



Norges miljø- og biovitenskapelige universitet
NMBU Veterinærhøgskolen
Institutt for produksjonsdyrmedisin
Stasjonærklinisk seksjon

Fordypningsoppgave 2021/ 2022

Differensiering Produksjonsdyrmedisin og Mattrygghet

Enighet mellom observatører ved klinisk vurdering av luftvegssymptomer på kalv

Evaluation of the agreement between observers in clinical assessment of respiratory signs in calves

Cathrine Kaarfald Elvland
Ingvild Karlsen Paulsrud
Kull 2016

Veslemøy Sunniva Oma
Maria Stokstad

Innhold

SAMMENDRAG	3
DEFINISJONER OG FORKORTELSER.....	4
INNLEDNING	5
FORMÅLET MED OPPGAVA	5
OM PROSJEKTET NORWEGIAN AIRWAYS.....	7
LUFTVEGSLIDELSER HOS STORFE NORGE I DAG	8
<i>Forekomst, sjukdomstegn og risikofaktorer.....</i>	<i>8</i>
<i>Smittestoff og smittemåte</i>	<i>9</i>
VANLIGE AGENS I LUFTVEGER HOS KALV	9
<i>Bovin respiratory syncytial virus.....</i>	<i>9</i>
<i>Bovine coronavirus.....</i>	<i>9</i>
<i>Pasteurella multocida.....</i>	<i>10</i>
<i>Mannheimia hemolytica.....</i>	<i>10</i>
<i>Histophilus somni</i>	<i>10</i>
<i>Mycoplasma bovis.....</i>	<i>10</i>
TERAPIANBEFALING I DAG	11
METODE	12
OBSERVATØRENE, POPULASJON OG SCORINGSSYSTEM	12
ANALYSE	14
RESULTATER	17
DIAGRAM	17
<i>Allmenntilstand</i>	<i>17</i>
<i>Respirasjonsfrekvens</i>	<i>18</i>
<i>Ørestilling.....</i>	<i>18</i>
<i>Neseflod</i>	<i>19</i>
<i>Øyeflod.....</i>	<i>19</i>
<i>Hoste.....</i>	<i>20</i>

<i>Lungeauskultasjon</i>	21
BESKRIVELSER AV LUNGEAUSKULTASJON	23
PROSENTVIS ENIGHET (PERCENTAGE AGREEMENT, PA).....	24
COHENS KAPPA	25
FLEISS KAPPA.....	27
DISKUSJON	27
PERCENTAGE AGREEMENT	29
COHENS KAPPA	29
FLEISS KAPPA.....	29
FEILKILDER	32
<i>Undersøkelse av kalvene</i>	32
<i>Påvirkning av kalvene</i>	32
<i>Fjøset og miljøforhold</i>	32
<i>Homogen besetning</i>	32
KONKLUSJON	33
TAKK TIL BIDRAGSYTERE	34
SUMMARY	34
KILDER	35

Sammendrag

Tittel: Enighet mellom observatører ved klinisk vurdering av luftvegssymptomer hos kalv.

Forfattere: Cathrine Kaarfald Elvland og Ingvild Karlsen Paulsrud

Veileder: Veslemøy Sunniva Oma og Maria Stokstad

Luftvegslidelser er vanlig hos kalver og medfører store dyrevelferdsmessige og økonomiske konsekvenser. Sjukdomsgruppa er den som gir flest rapporterte sjukdomstilfeller hos kalv i Norge. I denne studien har vi undersøkt enigheten mellom observatører ved klinisk vurdering av luftvegssymptomer hos kalv. Dette ble gjort for å se om klinikere er enige om hvilke dyr som er sjuke og derfor får behandling. Ved hjelp av et scoringsskjema, bestemt på forhånd, vurderte seks observatører de samme 28 kalvene i én besetning.

For å undersøke enigheten mellom observatørene ble observasjonene omgjort til en totalscore. På den måten kunne enigheten utregnes ved hjelp av digitale hjelpemidler. Det ble beregnet percentage agreement (PA) og kappa- verdier (Cohens kappa & Fleiss kappa) for å få et mål på enigheten mellom observatørene.

I denne studien var det generelt lav enighet mellom observatørene. Det vil si at klinikerne scoret forskjellig på symptomene hos de ulike kalvene. Dette kan skyldes at kalvene hadde få symptomer, og besetningen blir dermed for homogen til å gi god enighet mellom observatører.

Definisjoner og forkortelser

κ	Cohens kappa/Fleiss kappa - Mål på enighet mellom ulike observatører
BI	Bias index
BRD	Bovine respiratory disease - Fellesbetegnelse på respiratorisk sykdom hos storfe
PA	Percentage agreement - Prosentvis enighet
PI	Prevalance index
WCSS	Wisconsin calf scoring system - Scoringssystem basert på luftvegssymptomer hos kalv

Innledning

Formålet med oppgava

I denne oppgava er formålet å undersøke enigheten mellom klinikere ved vurdering av luftvegssymptomer hos kalv. Luftvegssjukdom er utbredt i storfebesetninger både i Norge og internasjonalt (Klem et al., 2014), og forekomsten av luftvegsinfeksjoner hos kalv er økende i det norske storfeholdet (Falk et al., 2022). Det er et velferdsproblem hos kalvene og et stort økonomisk tap for bonden. Det kan utarte seg i form av død, dårlig tilvekst, nedsatt fruktbarhet, redusert melkeproduksjon i første laktasjon og behandlingskostnader (Blakebrough-Hall; et al., 2020; McGuirk, 2014; Stokstad et al., 2020). Det kan være vanskelig å sette diagnosen (Timsit et al., 2016), og forsinket behandling kan føre til langvarig bruk av antibiotika, tilbakefall av sjukdom og kroniske forandringer i lungene til det affiserte dyret. Å kunne sette en korrekt og pålitelig diagnose av BRD er viktig for å begrense økonomisk tap, men også redusere bruken av antibiotika i de tilfellene det ikke er nødvendig. Derfor er det ønskelig å utføre en diagnostisk testevaluering for vurdering av symptomer på luftvegssjukdom.

Ved behandling av luftvegssjukdom er det viktig å sette i gang behandlingen tidlig for å bedre prognosen for kalven. Men hvordan blir dette vurdert ute i felt? Er det forskjell på hva klinikere ute i felt definerer som sjuk kalv, og som dermed kan føre til forskjeller i hvilke dyr som får behandling? Når man skal vurdere mistanke om luftvegssjukdom på dyr generelt, bør man undersøke: allmenntilstand, temperatur, hoste, øye – og neseflod, unormal respirasjon, lungelyder ved auskultasjon og respirasjonsfrekvens (Love et al., 2014; McGuirk, 2014). Det kan tenkes at veterinærer ute i felt ikke vurderer kalvene likt, siden en klinisk undersøkelse i stor grad er subjektiv. På den måten kan noen behandle kalven på «falsk» diagnose/sjukdom, mens andre har høyere terskel for å definere kalven som sjuk før iverksettelse av behandling.

For å vurdere enigheten mellom klinikere i scoring av luftvegssymptom kan man for eksempel bruke Wisconsin calf scoring system (WCSS) (McGuirk, 2014). WCSS ble utviklet i forbindelse med høy luftvegsproblematikk hos melkekalver i USA, og hadde som mål å gjøre det mulig å oppdage sjukdom og igangsette behandling tidligere (McGuirk, 2014).

Score	Rektaltemperatur	Hostescore	Neseflodscore	Øye - eller ørescore
0	37,8 – 38,2	Ingen hoste	Normal, serøst utflod	Normal øye og øre
1	38,3 – 38,8	Indusert et host	Liten mengde unilateral, skyet utflod	Noe øyeutflod eller ørebevegelse
2	38,9 – 39,3	Indusert gjentatt hoste eller sporadisk	Bilateralt, skyet eller mukøs	Moderat bilateral øyeflod eller henger litt med det ene øret
3	≥39,4	Gjentatt sporadisk hoste	Rikelig, bilateralt mucopurulent neseflod	Kraftig øyeutflod, alvorlig skjev hodestilling, eller henger med begge ørene

Figur 1 - Wisconsin calf scoring system (McGuirk, 2014).

Ved å ta i bruk en slik tabell kan klinkerne evaluere og score for de samme symptomene på kalvene på en mer standardisert måte, og man vil få en totalscore per kalv som senere kan brukes til å beregne enigheten.

En måte å vurdere enighet (inter-rater agreement) er å beregne Cohens kapp, κ . Cohens kapp er et mål for samsvar mellom observatører, som tar hensyn til tilfeldigheter (Flight & Julious, 2015). På den måten vil Cohens kapp i teorien være et mer robust mål enn vanlig prosentregning (prosentvis enighet, PA). Cohens kapp er begrenset til å se på enigheten mellom to klinikere. For å vurdere enigheten i mellom flere enn to observatører ble Fleiss Kappa beregnet (Zapf et al., 2016).

Tidligere studier fra New Zealand (Buczinski S et al., 2016a) undersøkte enigheten mellom observatører for klinisk scoring system (WCSS), for påvisning av bovin respiratorisk sykdom (BRD). Der kalvene klassifiseres som friske eller sjuke basert på forskjellig cut- off verdier. I denne studien var det tre observatører: to veterinærstudenter på tredjeåret og en veterinærstudent på siste året. Studentene besøkte et kalveoppdrett tre ganger: dagen etter ankomst (besøk 1), 12 dager etter ankomst (besøk 2) og 27 dager etter ankomst (besøk 3). På første og andre besøket deltok kun observatør 1 og 2. På det tredje besøket undersøkte alle observatører de samme kalvene, og da ble 20 kalver undersøkt. Før første besøk fikk studentene en to- timers lang gjennomgang med en erfaren veterinær som gikk gjennom teorien og praksisen rundt bruk av WCSS. Resultatet i studien viste lav enigheten mellom observatørene. Dette var observatører med lite erfaring og det ble sett på relativt få

kalver med lite sjukdom, og en for homogen besetning ble nevnt som en forklaring på den lave enigheten.

En fersk studie fra Canada (publisert i november 2021), så også på enigheten mellom klinikere ved vurdering av symptomer på luftvegssjukdom hos kalv (Berman et al., 2021). I denne studien samlet de inn totalt 524 videoer av tilfeldige kalver fra 48 ulike produsenter. En forsker, som ikke var med i vurderingene av inter-agreement, gikk gjennom videoene og vurderte om det var kliniske tegn til luftvegssjukdom på kalvene eller ikke. De kliniske tegnene som ble tatt med i denne studien var øyeflod, neseflod, ørestilling og hodeholdning, unormal respirasjon og hoste. Det vil si totalt fem kliniske tegn. For hver av de fem kliniske tegnene plukket forskeren ut 15 videoer med tilstedeværelse av symptomet og 15 videoer uten symptomet. Totalt 30 videoer for hvert symptom ble presentert i tilfeldig rekkefølge for 18 observatører (6 veterinærer, 6 teknikere og 6 produsenter). Observatørene vurderte de kliniske tegnene ut fra et eget scoringssystem basert på Wisconsin (McGuirk, 2014) og California scoringssystem (Love et al., 2014). I denne studien fant de at det var høyere enighet ved bruk av skala med to kategorier, det vil si sjuk/frisk, sammenlignet med skalaer med tre eller fire kategorier. Symptomet med høyest inter-agreement var induisert hoste, mens unormal respirasjon var symptomet med lavest enighet. Konklusjonen på studien var at det varierte en del mellom symptomene som ble vurdert.

Også i en artikkel publisert i 2013 ble det sett på nøyaktigheten av scoring og vurdering av sjukdom hos kalver som i forsøk ble utsatt for *Mycoplasma bovis* (Amrine et al., 2013). I denne studien var det ni veterinærer, med ulik erfaring, som vurderte kliniske tegn på kalver innen 48 timer før kalvene ble avlivet og vurdert opp mot funn i lungene ved obduksjon. Også i denne studien var enigheten mellom observatørene lav, κ var mellom 0,1 og 0,21 for de individuelle vurderingene (Amrine et al., 2013).

Om prosjektet Norwegian Airways

Norwegian Airways er et prosjekt mellom Veterinærinstituttet, NMBU Veterinærhøgskolen, Animalia, Tine og Geno, Animalia, Mississippi State University, Statens Veterinärmedicinska Anstalt og Sveriges Landbruksuniversitet (Klem, 2021). Målet med prosjektet er å forbedre dyrevelferden og bærekraften i storfeholdet ved å redusere tilfeller av luftvegssjukdom og antibiotikabehandlinger. Man ønsker å studere bakterier fra øvre og nedre luftveger, fra både

friske og sjuke flokker, fra melkeproduksjon og kjøttfeproduksjon. Det er lagt spesifikt vekt på *Mycoplasma bovis*, en luftvegspatogen som finnes i naboland, men som enda ikke har blitt registrert i Norge. Prosjektet vil bidra til kunnskap om diagnostiske muligheter og om hvordan luftvegssjukdom spres i flokker.

Denne fordypningsoppgava er et sideprosjekt til Norwegian Airways. Det er en usikkerhet i prosjektet hvor samstemte veterinærene som vurderer luftvegssjukdom er, både med hensyn til erfaringen til de som vurderer og alvorligheten av sjukdom på kalvene som undersøkes. Det er derfor gunstig for prosjektet å få et tall på hvor enige eller uenige de eventuelt er.

Luftvegslidelser hos storfe Norge i dag

Forekomst, sjukdomstegn og risikofaktorer

Luftvegssjukdom hos storfe er en kjent problematikk i Norge. Luftvegslidelse er den mest diagnostiserte sjukdommen og den vanligste årsakene til død hos kalv i Norge i dag (Stokstad et al., 2020).

Luftvegslidelse er en multifaktoriell tilstand der sjukdom oppstår som et resultat av samspill mellom smittestoff, miljø og dyr (Stokstad et al., 2020). Økning i smittepress kan skje ved kontakt med andre besetninger, når besetninger slås sammen eller utvides, eller når dyr innad i besetninger blir sjuke (Tine, 2014). I store besetninger er det gjerne kalvinger spredt gjennom hele året, og på den måten blir det kontinuerlig tilgang på speddyr som potensielt kan bli sjuke og oppformere smittestoff. Sjukdomsutbrudd er vanligst i fuktige, varme fjøs med et tett dyrehold. Det kan arte seg som alt fra milde luftvegssymptomer til mer alvorlige lungebetennelser, og i verste fall føre til død. Alvorlige tilfeller kan gi nedsatt allmenntilstand, høy feber og respirasjonsvansker. Mens mindre alvorlige tilfeller kan gi mildere symptomer, som neseflod og hoste. Alvorlige symptomer er vanligst hos kalver og ungdyr. Redusert tilvekst kan oppstå som en konsekvens av et lavere fôropptak på grunn av sjukdom hos kalvene. Det er vist at det blir funnet flere dyr med lungeforandringer på slakteriene enn det som rapporteres klinisk i felt. Dette tyder på at de dyrene man klarer å plukke ut ved klinisk undersøkelse bare er en del av problematikken med BRD (Stokstad et al., 2020; Timsit et al., 2016).

Smittestoff og smitemåte

Årsaker til luftvegsinfeksjon skyldes ofte flere agens som virus og bakterier. Ofte kommer virus til som primær årsak, deretter bakteriell infeksjon sekundært når lungene er svekket av viruset. De virusene som er mest aktuelle i Norge er bovint coronavirus (BCoV), bovint respiratorisk syncytialvirus (BRSV) og bovint parainfluenza 3 (PIV3) (Stokstad et al., 2020). I andre land finnes også bovint herpesvirus 1 (BHV1) og bovint virusdiaré virus (BVDV). Norge er fri for disse virusene og de blir dermed ikke nevnt videre. Av bakterielle agens er *Pasteurella multocida*, *Mannheimia haemolytica*, *Histophilus somni*, *Mycoplasma bovis* og *Trueperella pyogenes* beskrevet å kunne gi luftvegssjukdom hos storfe (Campbell, 2016). De nevnte, uten om *Mycoplasma bovis*, finnes i normalfloraen i øvre luftveger hos kalv, men kan gi sjukdom sekundært til virusinfeksjoner.

Vanlige agens i luftveger hos kalv

Bovin respiratory syncytial virus

Bovin respiratorisk syncytial virus er RNA-virus som tilhører *Paramyxoviridae*. Det er et av de mest smittsomme luftvegsviruset hos storfe (Mattilsynet, 2014), og viruset er enzootisk i mange land verden over (Valarcher, 2007), inkludert Norge. Viruset smitter ikke til andre dyr eller til menneske, men det kan spres hurtig i en naiv flokk med storfe. BRSV replikerer i epitelceller i respirasjonstraktus og skilles ut via neseflod (Toftaker, 2019). Symptomer på infeksjon med BRSV er hoste, neseflod, feber og nedsatt matlyst (Valarcher, 2007). Symptomene kan vises alt fra 2 til 11 dager etter smittetidspunktet. Alvorligheten av sjukdommen varierer fra subklinisk til dødelig. Inflammasjonsreaksjonen kan gi nekrotiserende bronkiolitt og eventuelt interstitiell pneumoni (Toftaker, 2019) Det er også vanlig med sekundære bakterielle infeksjoner. Det er vist at storfe er seropositive for BRSV i lang tid, kalver i minst åtte måneder etter naturlig infeksjon.

Bovine coronavirus

Bovint coronavirus er enzootisk i mange land, også i Norge. Viruset gir luftvegssjukdom og kan i tillegg forårsake kalvediaré og vinterdysenteri, samt smittsom diaré på voksne. Bovint coronavirus kan replikere og skilles ut både fra epitel i respirasjonstraktus og epitel i gastrointestinaltraktus. Luftvegssjukdom forårsaket av bovint coronavirus kan oppstå hos kyr i alle aldre, avhengig av smittestatus, dyretetthet og miljø. Alvorligheten av sjukdom varierer

fra subklinisk til alvorlig, og med svært ulik klinisk opptreden (Toftaker, 2019). De kliniske symptomene er tydeligst hos unge dyr. Typiske symptomer på infeksjon er hoste og feber, noen kan også ha diaré samtidig med respirasjonssjukdom. Inkubasjonstiden ved luftvegssjukdom er fra tre til åtte dager.

Pasteurella multocida

Pasteurella multocida er en gramnegativ bakterie i familien *Pasteurellacea*. Disse bakteriene er tynne og små kokkoide staver. Bakterien er kommensal i øvre luftveger hos kalv. Utvikling og alvorlighet av sjukdommen vil variere og er antakelig avhengig av hvilke andre virus og bakterier som også er tilstede (Confer, 2009). Lesjonene vil normalt være plassert cranioventralt og ha ulike varianter av bronkopneumoni og gjerne med fibrindannelse (Campbell, 2016).

Mannheimia hemolytica

Mannheimia hemolytica er en gramnegativ stavbakterie. Bakterien tilhører også familien *Pasteurellacea*, og er kommensal i øvre luftveger hos kalv (Campbell, 2016). Gir kliniske symptomer på bakteriell pneumoni, der nedstemthet og feber kan oppstå. Det avhenger av mengde smittestoff og helsestatus på kalven hvor alvorlig sjukdommen utvikles. Kalven kan få serøs til mukopurulent neseflod med fuktig hoste. *M. Hemolytica* kan føre til en alvorlig, akutt og hemorragisk fibrinonektorisk pneumoni. Lesjonen vil normalt være plassert cranioventralt, i form av bronkopneumoni, med røde og svart/brune områder.

Histophilus somni

Histophilus somni er en gramnegativ bakterie som kan knyttes til mange ulike lesjoner og lokalisasjoner. Bakterien regnes som kommensal i øvre luftveger hos storfe (Underwood; et al., 2015). *Histophilus somni* kan invadere og skade lungenes forsvarsmekanismer, for så å utvikles til en fibrinøs bronkopneumoni. Påfølgende kan det gi septikemi, fibrinøs pleuritt, myokarditt, abort og systemisk vaskulitt og tromboemboli med meningoencephalitt (Campbell, 2016). Kliniske tegn kan komme plutselig, og kan føre til akutt død. Første kliniske tegn er gjerne nedsatt allmenntilstand, ataksi, høy feber og dyspné.

Mycoplasma bovis

Mycoplasma bovis er en bakterie som gir kroniske jur-, øre-, ledd- og luftvegsinfeksjoner hos storfe (Veterinærinstituttet, 2022). Den er opphav til flere typer sjukdom, og kan føre til

store økonomiske tap. Bakterien kan bli skilt ut fra smittede dyr på slimhinnene, og fra alle slags sekreter; inkludert melk og neseflod. Kliniske tegn å være oppmerksom på hos kalv er luftvegsinfeksjon med hoste, skjev ørestilling og skjev hodestilling som følge av infeksjon i øret, konjunktivitt og artritt (Veterinærinstituttet, 2022). *Mycoplasma bovis* har aldri blitt påvist i Norge. Likevel er den av vår interesse siden bakterien mangler cellevegg, og er dermed naturlig resistens mot betalaktamer. Med tanke på dens motstandsdyktighet mot penicillin og andre betalaktamantibiotika, er bakterien vanskelig å bekjempe ved introduksjon til en besetning. I 2020 gjennomførte Veterinærinstituttet et overvåkningsprogram i regi av Mattilsynet for å kartlegge bakterien, og på den måten opprettholde fristatusen av *Mycoplasma bovis* i Norge.

Trueperella pyogenes

Trueperella pyogenes er en grampositiv klubbeformet bakterie. Bakterien forekommer gjerne på slimhinner hos storfe, blant annet i nasopharynx (Confer, 2009). Det er ikke uvanlig at bakterien forårsaker sekundærinfeksjon ved pneumoni. Ved kroniske luftvegsinfeksjoner kan det dannes abscesser, og i disse abscessene er det vanlig å finne bakterien *T. pyogenes*. Det tar ca. 3 uker å utvikle abscessene.

Terapianbefaling i dag

Terapianbefalinger utgitt av Statens legemiddelverk for bruk av antibakterielle midler til produksjonsdyr, er fra 2012. Hos kalv med bakteriell luftvegsinfeksjon er tidlig antibakteriell behandling ofte avgjørende for prognosen (Stokstad et al., 2020). De mest aktuelle bakterielle agensene ved luftvegsinfeksjon er *Pasteurella multocida*, *Mannheimia haemolytica*, *Histophilus* og *Trueperella pyogenes*. Alle er i utgangspunktet penicilinfølsomme.

Legemiddelverket anbefaler behandling med benzylpenicilline i minst 5- 7 dager, i noen tilfeller 10 dager. For dyr med kroniske symptomer eller med gjentatte tilbakefall, så er prognosen dårlig. Det er kun i Norge og Sverige penicillin er førstevalget ved luftvegsinfeksjoner hos storfe (Klem et al., 2020).

Diagnose	Behandling
Bakteriell luftvegsinfeksjon hos kalv	Penicillin i.m 5- 7 dager. Supplere med NSAIDs, og væske hvis dehydrerte.

Tabell 1 - Terapianbefaling, Legemiddelverket 2012

Terapianbefalingene for behandling av BRD er bare delvis fulgt av praktiserende veterinær (Klem et al., 2020). Det er utbredt bruk av mer bredspektrede antimikrobielle midler. Grunnen til dette er antakelig sammensatt. Det blir satt spørsmål ved effekten av benzylpenicillin mot BRD, praktiske forhold hos bonden, gjentatte behandlinger og gjerne behandling av flere dyr samtidig. Det er heller ikke vanlig å gjøre laboratediagnostikk ved BRD, noe som hadde lagt grunnlaget for mer målrettet behandling. Ved å vurdere enigheten blant klinikere ved diagnostisering av luftvegssjukdom kan man gjøre en diagnostisk test evaluering, noe som senere kan gjøre klinikere oppmerksom på hvor nøyaktig og unøyaktig symptomene er. Målet må være å øke sannsynligheten for at det kun er de kalvene som faktisk trenger antibiotika som får det.

Målet med oppgava

I denne oppgava er målet å undersøke enigheten mellom klinikere ved vurdering av luftvegssymptomer hos kalv. Det er ønskelig å utføre en diagnostisk testevaluering for luftvegssjukdom, fordi korrekt diagnose er avgjørende for behandling og prognose hos kalven. Dette er viktig for god dyrevelferd, begrense økonomisk tap og redusere bruken av antibiotika.

Metode**Observatørene, populasjon og scoringssystem**

Det ble satt av én dag til datainnsamling. Totalt seks klinikere dro ut i en besetning med NRF melkekyr. I besetningen var det 28 kalver som klinikerne, etter tur, gjorde en klinisk vurdering av. Kalvene som ble undersøkt varierte i alder, fra noen dager gamle til fravendt melk, og det var både kvige- og oksekalver. Det var stor variasjon i oppstalling. De fleste stod i enkeltbokser, men de eldre var i fellesbinger.

Det var klinikere med ulik erfaring; én student på femte året og to studenter på siste året av veterinærstudiet, i tillegg til tre erfarne veterinærer. Med utgangspunkt i Wisconsins tabell for scoring av luftvegssymptomer (McGuirk, 2014) ble det utviklet et eget scoringssystem i forkant av besøket. Det var på forhånd bestemt hvilke symptomer som skulle vurderes; allmenntilstand, respirasjonsfrekvens, nese – og øyeflod, hoste, ørestilling og auskultasjon av

lunger. Scoring av de ulike symptomene, plassering av punkter for auskultasjon, observasjon av respirasjon osv. ble gjennomgått før forsøket startet. Klinikerne ble bedt om å notere ned så konkrete observasjoner som mulig. Det ble ikke organisert noe tidsintervall pr kalv, klinikerne brukte så lang tid som nødvendig. Klinikerne ble anonymisert så vurderingene ikke kunne kobles opp mot erfaring før etter at dataene ble analysert.

Score	Allmenntilstand	Respirasjonsfrekvens (resp/min)	Neseflod	Øyeflod	Hoste	Ørestilling
0	Kvikk og alert	≤ 49	Normal	Normal	Ingen hoste	Normal
1	Mildt nedstemt,	50-54	Serøs, mukøs	Serøs, mukøs	Provosert hoste	Unormal
2	Nedstemt, nedsatt matlyst	55-64	Mucopurulent eller purulent	Mucopurulent eller purulent	Sporadisk hoste	-
3	Allment påvirket	65-74	-	-	Mer enn en sporadisk hoste i løpet av undersøkelsen	-
4	Liggende	75-85	-	-	-	-

Tabell 2- Scoringssystem luftvegssymptomer

Auskultasjon: Klinikerne noterte avvik for de enkelte punktene.

H1	Høyre dorsalt
H2	Høyre høyde med skulderledd
H3	Høyre ved albuen
V1	Venstre side dorsalt
V2	Venstre høyde med skulderledd
V3	Venstre ved albuen

Tabell 3 - Punkter for auskultasjon

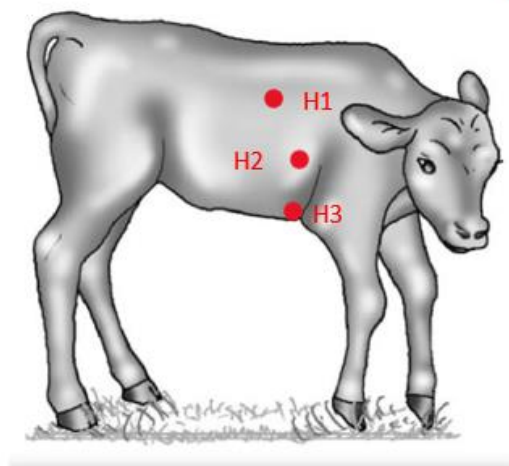


Figure 13.19 Contracted flexor tendons of the thoracic limbs of a

Figur 2 - punktene som ble auskultert. Illustrasjon hentet fra Clinical examination of farm animals.

Analyse

Dataene ble samlet inn og hvert av symptomene ble gjort om til en score i følge av vårt scoringssystem, se tabell 2. Scoringen per symptom varierer fra 0 – 4, der for eksempel allmenntilstand har scoring opp til 4, mens ørestilling har opp til 1. Det ble så regnet ut totalscore for sjukdom per kalv for hver enkelt kliniker, der scoringen fra symptomene blir summert. For eksempel vil en kalv med serøs neseflod (score 1) med respirasjonsfrekvens på 52/min (score 1), og ellers ingen avvik, får en totalscore på 2. Maks totalscore som en kalv kan få er 17.

For å vurdere enigheten mellom klinikerne ble prosent enighet (PA) først beregnet. Den er definert som antall observasjoner med lik scoring per symptom delt på totalt antall observasjoner. PA anses å være god om den er $\geq 0,75$ (Buczinski S et al., 2016b).

For å vurdere enighet mellom klinikere ble Cohens kappa, κ , beregnet for totalscore av sjukdom mellom hver kliniker. Cohens kappa tar høyde for tilfeldighet og vekter uenigheter i beregningen (Buczinski S et al., 2016a). κ ble vurdert med ulike cut-off, verdier for sjukdom, for å se hvor høy score som skal til før man kan anse klinikerne som enige i at kalven er sjuk. Her ble det satt grense på totalscore >2 og >3 . Cohens kappa ble vurdert mellom hver av

klinikerne parvis. Vurderingen av Cohens kappa for hva som er god og dårlig enighet, er vist i figur 3.

TABLE 3. Interpretation of Cohen's kappa.

Value of Kappa	Level of Agreement	% of Data that are Reliable
0-.20	None	0-4%
.21-.39	Minimal	4-15%
.40-.59	Weak	15-35%
.60-.79	Moderate	35-63%
.80-.90	Strong	64-81%
Above .90	Almost Perfect	82-100%

Figur 3 - Tabell hentet fra Biochemia media(McHugh, 2012)

For å kunne vurdere enigheten mellom grupper (studenter vs veterinærer), ble Fleiss kappa beregnet. Fleiss kappa kan beregnes mellom flere enn to observatører. I tillegg kan man på den måten også regne κ mellom alle klinikerne på utvalgte symptomer. Vurderingen av Fleiss kappa har tilsvarende inndeling som Cohens kappa i tabellen over (figur 3).

Prevalens indeks (PI) og bias indeks (BI) ble også estimert. PI beregnes som den positive forskjellen mellom symptomer observatørene er enige om, delt på totalt mulige enigheter. En høy PI indikerer en ubalansert matrise, for eksempel i en homogen besetning med mange friske dyr. Dette vil igjen føre til en lav κ - verdi. BI beregnes som differansen i uenigheten mellom observatørene, delt på totalt mulige enigheter. På den måten er BI en indikator på uenighet mellom differansen på uenighetene. En høy BI indikerer at en observatør scorer vesentlig høyere eller lavere enn den andre observatøren (Buczinski S et al., 2016a).

For å beregne PI og BI kan man sette opp 2x 2 matrise med de to ulike observatørene, der 1 (*frisk*) og 2 (*syk*) symbolisere de ulike scorene. A og d er observatørene enige om scoringen på kalvene, mens b og c er uenige. N er totalt observasjoner ($N = 28$ kalver)

	1	2	Total
1	a	b	g_1
2	c	d	g_2
Total	f_1	f_2	N

Figur 4- data fra to observatører som vurderer to kategorier(Flight & Julious, 2015)

Utregningene på PI og Bi blir dermed:

$$PI = \frac{(a - d)}{N}$$

Formel 2 - Formel for beregning av Prevalence Index(Flight & Julious, 2015)

$$BI = \frac{(b - c)}{N}$$

Formel 1 - Formel for beregning av Bias Index(Flight & Julious, 2015)

De statistiske analysene ble gjennomført ved bruk av Microsoft Office Excel (Cohens kappa) og Stata versjon 17 (Fleiss Kappa). Som tidligere nevnt, ble Cohens kappa beregnet med cut-off på >2 og >3 av totalscoren. På den måten kan vi definere hvilke kalver klinkerne var enige om var sjuke. Formelen på Cohens kappa:

$$\kappa = \frac{p_o - p_e}{1 - p_e}$$

Formel 33 - Cohen's Kappa(Flight & Julious, 2015)

Cohens kappa, κ : der p_o er andelen kalver klinikerne er enige om, delt på antall mulige enigheter (antall kalver). Mens p_e tar hensyn for tilfeldigheter bak klinikerens enighet.

Fleiss kappa- verdiene ble utregnet i statistikkprogrammet Stata. Det ble beregnet enighet mellom alle klinkerne på totalscoren og ved score >4. I tillegg ble enigheten mellom veterinærene og studentene, beregnet hver for seg, og på de ulike symptomene.

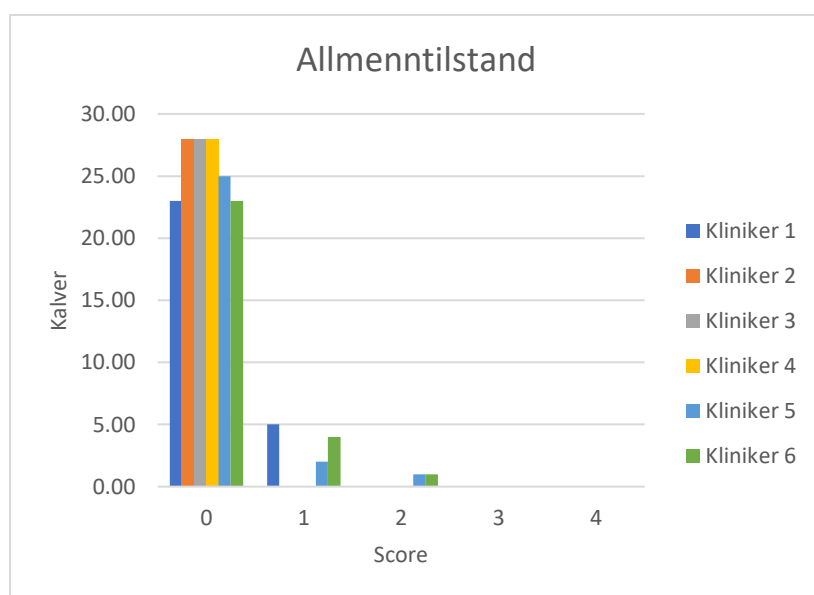
Ved vurderingen av auskultasjon ble seks punkter auskultert og eventuelt avvik per punkt notert. Senere ble hvert avvik gjort om til ett poeng, totalscore per kalv kunne dermed gå fra 0 til 6. Det var stor forskjell på oppførselen til kalvene i løpet av dagen, og små variasjoner i vurderingene ville gitt stort utslag på beregningen av kappa. For å kunne ta med auskultasjon i beregningen justerte vi scoren til 0 – 1, der avvik på auskultasjonen gir en score på 1. Auskultasjonen ble tatt med i beregningen av Fleiss kappa, men ikke av Cohens kappa.

Resultater

Diagram

Hvert symptom vi undersøkte har blitt satt inn i søylediagram for å vise scoringen hos klinikerne.

Allmenntilstand



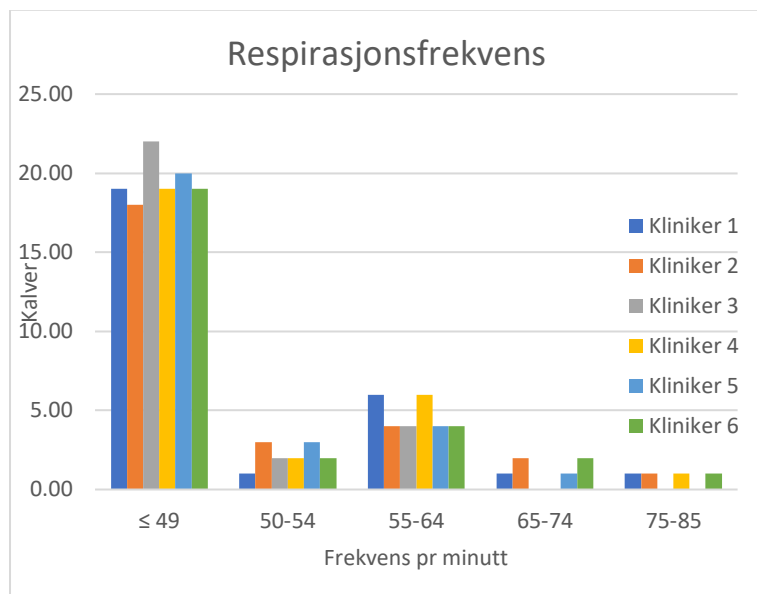
Allmenntilstand ble vurdert ut i fra, scoring:

0. *Kvikk og alert*
1. *Mildt nedstemt*
2. *Nedstemt, nedsatt matlyst*
3. *Allment påkjent*
4. *Liggende*

Figur 5 – Diagram almenntilstand

Søylediagrammet viser til at de fleste kalvene ble tolket som friske kalver med lite avvik. Klinikerne 2, 3 og 4, er helt samkjørte og har scoret alle kalvene med score 0. Mens klinikerne 1, 5 og 6 har scoret noen kalver høyere. Det vil si at de har tolket noen av kalvene som mildt nedstemte.

Respirasjonsfrekvens



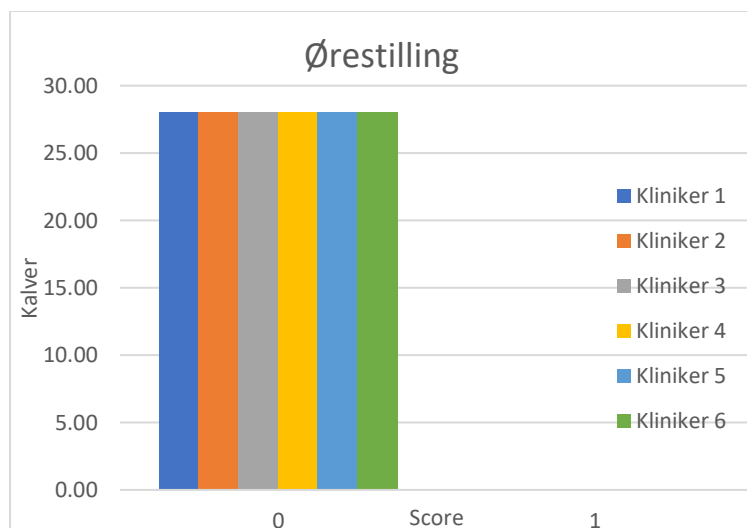
Respirasjonsfrekvensen ble scoret ut i fra antall respirasjon per minutt, scoring:

0. Under 49 /min
1. 50- 54
2. 55- 64
3. 65- 74
4. 75- 85

Figur 6 – Diagram respirasjonsfrekvens

Her er det noe mer variert scoring, fra scoring på 0 – 4. De fleste klinkerne har på dette punktet scoret i alle forskjellige skalaer. Den største andelen ble scoret på 0, med respirasjon under 49/min.

Ørestilling



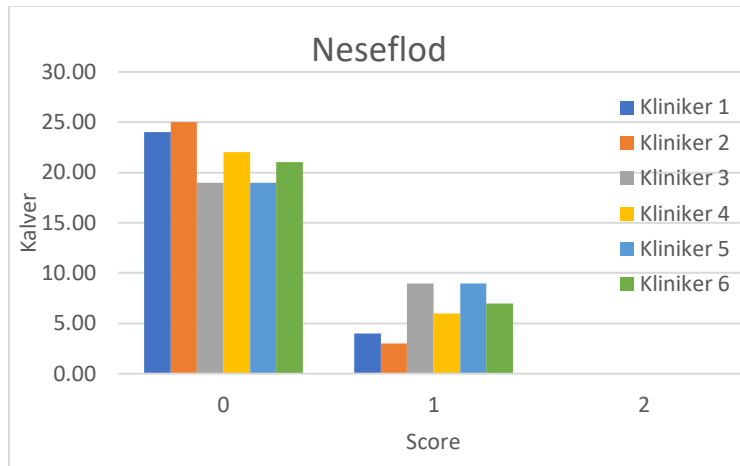
Ørestilling ble vurdert ut i fra, scoring:

0. Normal
1. Avvik

Figur 7 – Diagram ørestilling

Ørestillingen ble tatt med i scoringen siden otitis media kan oppstå ved *Mycoplasma bovis* infeksjon hos kalver. Da vil kalvene få en skjev ørestilling, eventuelt skakke litt på hodet. Det var 100% enighet mellom klinkerne. Dette punktet blir ikke diskutert noe videre i oppgaven.

Neseflod



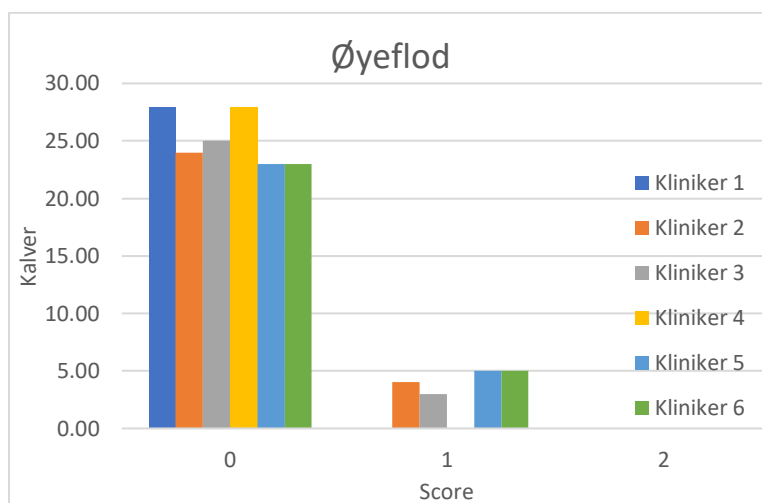
Neseflod ble vurdert ut i fra, scoring:

0. Normal
1. Serøs, mukøs
2. Mukopurulent eller purulent

Figur 8 – Diagram neseflod

Klinkerne har fordelt sin scoring i to grupper, på scoring 0 og 1. Ingen registrert score 2. Det er litt forskjell mellom klinkerne på antall kalver de observerte med serøs neseflod. Blant annet registrerte klinikerne 1 og 2 færre kalver med serøs neseflod sammenlignet med kliniker 3 og 5.

Øyeflod

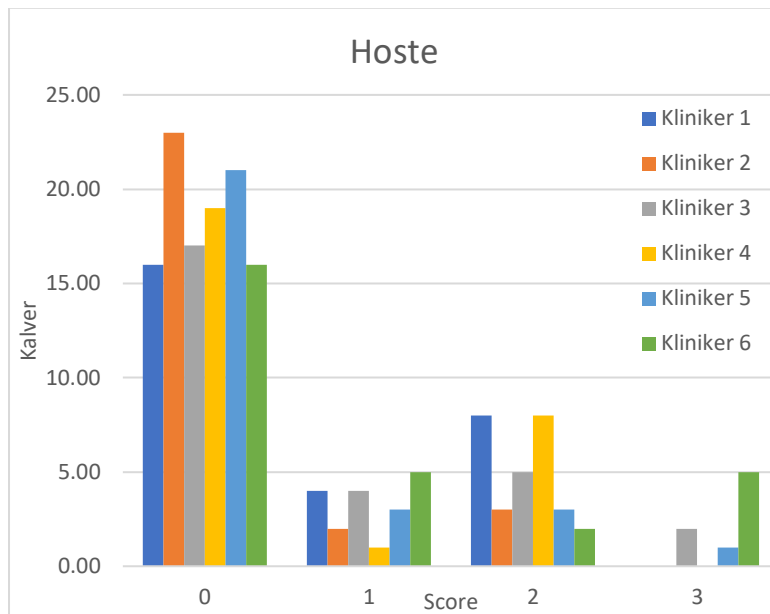


Øyeflod ble vurdert ut fra scoring:

0. Normal
1. Serøs, mukøs
2. Mucopurulent, purulent

Figur 9 – Diagram øyeflod

Klinkerne 1 og 4 observerte ingen kalver med øyeflod, mens de andre klinkerne observerte ≤ 5 kalver med serøs/mukøs øyeflod.

Hoste

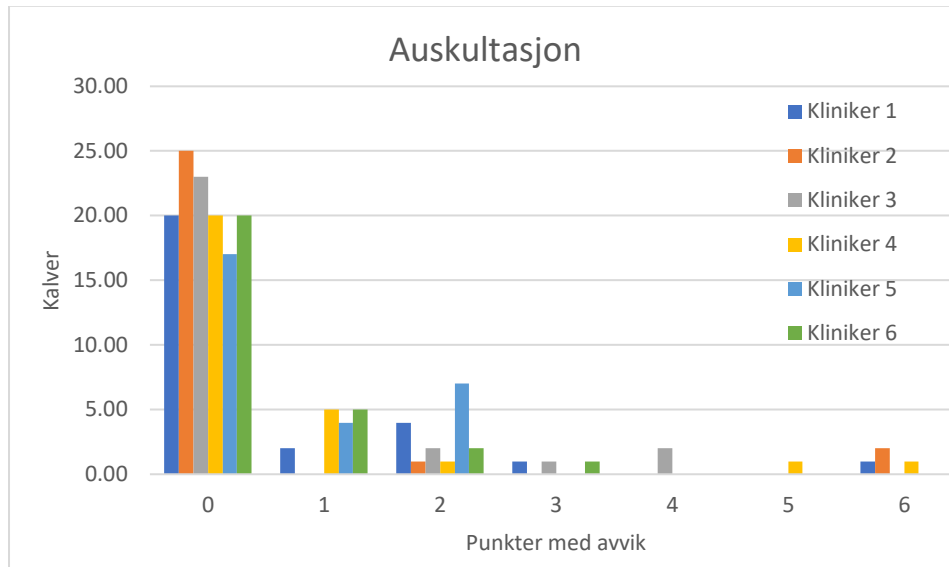
Hoste ble vurdert ut fra, scoring:

0. *Ingen hoste*
1. *Provosert hoste*
2. *Sporadisk hoste*
3. *Mer enn en sporadisk hoste i løpet av undersøkelsen*

Figur 10 – Diagram hoste

Søylediagrammet viser til at det ble observert en del hoste i besetningen, men av ulik grad.

Lungeauskultasjon



Figur 11 – Diagram lungeauskultasjon

Auskultasjon ble vurdert ut fra antall punkter med avvik

1. *Ett punkt med avvik*
2. *To punkter med avvik*
3. *Tre punkter med avvik*
4. *Fire punkter med avvik*
5. *Fem punkter med avvik*
6. *Seks punkter med avvik*

Lungeauskultasjon var den faktoren som var mest varierende. Her var det lite samsvar mellom klinikerne. Ut i fra søylediagrammet ser man vesentlig forskjell på antall kalver som ble registret med avvik per kliniker. For eksempel kliniker 2 undersøkte 25 kalver uten avvik, mens kliniker 5 registrerte 17 kalver uten avvik. Det søylediagrammet ikke viser er hvilke kalver som klinikerne er enige om. Dette vises i tabellen under (tabell 4).

Lungeauskultasjon, scoring for hver enkelt kalv

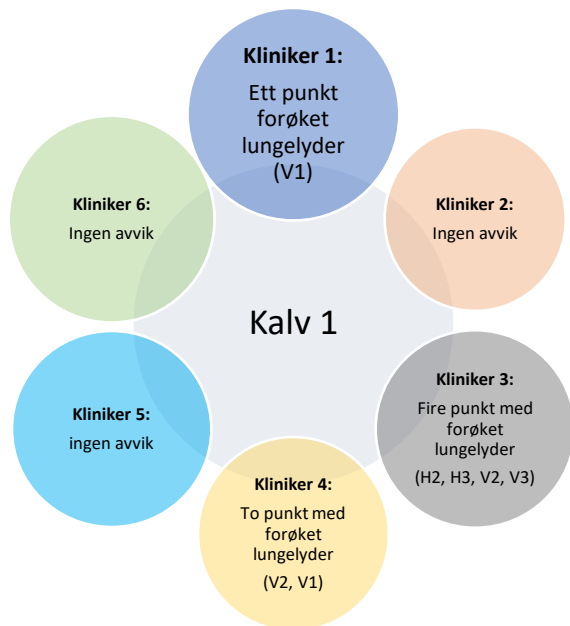
<i>Kalv nr</i>	Kliniker 1	Kliniker 2	Kliniker 3	Kliniker 4	Kliniker 5	Kliniker 6
1	1	0	4	2	0	0
2	0	0	2	1	0	0
3	0	0	0	0	0	0
4	2	0	0	1	2	2
5	0	0	0	1	1	0
6	0	6	0	0	2	0
7	0	0	0	0	0	1
8	1	0	0	0	0	0
9	2	0	0	0	1	0
10	2	0	0	0	0	1
11	3	0	0	0	0	0
12	0	0	0	0	0	0
13	6	0	0	5	0	0
14	0	0	0	0	0	0
15	2	0	2	0	0	0
16	0	0	0	0	1	0
17	0	0	0	0	2	0
18	0	0	0	0	0	1
19	0	0	0	1	0	1
20	0	0	0	1	2	0
21	0	0	0	0	0	1
22	0	6	4	0	2	3
23	0	2	3	6	0	0
24	0	0	0	0	1	0
25	0	0	0	0	2	0
26	0	0	0	0	0	0
27	0	0	0	0	2	0
28	0	0	0	0	0	2

Tabell 4 - Scoring av lungeauskultasjon, tallet indikerer antall punkter med registrerte avvik

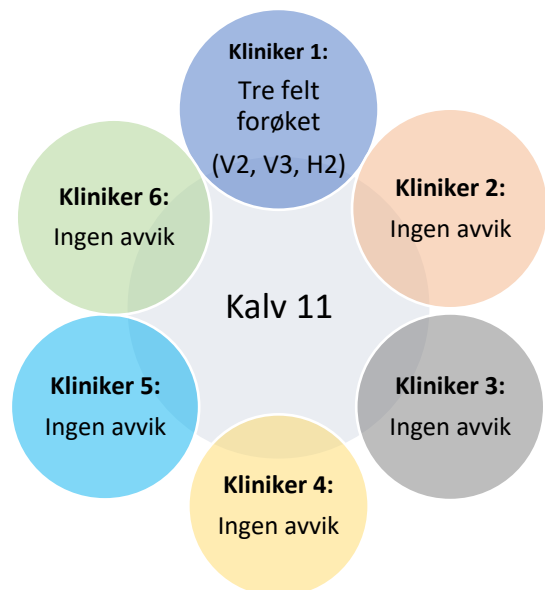
Tabellen viser til hvilke kalver som har fått ulik score hos de forskjellige klinikerne. Her er det store forskjeller.

Beskrivelser av lungeauskultasjon

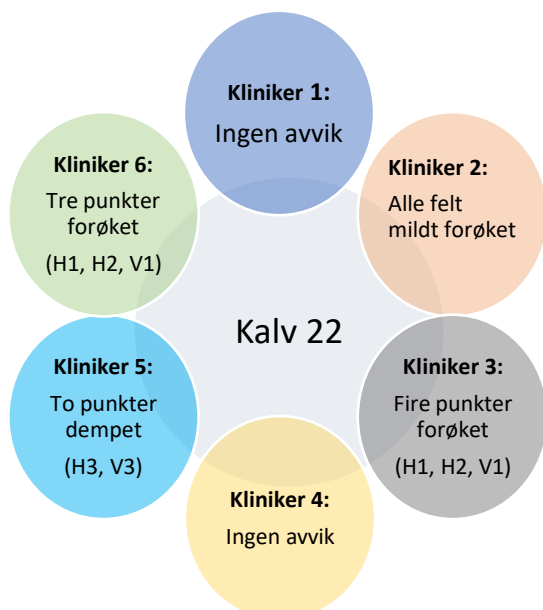
Figurene under er lagt ved for å illustrere hvor store forskjeller det var av funn på lungeauskultasjonen på enkelte kalver. Her viser vi eksempler på tre kalver med varierende enighet på avvikene. Hver boble er funn på auskultasjonen per kliniker, der kalvens nummer står sentralt i figuren. V2 betyr avvik på venstre side på punkt nr. 2, mens H står for funn på høyre side.



Figur 1212 - Kalv 1, funn på lungeauskultasjon



Figur 1313 - Kalv 11, funn på lungeauskultasjon

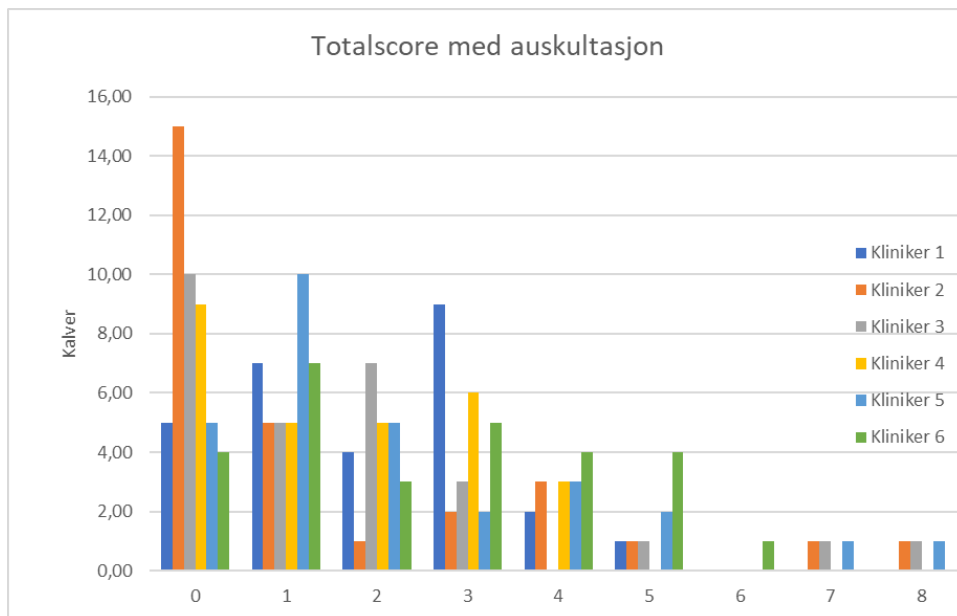


Figur 1414 - Kalv 22, funn på lungeauskultasjon

Totalscore

Diagram som viser fordeling av totalscore for hver kliniker, basert på alle symptomene.

Høyeste score som ble gitt var 8, av en maks sum på 17.



Prosentvis enighet (percentage agreement, PA)

Tabellen viser til prosentvis enighet mellom alle klinkerne for hver enkelt kategori.

	Percentage agreement, PA
Allmentilstand	86 %
Respirasjonsfrekvens normal	69 %
Ørestilling	100 %
Neseflod	68 %
Øyeflod	83 %
Hoste	56 %
Auskultasjon	56 %
Totalscore, mellom alle klinkerne	26 %
Totalscore, mellom alle veterinærer	30 %
Totalscore, mellom alle studenter	20 %

Tabell 5 - Percentage agreement**Cohens kappa**

Cohens kappa, κ , er et mål for samsvar mellom observatører, som tar hensyn til tilfeldigheter. Tolkning av kappa kan gjøres ut ifra tabellen under. Verdien på kappa går fra - 1– + 1, der 1 betyr total enighet.

TABLE 3. Interpretation of Cohen's kappa.

Value of Kappa	Level of Agreement	% of Data that are Reliable
0-.20	None	0-4%
.21-.39	Minimal	4-15%
.40-.59	Weak	15-35%
.60-.79	Moderate	35-63%
.80-.90	Strong	64-81%
Above .90	Almost Perfect	82-100%

Figur 15 - Tabell hentet fra Biochemia media(McHugh, 2012).**Bestemte grense for sykdom >2 i totalscore.**

Kalver score >2	PA	κ	Bias index	Prevalence index
Kliniker 1 og 2	7 %	0,50	-0,04	0,54
Kliniker 1 og 3	5 %	0,44	0,04	0,61
Kliniker 1 og 4	5 %	0,36	0	0,57
Kliniker 1 og 5	9 %	0,62	-0,07	0,50
Kliniker 1 og 6	9 %	0,43	-0,18	0,39
Kliniker 2 og 3	7 %	0,58	0,071	0,57
2 og 4	4 %	0,10	0,036	0,54
2 og 5	9 %	0,55	-0,036	0,46
2 og 6	5 %	0,04	-0,143	0,36
Kliniker 3 og 4	4 %	0,21	-0,04	0,61
3 og 5	4 %	0,11	-0,11	0,56
3 og 6	5 %	0,17	-0,21	0,43

Kliniker 4 og 5	4 %	0,05	-0,07	0,50
4 og 6	7 %	0,27	-0,18	0,40
Kliniker 5 og 6	9 %	0,29	-0,11	0,32

Tabell 6- Cohens kappa, sjuke kalver >2

Bestemte grense for sjukdom >3 i totalscore

Kalver score >3	PA	κ	Bias index	Prevalence index
Kliniker 1 og 2	2 %	0,13	-0,07	0,71
Kliniker 1 og 3	4 %	0,63	0	0,79
Kliniker 1 og 4	2 %	0,34	0,03	0,82
Kliniker 1 og 5	5 %	0,71	-0,07	0,71
Kliniker 1 og 6	0 %	-0,18	-0,14	0,64
Kliniker 2 og 3	2 %	0,13	0,07	0,71
2 og 4	0 %	-0,11	0,11	0,75
2 og 5	5 %	0,51	0	0,64
2 og 6	4 %	0,16	-0,07	0,57
Kliniker 3 og 4	0 %	-0,09	0,036	0,82
3 og 5	4 %	0,42	-0,07	0,71
3 og 6	2 %	0,059	-0,14	0,64
Kliniker 4 og 5	1 %	0,20	-0,11	0,75
4 og 6	0 %	-0,13	-0,18	0,67
Kliniker 5 og 6	2 %	-0,05	-0,07	0,57

Tabell 7 - Cohens Kappa, sjuke kalver >3

Anonymiteten avslørt etter bearbeidet data:

- Veterinærer: Kliniker 2, 3 og 4
- Studenter: Kliniker 1, 5 og 6

Studenter vs veterinærer	κ	Bias index	Prevalence index
Score >2	0,57	0,21	0,43
Score >3	0,20	0,11	0,75

Tabell 8 - Cohens kappa, veterinærer vs studenter

Fleiss kappa

Resultater med flere observatører

	κ
Totalscore, mellom alle klinkerne	0,08
Totalscore, mellom alle veterinærer	0,07
Totalscore, mellom alle studenter	0,02
Totalscore (cut- off >3)	0,18
Totalscore (cut – off >4)	0,06
Totalscore (cut- off >5)	0,05
Veterinærer (cut – off >4)	0,15
Studenter (cut- off>4)	-0,12

Tabell 9 – Fleiss kappa

Det ble også beregnet Fleiss kappa for hver enkelt kategori mellom alle klinkerne. Generelt var det lav kappa på alle kategoriene/symptomene, med enighet på 0- .20, som tilsvarer ingen enighet (figur 15). Unntaket er respirasjon, som hadde en kappaverdi på 0,35. Verdien tilsvarer en minimal enighet mellom klinkerne.

Diskusjon

Allmenntilstand

Som vist i figur 5 var det noe ulikt hvordan allmenntilstand ble vurdert. Her kan blant annet kalvens aktivitetsnivå påvirke klinikernes vurdering. Noen kalver ble raskt slitne etter håndtering, og kunne dermed fremstå nedstemte. Kalvene fikk heller ikke lang pause før neste kliniker. Selve undersøkelsen av kalvene tok i gjennomsnitt fem minutter, på den måten fikk vi ikke undersøkt kalvens matlyst nærmere.

Respirasjonsfrekvens

Kalvens aktivitetsnivå varierte i løpet av dagen. Noen kalver var oppjaget fra forrige klinikers undersøkelse, mens andre kalver var slitne og sov før neste undersøkelse. Dermed blir det vanskelig å kunne dra en endelig konklusjon på enigheten ut fra scoringen. Mest sannsynlig vil det ikke være her de største forskjellene vil oppstå mellom klinikerne siden respirasjonsfrekvens er et objektivt mål. På noen kalver var det vanskelig å telle respirasjonen på avstand, og dermed telte man ved hjelp av auskultasjon. Om dette har innvirkning på scoringen, er lite trolig.

Neseflod

Om dyrene varierte mengde neseflod gjennom dagen er uvisst, men det kan tenkes hvis kalven renses nesen hyppig. På den måten kan kalven ha renses nesen før neste kliniker skal undersøke den.

Øyeflod

Mengden øyeflod vil også kunne variere gjennom dagen. I deler av fjøset var det ganske mørkt, slik at det var litt vanskeligere for klinikerne å vurdere neseflod og øyeflod.

Hoste

Metoden for å framprovosere hoste hos kalven ble gjennomgått i plenum, men det er vanskelig å påse at alle gjør dette helt identisk med likt trykk. Grad av sporadisk hosting hos kalvene varierer. Derfor kan det være tilfeldig hvor mye hosting klinikerne observerte under undersøkelsen, men sjansen er større for å registrere hosting jo lengre tid man bruker.

Lungeauskultasjon

Det er stor variasjon i hvordan klinikerne beskriver funn ved auskultasjon av lungene. Selv om man i dette forsøket hadde definerte områder som skulle lyttes på, var det stor forskjell på hva som ble beskrevet. Som vist i figur 12, 13 og 14, er opplevelsen av auskultasjon veldig subjektiv. For eksempel ser man på kalv 1, figur 12, at det er tre av seks klinikerne som mener at det er forøkte lungelyder, men de tre klinikerne er ikke helt enige om hvor de hører avvik. På kalv 11, figur 13, er det en kliniker som har funnet avvik på tre punkter, de fem andre fant ingen avvik. Figur 14 viser beskrivelsene av kalv 22. Her er det fire av seks klinikerne som hevder at de hører avvik, men det er ulikt hvordan de karakteriseres. Det varierer fra avvik på alle seks punktene, en noterer dempede punkter, andre noterer forøkte punkter. Dempede

lyder et sted, og forøkede lyder et annet sted, kan være to sider av samme sak. Hva som er riktig beskrivelse, og hvilken betydning det har i praksis at dette er fullstendig samkjørt, er vanskelig å avgjøre. Det viktigste er antakelig å oppfatte avvik, så vil spesifiseringen kunne påvirke hvordan man eventuelt vurderer prognosen. For å få en mer helhetlig og reell vurdering av dette burde man laget en spesifikk diagnostisk testanalyse av auskultasjon. Da burde man sett på flere kalver, kalver i lik alder, alle oppstallet under helt like forhold med lite støy, og punktene som skulle vurderes burde vært markert på hver enkelt kalv.

Precentage agreement

Ved beregning av prosentvis enighet ble det sett på alle symptomene for å se hvilke symptomer klinikerne var mest enige om. Generelt var PA relativt høy for alle symptomene, fra 0,56 – 1. Den som skilte se ut var ørestilling, noe som kan forklares med at det var kun to variabler, og *M. bovis* som vanligvis forårsaker avviket ikke er påvist i Norge. Ellers var det høyest enighet ved vurdering av allmenntilstand, og lavest ved vurdering av auskultasjon og hoste. Ved beregning av PA for totalscore på kalvene for alle klinikerne var verdien vesentlig lavere. Veterinærene hadde noe høyere PA enn studentene, men uansett var PA lav sammenlignet med enigheten per symptom.

Cohens kappa

Selv om scoringssystemet var gjennomgått på forhånd og alle kalvene ble undersøkt på samme dag, var det lav kappa mellom så å si alle klinikerne ved vurdering av totalscore. Ved grense >2 var det sjeldent kappa var høyere enn 0,5, det vil si svak enighet (figur 15). Det er så vidt unntak. Kliniker 1 og 5, begge studenter, hadde $\kappa = 0,62$. Det vil si moderat enighet. Ved å øke grensa for hva som regnes som sjuke kalv til >3 totalscore ville vi anta at kappa økte. Dette var derimot ikke tilfelle, da den i dette forsøket heller ble svakere. Her var de fleste κ under 0,3, det vil si minimal enighet. Det var også flere tilfeller av negativ verdi på κ , det vil si en stor uenighet (McHugh, 2012). Når man ser på kappa mellom klinikerne og studenter så er $\kappa = 0,569$ for grensa > 2. Dette tilsvarer en svak enighet mellom gruppene (figur 15). Men også her blir kappa lavere ved høyere grense. Utreget PI er jevnt over høy. Den høye PI-verdien viser til ubalanse i matrisen og en for homogen populasjon.

Fleiss kappa

Fleiss kappa (tabell 9) viser til hvor enig hele gruppen med klinikere var på scoringen av kalvene. Enigheten var lav. Generelt ble enigheten lavere når man økte grensen for cut-off.

Veterinærene som en gruppe var mer enig enn studentene. I tillegg økte deres kappaverdi ved høyere cut-off (>4), mens studentene fikk da lavere enighet. Av kappaverdiene basert på symptomene hadde respirasjon høyest verdi, til tross for stor spredning i søylediagrammet. Det vil si at klinikerne var mer enige i scoringen av de ulike kalvene i denne kategorien.

Resultatene viser at vi generelt får en høyere PA enn kappa. PA tar ikke hensyn til tilfeldigheter eller skjevheter i populasjonen, og blir i dette forsøket misvisende. Verdiene av PA blir derfor falskt høye hvis man sammenligner med kappa.

Det er flere grunner til at kappa blir så lav i dette forsøket. Det er en veldig homogen besetning (Schlageter-Tello et al., 2014), der de fleste kalvene er friske. De få dyrene som klinikerne klassifiserte som sjuke, var de ikke enige om, og dette fikk dermed stor påvirkning på kappa. Dette var en veldig frisk besetning, og de små avvikene som fantes hos kalvene fikk større betydning i scoringen her enn normalt ute i felt. Dermed er det vanskelig å få et korrekt tall på hvor enige vi hadde vært i vurdering av kalver som har luftvegssjukdom. For å få en bedre vurdering av enigheten mellom klinikere på dette burde man gjort et nytt forsøk. Man burde da vurdert en populasjon med kalver med mer spredning i symptomer og større andel kalver som faktisk er sjuke. Ideelt sett dratt ut i flere besetninger med utbrudd av luftvegssjukdom.

Tidligere studier som har sett på vurdering av luftvegssjukdom har også konkludert med at det er lav enighet mellom klinikere. I artikkelen som ble publisert i 2021 (Berman et al., 2021) konkluderer de med at det kun er symptomene indusert hoste og ørestilling eller headtilt som vurderes å ha klinisk akseptabel enighet, κ og $AC1 > 0,6$ og $PA > 0,75$. Gwet's agreement coefficient type 1 ($AC1$) er en koeffisient som vurderer enigheten mellom flere observatører, men som er mer pålitelig i homogene populasjoner enn κ (Berman et al., 2021). I den studien ble kalvene undersøkt kun en gang, deretter ble de kliniske tegnene vurdert på film av 18 andre personer. Dermed er påvirkningen av kalvene av andre observatører så å si neglisjerbar, det samme gjelder stress og hvor slitne kalvene ble etter undersøkelsen. Dette er noe som antakelig hadde stor påvirkning på forsøket som ble utført i forbindelse med vår oppgave, i tillegg til at vi beregnet κ , til tross for at det var en homogen populasjon vi vurderte. Berman et al. så ikke på totalscore for sykdom på hver kalv, noe det ble lagt vekt på i både studien til Buczinski et al. (2016) og i denne oppgaven. Berman et al. så kun symptomene hver for seg, og skrev i konklusjonen at det var kun indusert hoste, ørestilling og rektaltemperatur som kan

brukes til å senere utvikle et pålitelig scoringssystem. I artikkelen til Buczinski et al. og i vår oppgave ble det beregnet totalscore og vurdert enighet mellom observatører ved ulike cut-off verdier for sjuke vs frisk kalv. De fant stor uenighet blant observatørene på dette, noe resultatene fra vårt eget forsøk også viste.

Til tross for ulike studiedesign viser også disse tidligere studiene at det er lav enighet mellom klinikere i vurdering av luftvegssymptomer (Amrine et al., 2013; Berman et al., 2021; Buczinski S et al., 2016a). Dette tyder på at klinisk vurdering av luftvegssymptomer hos kalv er veldig subjektivt og vanskelig å standardisere.

Cut- off og scoringssystemet

Ut i fra datamaterialet har det blitt beregnet forskjellige kappaverdier basert på ulike cut – off på totalscoren. Disse cut- offene ble satt relativt lavt på grunn av få symptomer og funn i besetningen. Men hva ville reel cut- off for sykdom og nødvendigvis behandling ha vært ute i felt?

Scoringssystemet vektlegger de ulike scoringene like mye. Det vil si at en liggende kalv (score 4), teller like mye som en annen kalv med høy respirasjon (75- 85/min = score 4). Scoringssystemet er følsomt, og en oppjaget kalv kan score høyt ut i fra tabellen, selv om kalven ikke nødvendigvis er syk. Med tanke på respirasjonen er det kanskje mer aktuelt å se på respirasjonsmønsteret som et avvik enn frekvensen. Et abdominalt respirasjonsmønster kan være et tydeligere tegn på sykdom enn høy respirasjonsfrekvens. Videre gir purulent neseflod score kun på 2, selv om det er et viktig klinisk symptom i utgangspunktet. Derimot i vårt scoringssystem ga serøs neseflod en score på 1. Dette er kanskje er litt vel strengt, i og med at friske kalver kan ha litt serøs neseflod naturlig. Det samme gjelder øyeflod. Hver kalv må derfor vurderes individuelt. I tillegg er temperaturen en viktig faktor. Klinkere ute i felt vektlegger trolig feber høyt. Vi valgte å se bort i fra temperatur i scoringssystemet for dette prosjektet siden måling av temperaturen er objektivt. Da kunne dyrets naturlige svingninger i temperatur også slått ut på scoringen. De verdiene vi har brukt som cut-off i forsøket er nok mye lavere enn det som ville vært vurdert som syke kalver ute i felt. Med andre ord, den reelle cut- off ute i felt er ikke lett å avgjøre, og det finnes heller ingen fasit. For å kunne estimere et cut- off for behandling, bør det baseres på observasjoner av flere klinkere og kalver ute i felt, som faktisk har luftvegssymptomer.

Feilkilder

Undersøkelse av kalvene

Til tross for felles gjennomgang av undersøkelsen, de ulike symptomene og scoringssystemet før start, er det sannsynlig at det ble noen individuelle forskjeller. Punktene for auskultasjon kan oppfattes ulikt og derfor vurderes forskjellig. Om man måtte holde fast kalven for å telle respirasjonen ved auskultasjon istedenfor å se den fra avstand påvirker også respirasjonsfrekvensen. Klinikernes oppfattelse av hvilende versus nedstemt kalv er også ulikt, antakelig mye avhengig av tidligere erfaring. I tillegg antar vi at siden vi gikk parvis rundt i fjøset, registrerte de som så på kalven samtidig det samme. For eksempel vil litt serøs neseflod kunne oppfattes av de to som så kalven først, mens kalven kan ha fjernet det til neste par vurderte den samme kalven.

Påvirkning av kalvene

Kalvene ble veldig påvirket av håndteringen. De ble oppjaget, noe som gjorde dem vanskelige å undersøke, for eksempel respirasjonsfrekvens. Noen telte respirasjonsfrekvensen ved observasjon, mens andre brukte stetoskop. Det er vanskelig å si hvor stor påvirkning dette har på resultatet. Respirasjonsfrekvensen ble også i stor grad påvirket av håndteringen. Noen av kalvene ble veldig oppjaget og fikk dermed høyere respirasjonsfrekvens. I løpet av dagen ble de yngste kalvene slitne og tok seg en lur, noe som gjør at de for noen av klinikerne kan slå ut på allmenntilstand, «liggende kalv», uten at de egentlig var nedstemt.

Fjøset og miljøforhold

Kalvene var oppstallet på ulike måter, de unge sto i enkeltbokser og de eldre stod i fellesbinger. På den måten ble det varierende hvordan hver enkelt kalv ble undersøkt. I det ene fjøset var det også en del støy fra vifter og melkingsanlegg, noe som blant annet hadde innvirkning på hvor godt man klarte å auskultere lungene. Forskjell i lysforhold kan ha hatt påvirkning på hvordan man registrerte utflod fra nese og øyne. Temperaturen i fjøset kan også ha innvirkning på kalvenes respirasjonsfrekvens, men det ble ikke utført målinger på romtemperaturen.

Homogen besetning

Dette var en besetning uten noe særlig luftvegsproblematikk. Dermed var det totalt sett lite funn, og lite variasjon i dataene. Når det er en homogen besetning og liten prevalens av

sjukdom og ulike symptomer, blir beregning av kappa noe upresis og vanskelig å vurdere (Buczinski S et al., 2016a). Dersom nesten alle observasjoner er i samme klasse, noe som gjelder vårt prosjekt, blir kappa lav selv om samsvaret mellom klinkerne kan være bra (Lydersen, 2018). Det tilsvarende konkluderer studiet fra New Zealand med (Buczinski S et al., 2016a). Man kan anta at i et utbrudd med luftvegssjukdom blir det tydeligere symptomer å vurdere, og dermed mer reelle tall.

Konklusjon

I studien vår har vi undersøkt enigheten mellom observatører ved klinisk vurdering av luftvegssymptomer på kalv, ved bruk av et scoringssystem. Studien viser en høy PA (0,56 – 1) for de ulike symptomene mellom klinkerne, men har en lavere PA for totalscore mellom klinkerne (<0,30). Basert på κ - verdiene er det generelt lav enighet mellom observatørene. Det er noe høyere enighet mellom de erfarne klinkerne enn studentene. Enigheten mellom veterinærene og studentene ble lavere når man satte en høyere grense for cut- off. Resultatene viser store forskjeller mellom bruk av PA og κ , der κ gir en mer reel verdi av enighet som tar hensyn til tilfeldigheter.

Det er vanskelig å konkludere med om de lave κ - verdiene skyldes en dårlig enighet eller om det er på grunn av at besetningen er for homogen. Men også tidligere studier har vist at det er vanskelig å få en høy enighet ved vurdering av luftvegssymptomer hos kalver. Ettersom denne studien baserer seg på få kalver med symptomer på luftvegssjukdom, så kan det ikke trekkes generelle slutninger av funn i dette forsøket som kan brukes i videre prosjekter.

Takk til bidragsyttere

Vi ønsker å takke Veslemøy Sunniva Oma for veiledning gjennom skriveprosessen, samt bearbeiding av data. Vi vil også takke Maria Stokstad for veiledning i planlegging og utførelse av forsøksdagen. Vi ønsker også å takke Lise Marie Ånestad, Thea Klem og Aksel Heggernes Njaa som bidro på feltdagen i prosjektet som klinikere. Til slutt vil vi takke bonden Jens Thori Kogstad for å stille sin besetning til disposisjon i forsøket.

Summary

Title: Evaluation of the agreement between observers in clinical assessment of respiratory signs in calves

Authors: Cathrine Kaarfald Elvland and Ingvild Karlsen Paulsrud

Supervisor: Veslemøy Sunniva Oma og Maria Stokstad

Respiratory disease are common in calves and have a major animal welfare and economic consequences. This disease group are the one that give the most reported cases of disease in calves in Norway. In this study, we examined the agreement between observers in clinical assessment of respiratory signs in calves. This was done to see if clinicians agree on which animals are sick and therefore receiving treatment. Using a pre- determined scoring system, six observers assessed the same 28 calves in one herd.

To examine the agreement between the observers, the observations were converted to a total score. In this way, the agreement could be calculated using digital aids. The percentage agreement (PA) and kappa values (Cohens kappa & Fleiss kappa) were calculated to get a measure of the agreement between the observers.

In this study, there was generally low agreement between the observers. That means the clinicians scored differently on the clinical signs of the different calves. This may be because the calves had few respiratory signs, and therefor the herd becomes too homogeneous to give a good agreement between observers.

Kilder

- Amrine, D. E., White, B. J., Larson, R., Anderson, D. E., Mosier, D. A. & Cernicchiaro, N. (2013). Precision and accuracy of clinical illness scores, compared with pulmonary consolidation scores, in Holstein calves with experimentally induced *Mycoplasma bovis* pneumonia. *American Journal of Veterinary Research*, 74 (2): 310-315. doi: 10.2460/ajvr.74.2.310.
- Berman, J., Francoz, D., Dufour, S. & Buczinski, S. (2021). Evaluation of inter-rater agreement of the clinical signs used to diagnose bovine respiratory disease in individually housed veal calves. *Journal of Dairy Science*, 104 (11): 12053-12065. doi: <https://doi.org/10.3168/jds.2021-20503>.
- Blakebrough-Hall, C., McMeniman, J. P. & González, L. A. (2020). An evaluation of the economic effects of bovine respiratory disease on animal performance, carcass traits, and economic outcomes in feedlot cattle defined using four BRD diagnosis methods. *Journal of Animal Science*. doi: 10.1093/jas/skaa005.
- Buczinski S, Faure C, Jolivet S & Abdallah A. (2016a). Evaluation of inter-observer agreement when using a clinical respiratory scoring system in preweaned dairy calves. *New Zealand Veterinary Journal*. doi: 10.1080/00480169.2016.1153439.
- Buczinski S, Faure C, Jolivet S & Abdallah A. (2016b). Evaluation of inter-observer agreement when using a clinical respiratory scoring system in pre weaned dairy calves. *New Zealand Veterinary Journal*. doi: 10.1080/00480169.2016.1153439.
- Campbell, J. (2016). *Bacterial pneumonia in Cattle*. msdvetmanual.com: MSD Manual Veterinary Manual. Tilgjengelig fra: <https://www.msdvetmanual.com/respiratory-system/respiratory-diseases-of-cattle/bacterial-pneumonia-in-cattle?query=Pasteurella%20multocida%20cattle>.
- Confer, A. W. (2009). Update on bacterial pathogenesis in BRD. *Cambridge University press*. doi: 10.1017/S1466252309990193.
- Falk, M., Reiersen, A., Klem, T. B., Jonsson, M. & Heier, B. T. (2022). *Dyrehelserapporten 2021*. Veterinærinstituttets rapportserie 19/2022.
- Flight, L. & Julious, S. A. (2015). The disagreeable behaviour of the kappa statistic. *Pharmaceutical Statistics*, 14 (1): 74-78. doi: <https://doi.org/10.1002/pst.1659>.

- Klem, T. B., Rimstad, E. & Stokstad, M. (2014). Occurrence and phylogenetic analysis of bovine respiratory syncytial virus in outbreaks of respiratory disease in Norway. *BMC Veterinary Research*, 10 (1): 15. doi: 10.1186/1746-6148-10-15.
- Klem, T. B., Stokstad, M. & .. (2020). *Norwegian Airways full application*: Veterinærinstituttet.
- Klem, T. B. B. (2021). *Norwegian Airways*: Veterinærinstituttet. Tilgjengelig fra: <https://www.vetinst.no/forskning-innovasjon/pagaende-forskningsprosjekter/norwegian-airways> (lest 20.08.21).
- Love, W. J., Lehenbauer, T. W., Kass, P. H., Van Eenennaam, A. L. & Aly, S. S. (2014). Development of a novel clinical scoring system for on-farm diagnosis of bovine respiratory disease in pre-weaned dairy calves. *PeerJ (San Francisco, CA)*, 2: e238-e238. doi: 10.7717/peerj.238.
- Lydersen, S. (2018). *Cohens kappa - et mål på samsvar mellom observatører*. tidsskriftet.no: Tidsskriftet - Den norske legeforening. Tilgjengelig fra: <https://tidsskriftet.no/2018/03/medisin-og-tall/cohens-kappa-et-mal-pa-samsvar-mellom-observatorer> (lest 11.05).
- Mattilsynet. (2014). *Bovin respiratorisk syncytial virus (BRSV)*: mattilsynet.no. Tilgjengelig fra: https://www.mattilsynet.no/dyr_og_dyrehold/dyrehelse/dyresykdommer/brsv/ (lest 03.12).
- McGuirk, S. M. (2014). Timely diagnosis of dairy calf respiratory disease using standardized scoring system. *Animal Health Research Reviews*. doi: 10.1017/S1466252314000267.
- McHugh, M. L. (2012). Interrater reliability: the kappa statistic. *Biochemia Medica*.
- Schlageter-Tello, A., Bokkers, E. A. M., Groot Koerkamp, P. W. G. & Van Hertem, T. (2014). Manual and automatic locomotion scoring systems in dairy cows: A review. *Preventive Veterinary Medicine*, 116 (1-2): 12-25. doi: <https://doi.org/10.1016/j.prevetmed.2014.06.006>.
- Stokstad, M., Klem, T. B., Myrmel, M., Oma, V. S., Toftaker, I., Østerås, O. & Nødtvedt, A. (2020). Using Biosecurity Measures to Combat Respiratory Disease in Cattle: The Norwegian Control Program for Bovine Respiratory Syncytial Virus and Bovine Coronavirus. *Frontiers in Veterinary Science*, 7. doi: 10.3389/fvets.2020.00167.
- Timsit, E., Dendukuri, N., Schiller, I. & Buczinski, S. (2016). Diagnostic accuracy of clinical illness for bovine respiratory disease (BRD) diagnosis in beef cattle placed in feedlots: A systematic literature review and hierarchical Bayesian latent-class meta-analysis,.

Preventive Veterinary Medicine, 135: 67-73. doi:

<https://doi.org/10.1016/j.prevetmed.2016.11.006>.

Tine. (2014). *Luftveisinfeksjon hos storfe - årsker og forebyggende tiltak*. Animalia.no.

Upublisert manuskript.

Toftaker, I. (2019). *Epidemiologic studies of bovine respiratory syncytial virus and bovine coronavirus in Norway*: Norwegian University of Life Sciences.

Underwood, W. J., Blauwiel, R., Delano, M. L., Gillesby, R., Mischler, S. A. & Schoell, A. (2015). *Biology and Diseases of Ruminants (Sheep, Goats and Cattle)*. I: *Laboratory Animal Medicine*: Elsevier Inc.

Valarcher, J.-F. T., Geraldine. (2007). Bovine respiratory syncytial virus infection. *Veterinary Research*, 38.

Veterinærinstituttet. (2022). *Mycoplasma bovis*. vetinst.no: Veterinærinstituttet. Tilgjengelig fra: <https://www.vetinst.no/sykdom-og-agens/mycoplasma-bovis> (lest 19.04).

Zapf, A., Castell, S., Morawietz, L. & Karch, A. (2016). Measuring inter-rater reliability for nominal data – which coefficients and confidence intervals are appropriate? *BMC Medical Research Methodology*, 16 (1): 93. doi: 10.1186/s12874-016-0200-9.



Norges miljø- og
biovitenskapelige
universitet

Postboks 5003
NO-1432 Ås
67 23 00 00
www.nmbu.no