, H., Pai, P.F., Dent, E.V., Kellerman, T.E., Wilson, D.A. & Reed, S.K. (2011). Assessment of repeatability of a wireless, inertial sensor-based lameness evaluation system for horses. *American Journal of Veterinary Research*, 72 (9), 1156–1163. <https://doi.org/10.2460/ajvr.72.9.1156>
- Keegan, K.G., MacAllister, C.G., Wilson, D.A., Gedon, C.A., Kramer, J., Yonezawa, Y., Maki, H. & Pai, P.F. (2012). Comparison of an inertial sensor system with a stationary

- force plate for evaluation of horses with bilateral forelimb lameness. *American Journal of Veterinary Research*, 73 (3), 368–374.
<https://doi.org/10.2460/ajvr.73.3.368>
- Keegan, K.G., Wilson, D.A., Kramer, J., Reed, S.K., Yonezawa, Y., Maki, H., Pai, P.F. & Lopes, M.A.F. (2013). Comparison of a body-mounted inertial sensor system-based method with subjective evaluation for detection of lameness in horses. *American Journal of Veterinary Research*, 74 (1), 17–24. <https://doi.org/10.2460/ajvr.74.1.17>
- Kelmer, G., Keegan, K.G., Kramer, J., Wilson, D.A., Pai, F.P. & Singh, P. (2005). Computer-assisted kinematic evaluation of induced compensatory movements resembling lameness in horses trotting on a treadmill. *American Journal of Veterinary Research*, 66 (4), 646–655. <https://doi.org/10.2460/ajvr.2005.66.646>
- Khera, P. & Kumar, N. (2020). Role of machine learning in gait analysis: A review. *Journal of Medical Engineering & Technology*, 44 (8), 441–467.
<https://doi.org/10.1080/03091902.2020.1822940>
- Kramer, J., Keegan, K.G., Kelmer, G. & Wilson, D.A. (2004). Objective determination of pelvic movement during hind limb lameness by use of a signal decomposition method and pelvic height differences. *American Journal of Veterinary Research*, 65 (6), 741–747. <https://doi.org/10.2460/ajvr.2004.65.741>
- Kutilek, P. & Hozman, J. (2011). Prediction of lower extremities movement using characteristics of angle-angle diagrams and artificial intelligence. *Proceedings of the 3rd International Conference on E-Health and Bioengineering*, EHB 2011, 143–147.
- Li, C., Ghorbani, N., Broomé, S., Rashid, M., Black, M.J., Hernlund, E., Kjellström, H. & Zuffi, S. (2021). hSMAL: Detailed Horse Shape and Pose Reconstruction for Motion Pattern Recognition. *arXiv:2106.10102 [cs]*, <http://arxiv.org/abs/2106.10102> [2021-10-26]
- Liautard, A.F.A. (1888). *Lameness of Horses and Diseases of the Locomotory Apparatus*. New York: W. R. Jenkins.
- Loomans, J.B.A., Stolk, P.W.Th., van Weeren, P.R., Vaarkamp, H. & Barneveld, A. (2007). A survey of the workload and clinical skills in current equine practices in The Netherlands. *Equine Veterinary Education*, 19 (3), 162–168.
<https://doi.org/10.2746/095777307X186875>
- Maliye, S. & Marshall, J.F. (2016). Objective assessment of the compensatory effect of clinical hind limb lameness in horses: 37 cases (2011–2014). *Journal of the American Veterinary Medical Association*, 249 (8), 940–944.
<https://doi.org/10.2460/javma.249.8.940>
- Maliye, S., Voute, L., Lund, D. & Marshall, J.F. (2013). An inertial sensor-based system can objectively assess diagnostic anaesthesia of the equine foot: Inertial sensor-based objective analysis of diagnostic anaesthesia of the foot. *Equine Veterinary Journal*, 45, 26–30. <https://doi.org/10.1111/evj.12158>
- Maliye, S., Voute, L.C. & Marshall, J.F. (2015). Naturally-occurring forelimb lameness in the horse results in significant compensatory load redistribution during trotting. *The Veterinary Journal*, 204 (2), 208–213. <https://doi.org/10.1016/j.tvjl.2015.03.005>

- May, S.A. & Wyn-Jones, G. (1987). Identification of hindleg lameness. *Equine Veterinary Journal*, 19 (3), 185–188. <https://doi.org/10.1111/j.2042-3306.1987.tb01371.x>
- McCracken, M.J., Kramer, J., Keegan, K.G., Lopes, M., Wilson, D.A., Reed, S.K., LaCarrubba, A. & Rasch, M. (2012). Comparison of an inertial sensor system of lameness quantification with subjective lameness evaluation: Comparison of inertial system with subjective lameness evaluation. *Equine Veterinary Journal*, 44 (6), 652–656. <https://doi.org/10.1111/j.2042-3306.2012.00571.x>
- Merkens, H.W. & Schamhardt, H.C. (1988). Evaluation of equine locomotion during different degrees of experimentally induced lameness II: Distribution of ground reaction force patterns of the concurrently loaded limbs. *Equine Veterinary Journal*, 20 (6), 107–112. <https://doi.org/10.1111/j.2042-3306.1988.tb04656.x>
- Morris, E.A. & Seeherman, H.J. (1987). Redistribution of ground reaction forces in experimentally induced equine carpal lameness. *Equine Exercise Physiology* 2, 1987, 553–563
- Morris, E.A. & Seeherman, H.J. (1991). Clinical evaluation of poor performance in the racehorse: the results of 275 evaluations. *Equine Veterinary Journal*, 23 (3), 169–174. <https://doi.org/10.1111/j.2042-3306.1991.tb02749.x>
- Mukaka, M.M. (2012). Statistics Corner: A guide to appropriate use of Correlation coefficient in medical research. *Malawi Medical Journal*, 24 (3), 69–71
- Mullard, J., Berger, J.M., Ellis, A.D. & Dyson, S. (2017). Development of an ethogram to describe facial expressions in ridden horses (FEReq). *Journal of Veterinary Behavior*, 18, 7–12. <https://doi.org/10.1016/j.jveb.2016.11.005>
- Müller-Quirin, J., Dittmann, M.T., Roepstorff, C., Arpagaus, S., Latif, S.N. & Weishaupt, M.A. (2020). Riding soundness - comparison of subjective with objective lameness assessments of owner-sound horses at trot on a treadmill. *Journal of Equine Veterinary Science*, 95, 103314. <https://doi.org/10.1016/j.jevs.2020.103314>
- Nationalencyklopedin (2021). *Artificiell intelligens*. <http://www.ne.se/uppslagsverk/encyklopedi/lång/artificiell-intelligens> [2021-10-27]
- Nyman, S. (2021). Frågor om statistik för hästveterinärer till Hästsektionens representant i Sveriges veterinärförbund. *Personligt meddelande*.
- Nyström, A. (2021). Hur fungerar Sleip AI? *Personligt meddelande*.
- Parkes, R.S.V., Weller, R., Groth, A.M., May, S. & Pfau, T. (2009). Evidence of the development of ‘domain-restricted’ expertise in the recognition of asymmetric motion characteristics of hindlimb lameness in the horse. *Equine Veterinary Journal*, 41 (2), 112–117. <https://doi.org/10.2746/042516408X343000>
- Penell, J.C., Egenvall, A., Bonnett, B.N., Olson, P. & Pringle, J. (2005). Specific causes of morbidity among Swedish horses insured for veterinary care between 1997 and 2000. *Veterinary Record*, 157 (16), 470–477. <https://doi.org/10.1136/vr.157.16.470>
- Persson-Sjodin, E., Hernlund, E., Pfau, T., Haubro Andersen, P., Holm Forsström, K. & Rhodin, M. (2019). Effect of meloxicam treatment on movement asymmetry in riding

- horses in training. (Katz, L. M., red.) *PLOS ONE*, 14 (8), e0221117.
<https://doi.org/10.1371/journal.pone.0221117>
- Pfau, T., Boulton, H., Davis, H., Walker, A. & Rhodin, M. (2016). Agreement between two inertial sensor gait analysis systems for lameness examinations in horses. *Equine Veterinary Education*, 28 (4), 203–208. <https://doi.org/10.1111/eve.12400>
- Pfau, T., Spicer-Jenkins, C., Smith, R.K., Bolt, D.M., Fiske-Jackson, A. & Witte, T.H. (2014). Identifying optimal parameters for quantification of changes in pelvic movement symmetry as a response to diagnostic analgesia in the hindlimbs of horses. *Equine Veterinary Journal*, 46 (6), 759–763. <https://doi.org/10.1111/evj.12220>
- Priester, W.A. (1970). A summary of diagnoses in the ox, horse, dog and cat from 12 veterinary school clinics in the U. S. and Canada. *Veterinary Record*, 86(22), 654–658. <https://doi.org/10.1136/vr.86.22.654>
- Reed, S.K., Kramer, J., Thombs, L., Pitts, J.B., Wilson, D.A. & Keegan, K.G. (2020). Comparison of results for body-mounted inertial sensor assessment with final lameness determination in 1,224 equids. *Journal of the American Veterinary Medical Association*, 256 (5), 590–599. <https://doi.org/10.2460/javma.256.5.590>
- Rhodin, M., Egenvall, A., Haubro Andersen, P. & Pfau, T. (2017). Head and pelvic movement asymmetries at trot in riding horses in training and perceived as free from lameness by the owner. (Lor, J. J., red.) *PLOS ONE*, 12 (4), e0176253.
<https://doi.org/10.1371/journal.pone.0176253>
- Rhodin, M., Pfau, T., Roepstorff, L. & Egenvall, A. (2013). Effect of lungeing on head and pelvic movement asymmetry in horses with induced lameness. *The Veterinary Journal*, 198, e39–e45. <https://doi.org/10.1016/j.tvjl.2013.09.031>
- Roethlisberger Holm, K. (2012). Rörelsesystemet del II: Undersökning i rörelse. *Besiktning av häst*. 1. uppl. Stockholm: Sveriges veterinärförbund
- Ross, M.W. (2011a). Lameness in horses: Basic facts before starting. I: Ross, M.W. & Dyson, S.J. (red.) *Diagnosis and Management of Lameness in the Horse*. 2. uppl. St. Louis, Mo: Elsevier Saunders
- Ross, M.W. (2011b). Movement. I: Ross, M.W. & Dyson, S.J. (red.) *Diagnosis and Management of Lameness in the Horse*. 2. uppl. St. Louis, Mo: Elsevier Saunders
- RStudio (2021). *RStudio*. Version: Build 372. RStudio, PBC. [2021-11-15]
- Sandström, K. (2009). *Jämförelse mellan subjektiv bedömning och en objektiv bedömningsmetod av hältor på rakt spår*. (Examensarbete 2010:2). Sveriges lantbruksuniversitet, Veterinärprogrammet.
<http://urn.kb.se/resolve?urn=urn:nbn:se:slu:epsilon-6-102>
- Schumacher, J., Schumacher, J., Schramme, M. & Moyer, W. (2014). Review of mistakes that can be made when interpreting the results of diagnostic analgesia during a lameness examination. *Proceedings of the AAEP Annual Convention*, Salt Lake City, UT, USA, (Vol 60), 26–37
- Serra Bragança, F.M., Rhodin, M. & van Weeren, P.R. (2018). On the brink of daily clinical application of objective gait analysis: What evidence do we have so far from

- studies using an induced lameness model? *The Veterinary Journal*, 234, 11–23.
<https://doi.org/10.1016/j.tvjl.2018.01.006>
- Social Science Statistics (2021a). *Pearson Correlation Coefficient Calculator*.
<https://www.socscistatistics.com/tests/pearson/default2.aspx> [2021-11-25]
- Social Science Statistics (2021b). *Spearman's Rho Calculator*.
<https://www.socscistatistics.com.com/tests/spearman/default2.aspx> [2021-11-25]
- Starke, S.D. & May, S.A. (2017). Veterinary student competence in equine lameness recognition and assessment: a mixed methods study. *Veterinary Record*, 181 (7), 168–168. <https://doi.org/10.1136/vr.104245>
- Starke, S.D., May, S.A. & Pfau, T. (2015). Understanding hind limb lameness signs in horses using simple rigid body mechanics. *Journal of Biomechanics*, 48 (12), 3323–3331. <https://doi.org/10.1016/j.jbiomech.2015.06.019>
- Starke, S.D. & Oosterlinck, M. (2019). Reliability of equine visual lameness classification as a function of expertise, lameness severity and rater confidence. *Veterinary Record*, 184 (2), 63–63. <https://doi.org/10.1136/vr.105058>
- Stashak, T. (1987). Diagnosis of lameness. I: *Adams' Lameness in Horses*. Philadelphia.: Lea and Febiger, 100–156
- Takeda, I., Yamada, A. & Onodera, H. (2021). Artificial Intelligence-Assisted motion capture for medical applications: a comparative study between markerless and passive marker motion capture. *Computer Methods in Biomechanics and Biomedical Engineering*, 24 (8), 864–873. <https://doi.org/10.1080/10255842.2020.1856372>
- Tóth, F., Schumacher, J., Schramme, M.C. & Hecht, S. (2014). Effect of anesthetizing individual compartments of the stifle joint in horses with experimentally induced stifle joint lameness. *American Journal of Veterinary Research*, 75 (1), 19–25.
<https://doi.org/10.2460/ajvr.75.1.19>
- Wang, Y., Li, J., Zhang, Y. & Sinnott, R.O. (2021). Identifying lameness in horses through deep learning. *Proceedings of SAC '21: The 36th ACM/SIGAPP Symposium on Applied Computing*, Virtual Event Republic of Korea, mars 22 2021. 976–985. Virtual Event Republic of Korea: ACM. <https://doi.org/10.1145/3412841.3441973>
- Warner, S.M., Koch, T.O. & Pfau, T. (2010). Inertial sensors for assessment of back movement in horses during locomotion over ground: Inertial sensors for back movement. *Equine Veterinary Journal*, 42, 417–424. <https://doi.org/10.1111/j.2042-3306.2010.00200.x>
- Weeren, P.R., Pfau, T., Rhodin, M., Roepstorff, L., Serra Bragança, F. & Weishaupt, M.A. (2017). Do we have to redefine lameness in the era of quantitative gait analysis? *Equine Veterinary Journal*, 49 (5), 567–569. <https://doi.org/10.1111/evj.12715>
- Weeren, P.R. (2013). History of locomotor system. I: Back, W. & Clayton, H. *Equine Locomotion*. 2. uppl. St. Louis, Mo: Elsevier Saunders, 1–30
- Weishaupt, M.A., Wiestner, T., Hogg, H.P., Jordan, P. & Auer, J.A. (2006). Compensatory load redistribution of horses with induced weight-bearing forelimb lameness trotting on a treadmill. *The Veterinary Journal*, 171 (1), 135–146.
<https://doi.org/10.1016/j.tvjl.2004.09.004>

- Weishaupt, M.A., Wiestner, T., Hogg, H.P., Jordan, P. & Auer, J.A. (2010).
Compensatory load redistribution of horses with induced weightbearing hindlimb
lameness trotting on a treadmill. *Equine Veterinary Journal*, 36 (8), 727–733.
<https://doi.org/10.2746/0425164044848244>
- Wikipedia (2021). *Datorseende/Computer vision*.
<https://sv.wikipedia.org/wiki/Datorseende>
https://en.wikipedia.org/wiki/Computer_vision [2021-10-26]
- Wlosinska, J. (2019). *Feminiseringen av veterinäryrket ur ett svenskt perspektiv*.
(Grundnivå G2E). Sveriges lantbruksuniversitet. Veterinärprogrammet.
<http://urn.kb.se/resolve?urn=urn:nbn:se:slu:epsilon-s-10417> [2021-11-26]
- Wyn-Jones, G. (1988). *Equine Lameness*. Oxford: Blackwell Scientific.

Tack

Jag vill först och främst rikta ett stort, varmt tack till de tolv veterinärer som gjort denna studie möjlig genom att generöst bidra med sin kunskap, tid och sitt engagemang.

Jag vill även tacka min handledare Elin Hernlund för att du är en konstant inspirationskälla, ger precis lagom mycket guidning och för att du alltid bjuder på din kunskap utan att få en att känna sig korkad och min biträdande handledare Anna Byström för outhärlig och outtröttlig vägledning i statistikens och vetenskapens värld. Jag vill också tacka teamet på Sleip AI för er aldrig sinande energi och tekniska dygnet-runt-support.

Slutligen vill jag tacka min fästman Wilhelm Laska för att du startade allt, och har stöttat mig hela vägen igenom.

Populärvetenskaplig sammanfattning

Hälta är hästens största medicinska problem världen över och sjukdomar i rörelseapparaten är den vanligaste anledningen till att svenska hästar behandlas av en veterinär. För att lyckas med behandling av hälta krävs i första hand att veterinären kan göra en korrekt identifiering av vilket ben som är halt.

Tidigare forskning har visat att veterinärers bedömning av hälta, även kallad subjektiv rörelseanalys, lider av flera brister. Veterinärerna är inte överens om hur halt hästen är eller vilket ben hästen är halt på. Vid hältor av låg grad och bakbenshältor är samstämmigheten i vissa studier inte mycket bättre än slumpen. Diagnostiska bedömningar misstolkas på olika sätt och bedömningen av dessa kan färgas av att veterinären vet att hästen bedövats och därför kan anse att hästen är mindre halt.

Tekniska metoder för att kvantifiera rörelseasymmetri, även kallat objektiv rörelseanalys, har länge varit ett fokus för ortopedisk forskning och har idag utvecklats tillräckligt långt för att användas i praktiskt kliniskt arbete. Idag existerar ingen godkänd och kontrollerad skala för hältgradering i något av de objektiva systemen utan endast ett tröskelvärde som ska skilja halt från ohalt.

Denna studie jämförde subjektiv och objektiv rörelseanalys av 10 hästar i trav på rakt spår. Bedömningen avsåg vilket ben som var halt samt av vilken grad och typ hältan var. Hästarna hade valts ut för att representera ett symmetriskt rörelsemönster samt en ökande grad av hälta jämnt fördelat mellan fram- och bakben. Lindrig–kraftig hälta hade verifierats med en positiv diagnostisk bedövning.

Den subjektiva analysen utfördes på video, av 12 hästveterinärer. Den objektiva analysen utfördes av Sleip AI, en mobil applikation för markörlös optisk *motion capture*, vilket är en ny typ av AI-assisterad objektiv teknik för rörelseanalys. Applikationen identifierar hästens kroppsdelar med datorseende och analyserar dess rörelsebanor med hjälp av maskininlärning. Veterinärerna besvarade även frågor om erfarenhet, säkerhet och användande av objektiv rörelseanalys i en enkät.

Samstämmigheten analyserades både inom veterinärgruppen och mellan gruppen och Sleip AI. Vidare analyserades också varje enskild veterinärs samstämmighet med gruppen samt med Sleip AI.

Metoden som användes heter Fleiss' och Cohens kappa (κ) resulterar i ett tal -1–1. Talet 1 betyder perfekt överensstämmelse, då man valt samma alternativ i alla frågor, talet -1 tolkas som att man inte valt samma alternativ i någon fråga och talet 0 betyder att man har samma överensstämmelse som slumpen. I denna studie tolkades överensstämmelsen som dålig för $\kappa \leq 0,3$, acceptabel för $\kappa = 0,31–0,5$, bra för $\kappa = 0,51–0,8$ och mycket bra för $\kappa > 0,8$. Korrelation mellan två företeelser beräknades med Pearsons korrelationskoefficient (r) och Spearmans rangkorrelationskoefficient (r_s) och ju högre värde desto starkare samband ses mellan företeelserna.

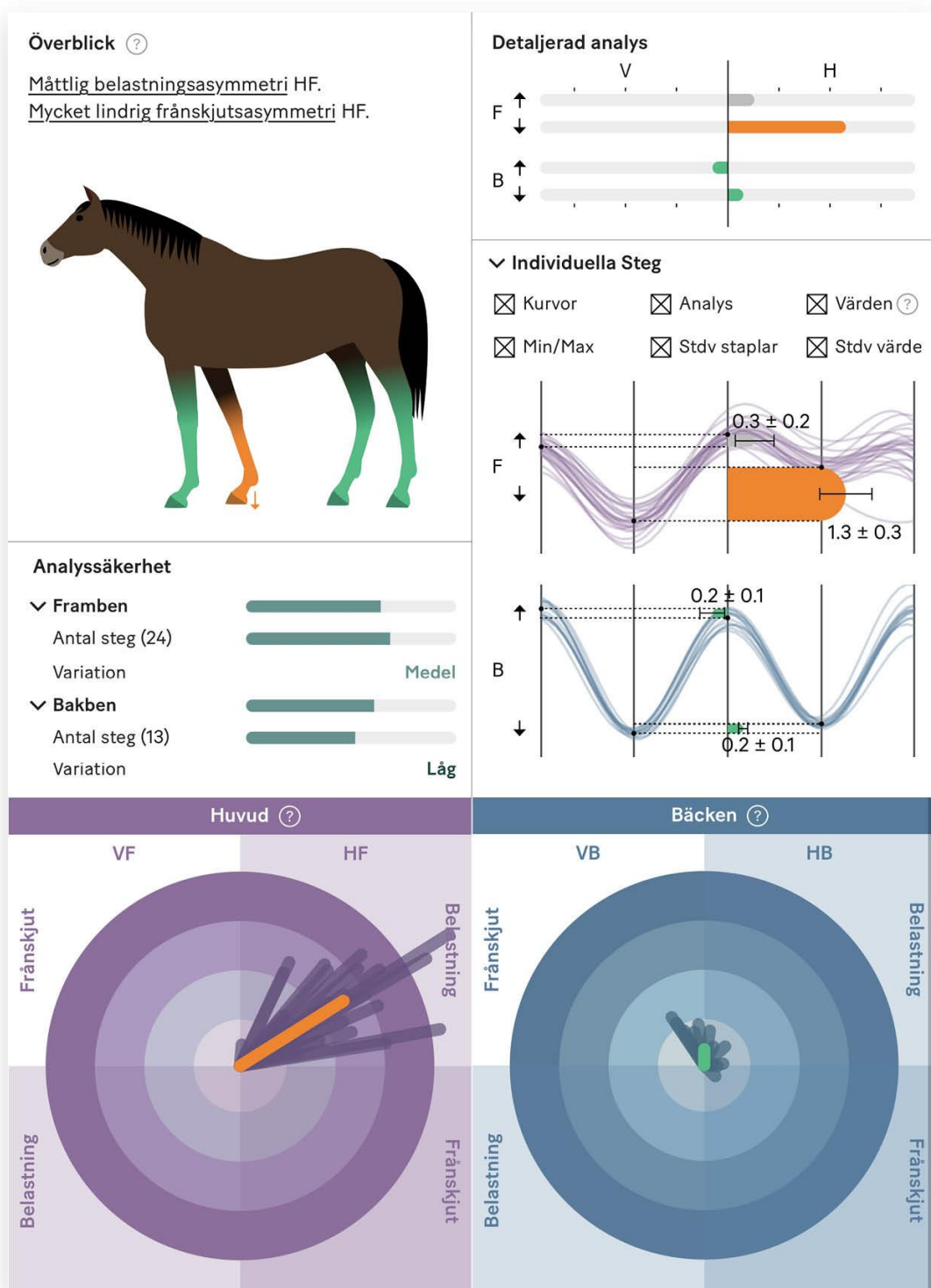
Graden av konsensus inom veterinärgruppen skiljer sig inte nämnvärt från tidigare studier. När veterinärerna skulle ange vilket ben som var halt var överensstämmelsen acceptabel ($\kappa = 0,45$), men för gradering av hältan var den dålig ($\kappa = 0,24$).

Samstämmigheten mellan veterinärgruppen och Sleip AI var bra ($\kappa = 0,63/0,75$), vilket är bättre än man sett i en tidigare studie. Det fanns ett samband där veterinärer som var överens med Sleip AI med avseende på vilket ben som var halt också var mer överens med resten av veterinärgruppen. De var dessutom säkrare i sin bedömning och mer konsekventa i sin hältgradering. Varken hur lång tid de varit yrkesverksamma, antal fall/år eller huruvida de använde objektiv rörelseanalys sågs ha något samband med den enskilde veterinärens överensstämmelse med gruppen eller Sleip AI.

På grund av att det inte var tillräckligt många hästar i studien samt dålig samstämmighet för hältgradering mellan veterinärerna kan denna studie inte presentera en fungerande modell för sambandet mellan veterinärers hältgradering och Sleip AIs gradering av asymmetri. För att i framtida studier utarbeta en fungerande modell bör varje bedömare gradera fler hästar, gärna samma häst flera gånger exempelvis före/efter diagnosisk bedövning.

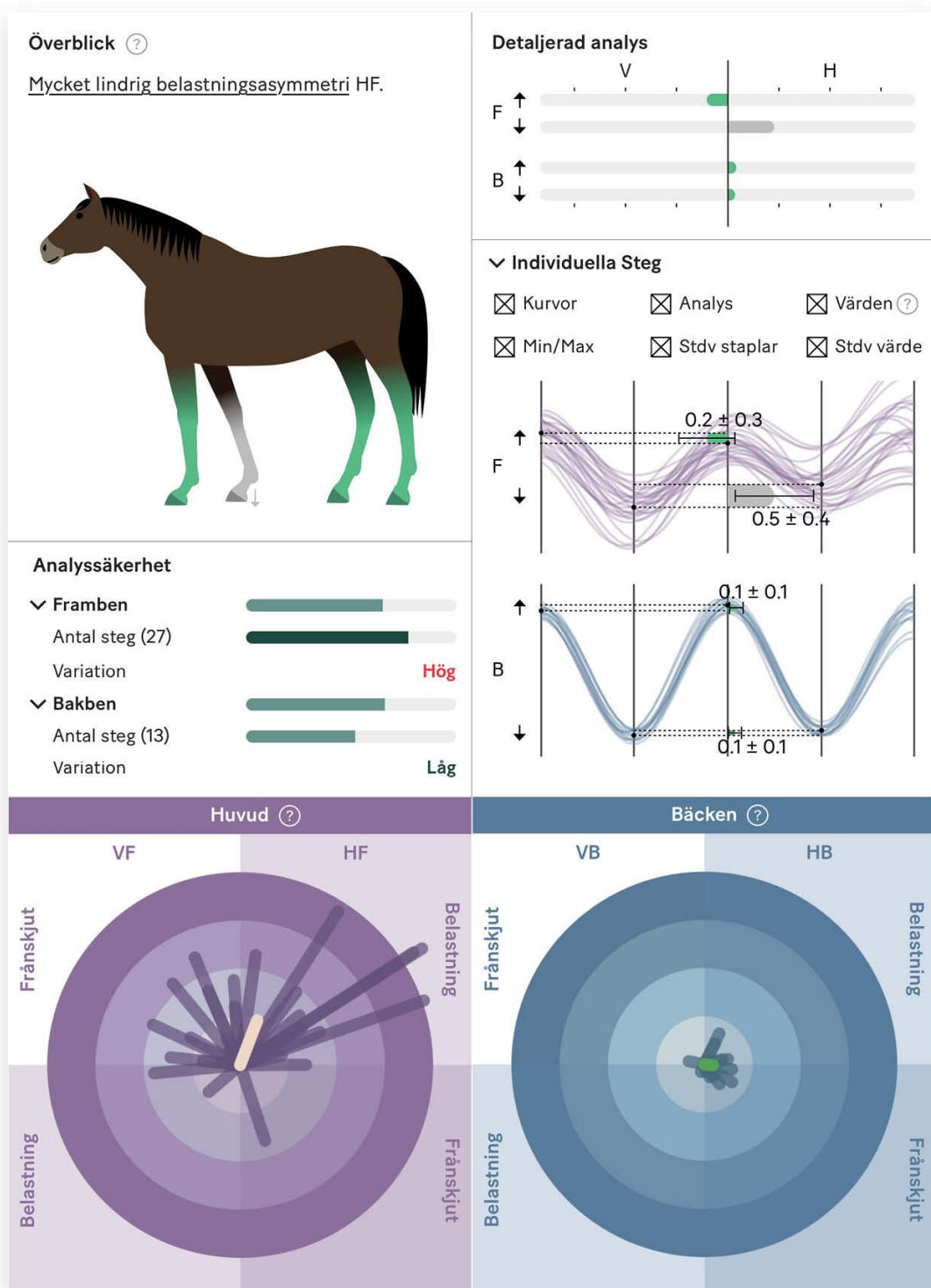
Bilaga 1

Häst 1



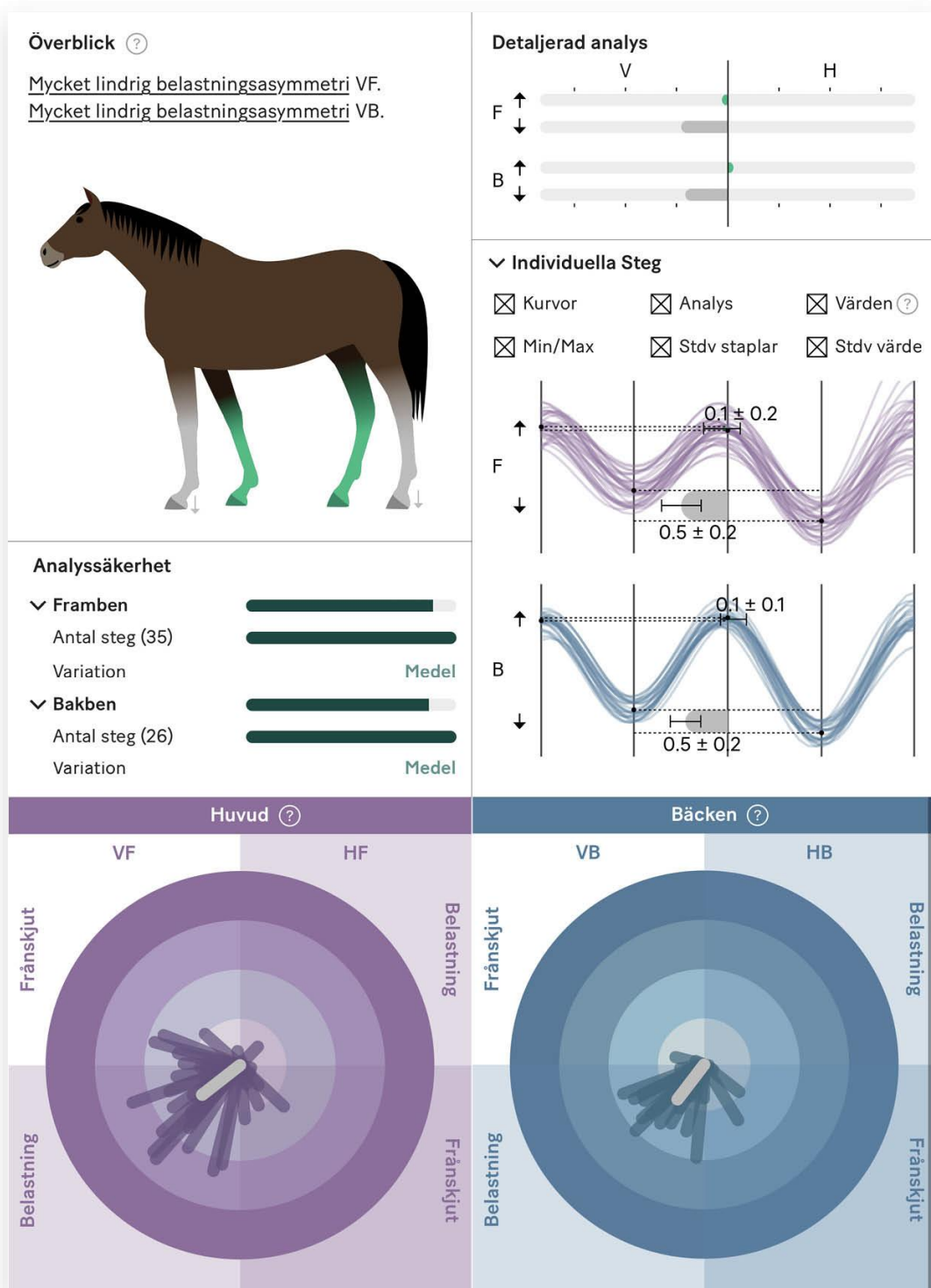
Figur 20. Analys av häst 1 från Sleip AI. V=vänster, H=höger, F=framben, B=bakben, Pil upp = frånskjutsasymmetri, Pil ned = belastningsasymmetri, Stdv=Standardavvikelse. Bildillustrationer med tillstånd från Sleip AI.

Häst 2



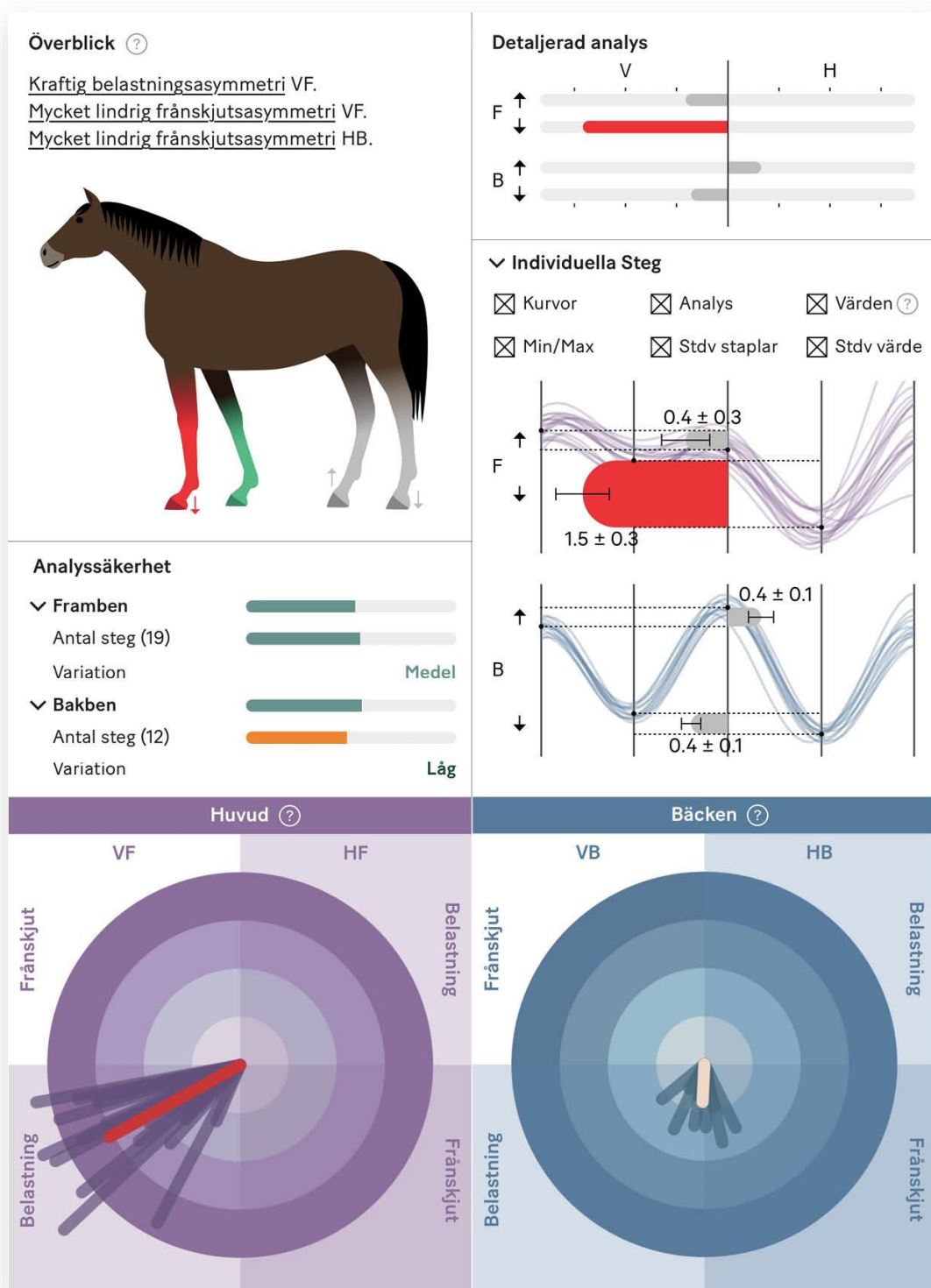
Figur 21. Analys av häst 2 från Sleip AI. V=vänster, H=höger, F=framben, B=bakben, Pil upp = frånskjutsasymmetri, Pil ned = belastningsasymmetri, Stdv=Standardavvikelse. Bildillustrationer med tillstånd från Sleip AI.

Häst 3



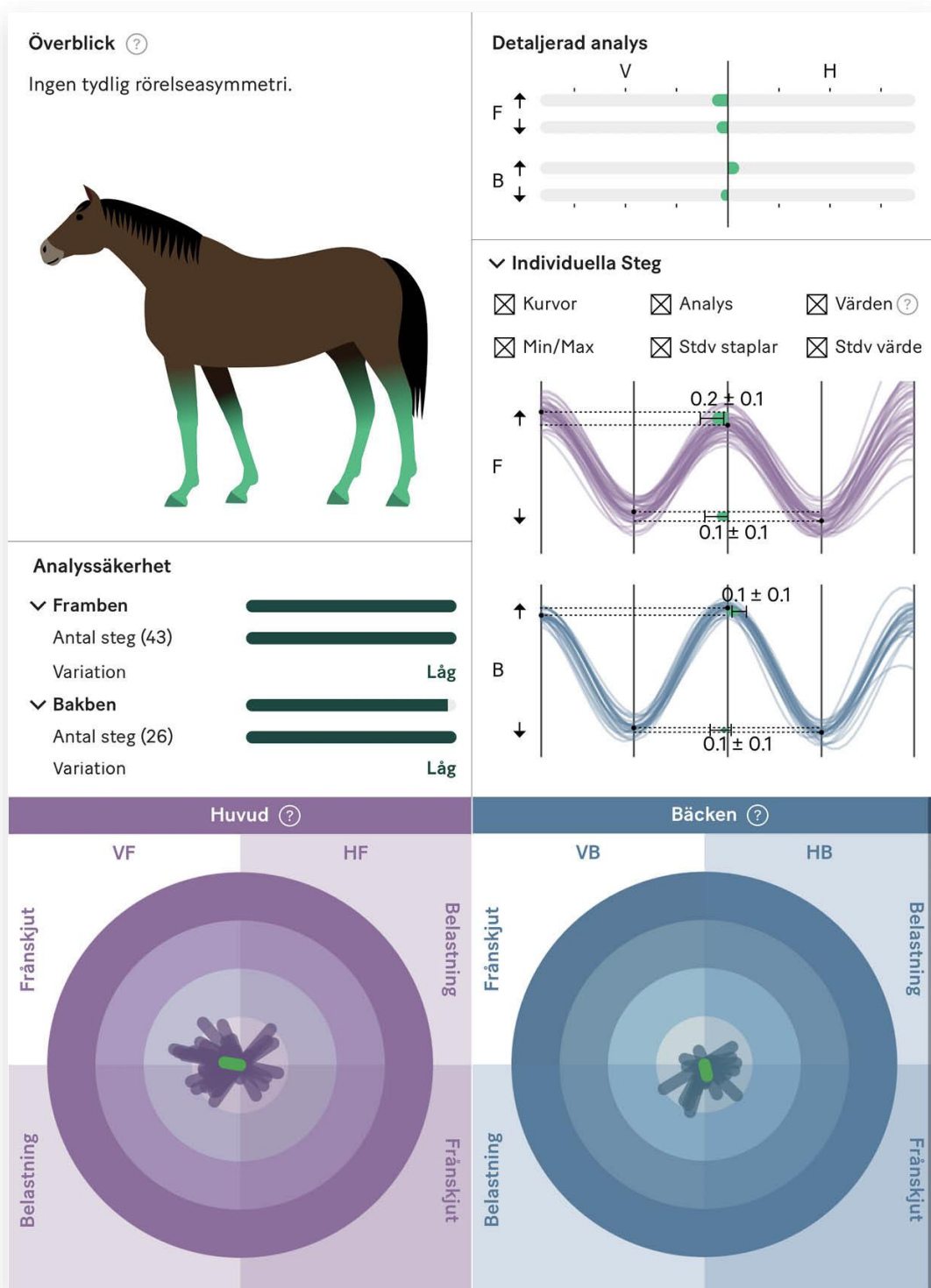
Figur 22. Analys av häst 3 från Sleip AI. V=vänster, H=höger, F=framben, B=bakben, Pil upp = frånskjutsasymmetri, Pil ned = belastningsasymmetri, Stdv=Standardavvikelse. Bildillustrationer med tillstånd från Sleip AI.

Häst 4



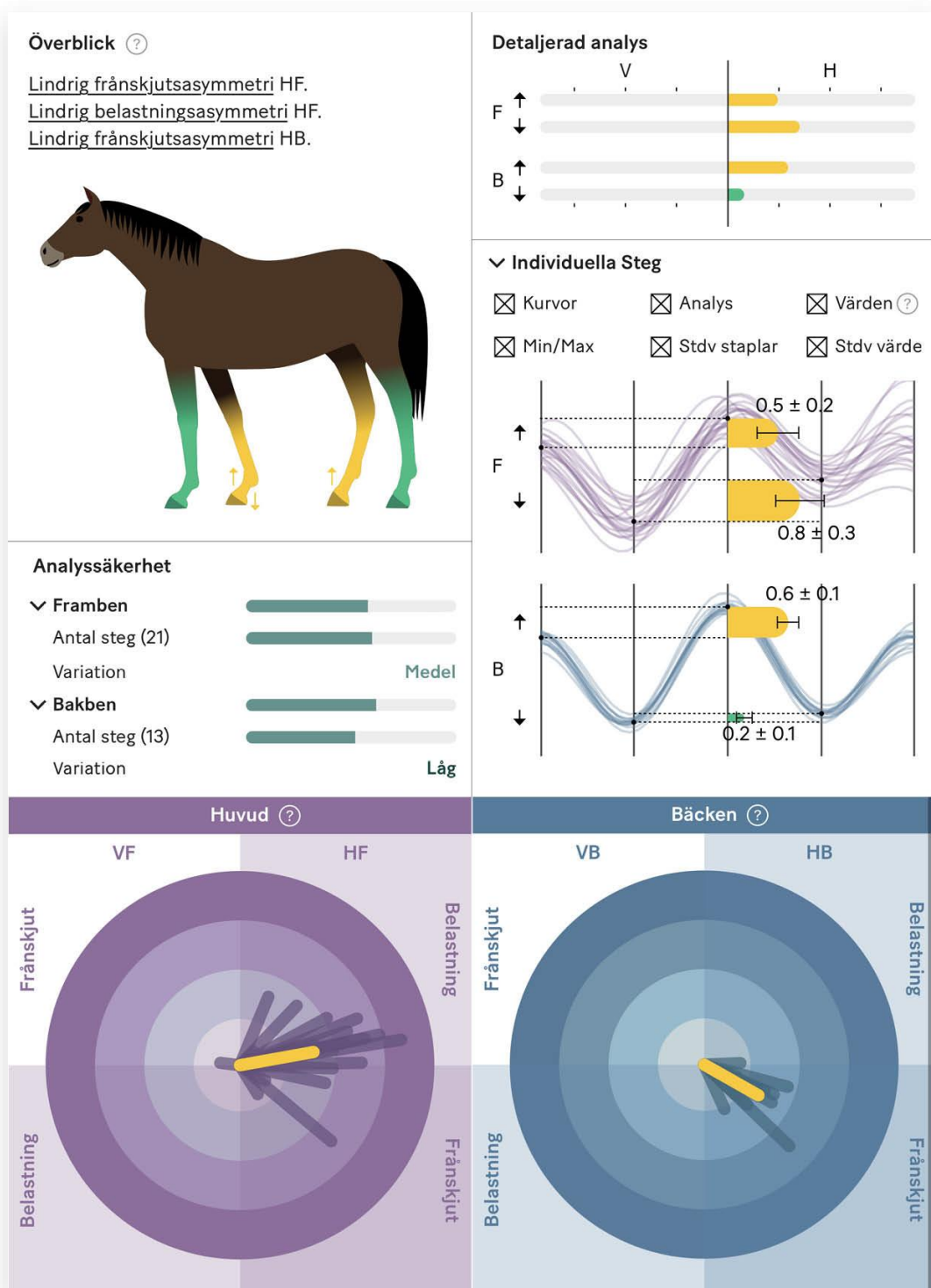
Figur 23. Analys av häst 4 från Sleip AI. V=vänster, H=höger, F=framben, B=bakben, Pil upp = frånskjutsasymmetri, Pil ned = belastningsasymmetri, Stdv=Standardavvikelse. Bildillustrationer med tillstånd från Sleip AI.

Häst 5



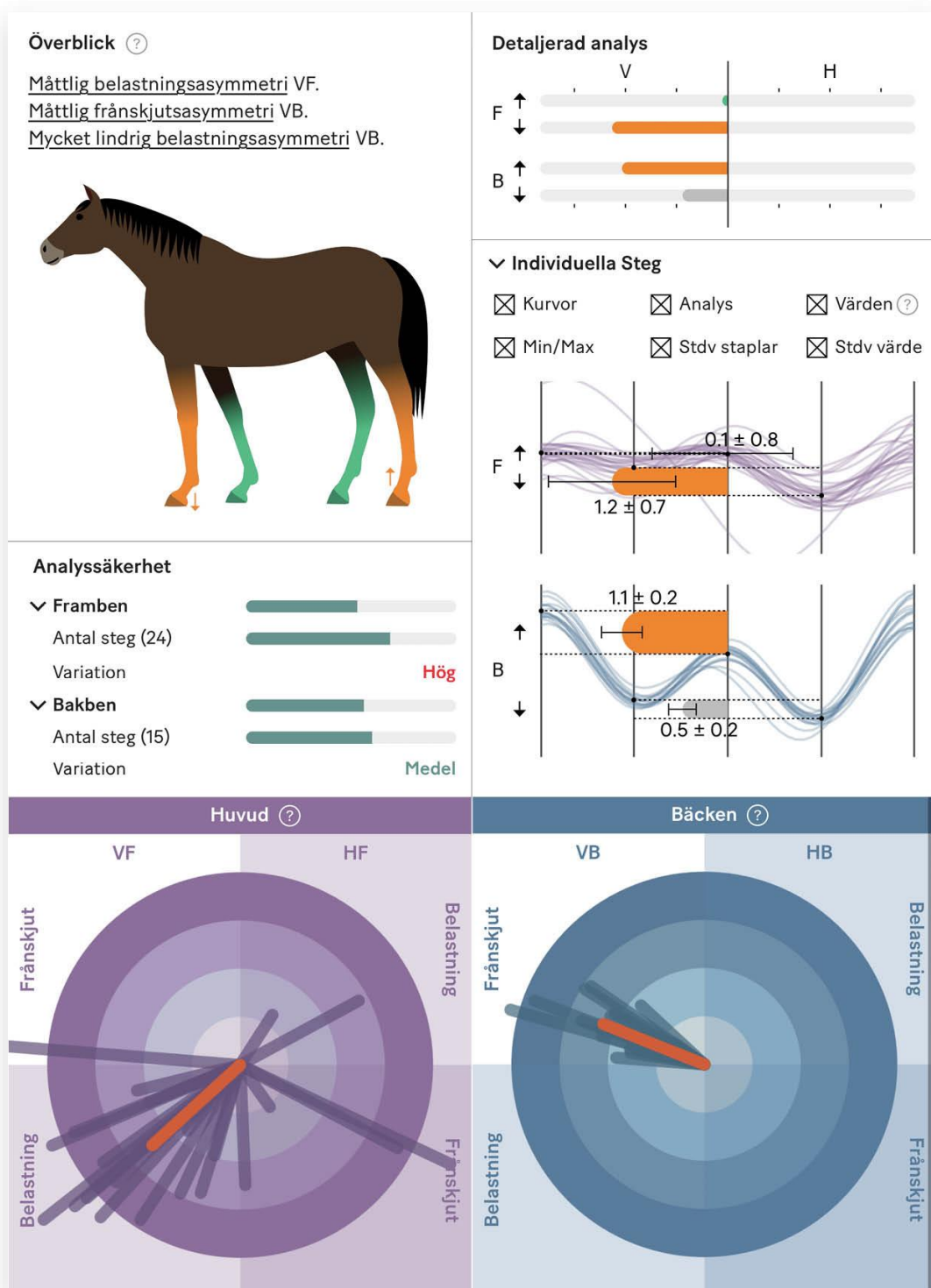
Figur 24. Analys av häst 5 från Sleip AI. V=vänster, H=höger, F=framben, B=bakben, Pil upp = frånskjutsasymmetri, Pil ned = belastningsasymmetri, Stdv=Standardavvikelse. Bildillustrationer med tillstånd från Sleip AI.

Häst 6



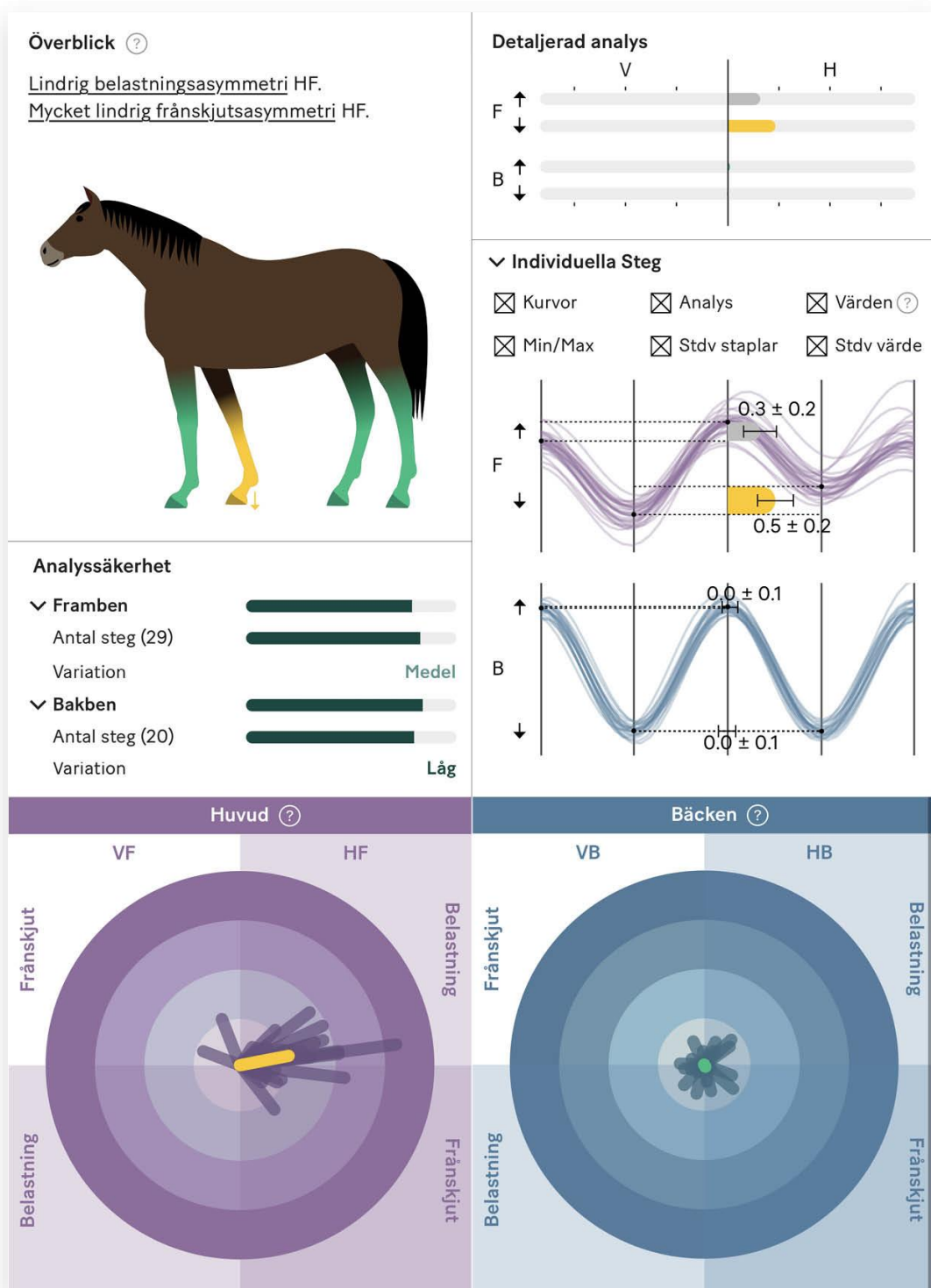
Figur 25. Analys av häst 6 från Sleip AI. V=vänster, H=höger, F=framben, B=bakben, Pil upp = frånskjutsasymmetri, Pil ned = belastningsasymmetri, Stdv=Standardavvikelse. Bildillustrationer med tillstånd från Sleip AI.

Häst 7



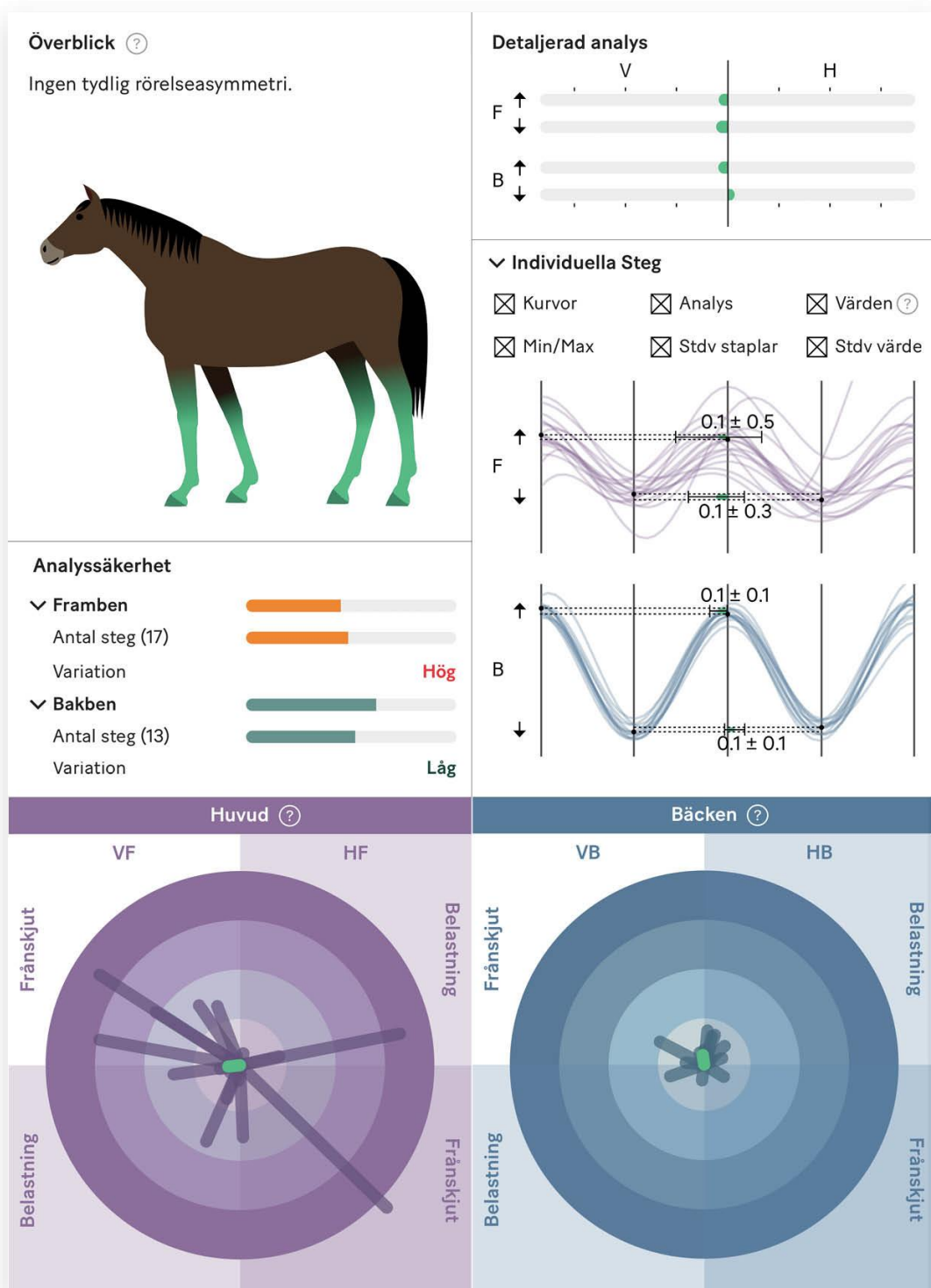
Figur 26. Analys av häst 7 från Sleip AI. V=vänster, H=höger, F=framben, B=bakben, Pil upp = frånskjutsasymmetri, Pil ned = belastningsasymmetri, Stdv=Standardavvikelse. Bildillustrationer med tillstånd från Sleip AI.

Häst 8



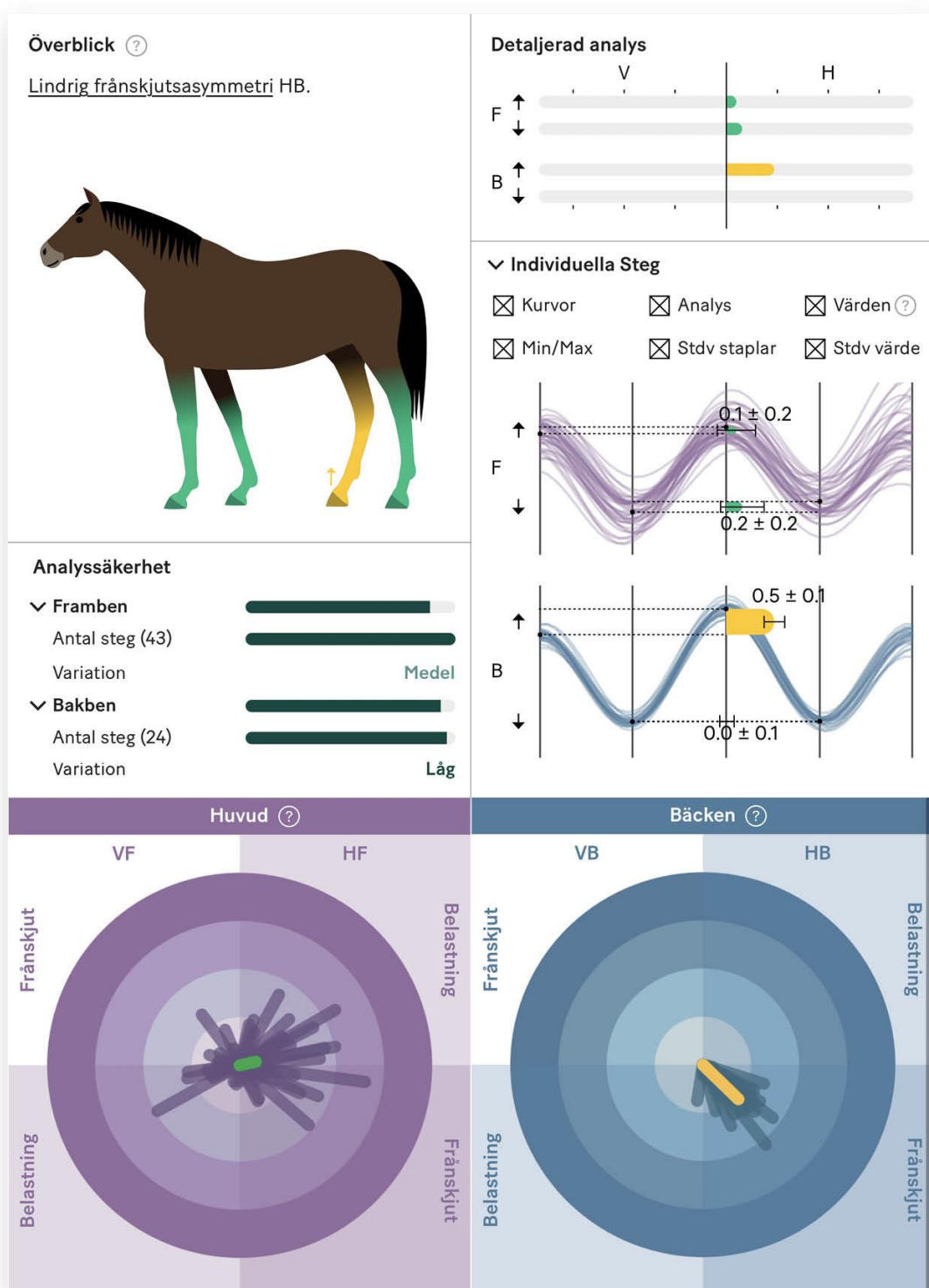
Figur 27. Analys av häst 8 från Sleip AI. V=vänster, H=höger, F=framben, B=bakben, Pil upp = frånskjutsasymmetri, Pil ned = belastningsasymmetri, Stdv=Standardavvikelse. Bildillustrationer med tillstånd från Sleip AI.

Häst 9



Figur 28. Analys av häst 9 från Sleip AI. V=vänster, H=höger, F=framben, B=bakben, Pil upp = frånskjutsasymmetri, Pil ned = belastningsasymmetri, Stdv=Standardavvikelse. Bildillustrationer med tillstånd från Sleip AI.

Häst 10



Figur 29. Analys av häst 10 från Sleip AI. V=vänster, H=höger, F=framben, B=bakben, Pil upp = frånskjutsasymmetri, Pil ned = belastningsasymmetri, Stdv=Standardavvikelse. Bildillustrationer med tillstånd från Sleip AI.