

**Skema til afrapportering af ViD projekter**  
Videncenter for Dyrevelfærd  
2013

**1. Projekttitle:**

Validering samt kvalitets- og fremtidssikring af velfærdsregistreringer i svinebesætninger

**2. Projektleder og projektdeltagere:** (titel, navn, affiliation)

Projektleder: Professor Nils Toft, Institut for Produktionsdyr og Heste, Københavns Universitet

Deltagere: Sine Norlander Andreasen, Hans Houe, Bjørn Forkman, Institut for Produktionsdyr og Heste, Københavns Universitet.

Kontaktperson: Tina Birk Jensen, Videncenter for Dyrevelfærd, Fødevarestyrelsen.

Desuden har et antal medarbejdere fra FVST deltaget både i forbindelse med indsamling af billedmateriale, samt ved efterfølgende vurdering af disse materialer.

**3. Populærvidenskabeligt dansk resumé (max 250 ord):**

Det overordnede formål med projektet er at kunne validere og kvalitetssikre velfærdsregistreringer i svinebesætninger. Projektet skal skabe mulighederne for at danne et objektive og sammenlignelige billede af både det nuværende, tidligere og fremtidige niveau af dyrevelfærdsmæssige overtrædelser.

Projektet består af tre faser. I første fase er der indsamlet video- og billedmateriale, der fremstiller forskellige situationer/scenarier med hensyn til rode- og beskæftigelsesmaterialer (RB), opstaldning (O), herunder sygestier, og dyrene (D), herunder brok, aflivning og syge dyr. I anden fase er det indsamlede billedmateriale anvendt ved en række workshops, hvor teknikere og dyrlæger har vurderet hvert enkelt tilfælde med hensyn til evt. overtrædelse og sanktion. I sidste fase er disse vurderinger anvendt i en matematisk model, hvor sensitivitet og specificitet for hver kontrollant udregnes.

I fase 1 blev der indsamlet 176 film og 127 billeder, disse blev reduceret til 63 cases (D), 74 cases (O) og 29 cases (RB). Der blev afholdt to workshops, ved den første blev O og RB testet af 5 dyrlæger og 2 teknikere, ved den anden blev D og RB testet af 4 dyrlæger og 1 tekniker. Analyser af sensitivitet og specificitet viser, at der generelt er en meget høj sensitivitet og specificitet (>90%) i vurderingerne af video- og billedmateriale. Feedback fra begge workshops var til gengæld, at materialet ikke tilgodeser behovet for en helhedsvurdering, men udelukkende fokuserer på enkelte situationer, hvilket ikke afspejler den virkelighed kontrollanterne føler de opererer i. Der skal derfor arbejdes med muligheder for at integrere materialet i en operationel helhedsvurdering af besætningerne.

**4. Populærvidenskabeligt engelsk resumé (max 250 ord):**

The primary purpose of the project was validation and quality assurance of the welfare recordings

in Danish pig herds. Furthermore, the project established a basis for an objective and comparable assessment of the former, current and future level of welfare violations. The project consisted of three phases. In phase 1, video and pictures illustrating various situations with a potential welfare problem. Three overall themes were illustrated: rooting and enrichment materials, housing (including sick pens) and the animals (including sick animals and euthanasia). In phase 2, the collected materials were presented to the welfare inspectors (veterinarians and technicians), who judged the potential violation and its consequence. In phase 3 these assessments were used in a mathematical model to estimate the sensitivity and specificity of the observers. From phase 1, 176 videos and 127 stills were obtained. These were subsequently reduced to 63 cases (Animals), 74 cases (Housing) and 29 cases (Rooting). These cases were used at two workshops, where 9 veterinarians and 3 technicians evaluated parts of the materials. The statistical analysis of the results showed that generally the sensitivity and specificity was very high (>90%) when evaluating the observers based on the video and pictures. However, feedback from both workshops was that the observers did not feel that the materials adequately reflected the overall farm assessment that the welfare inspectors felt were part of their daily work. Hence, further work is needed for integrating the materials in an overall farm assessment.

## **5. Videnskabeligt dansk resumé af projektets formål, udførelse, væsentligste resultater og konklusion (max 500 ord):**

### Baggrund:

Der er en stor forståelse for, at velfærdsvurderinger i husdyrbesætninger skal være så korrekte som muligt. Derfor foregår der også kontinuerligt træning og kalibrering gennem fælles møder og workshops for, på denne måde, at få ensrettet kontrollanternes opfattelse af hvad der udgør et problem. Ved at systematisere disse træningssessioner vil det være muligt at indhente information til validering af kontrollanterne. Både, hvor enige disse er, men også, hvor præcist de henholdsvis påpeger problemer (sensitivitet) og vurderer, at der ikke forekommer overtrædelser i besætningen (specificitet). Kun når man kender disse kvalitetsparametre, kan man udregne og sammenligne den sande forekomst af velfærdsproblemer i 2 forskellige velfærdsundersøgelser, som fx nulpunktundersøgelser og risikobaseret velfærdskontrol, når der korrigeres for evt. forskelle mellem kontrollanter.

### Formål:

Det overordnede formål med projektet er at kunne validere og kvalitetssikre velfærdsregistreringer i svinebesætninger. Projektet skal skabe mulighederne for at danne et objektivt og sammenligneligt billede af både det nuværende, tidligere og fremtidige niveau af dyrevelfærdsmæssige overtrædelser i danske svinebesætninger.

### Gennemførelse:

Projektet består af tre faser:

I første fase blev der indsamlet video- og billedmateriale, der fremstiller forskellige situationer/scenarier med hensyn til rode- og beskæftigelsesmaterialer (RB), opstaldning (O), herunder sygestier, og dyrene (D), herunder brok, aflivning og syge dyr. Dette materiale blev

indsamlet ved kontrolbesøg foretaget i forbindelse med nulpunktsundersøgelsen for svin, der blev udført af Fødevarestyrelsen (FVST) i 2011 (Houe, 2012), og repræsenterer derfor aktuelle situationer, der forekommer i tilfældigt udvalgte besætninger.

I anden fase blev det indsamlede billedmateriale anvendt ved 2 workshops, hvor teknikere og dyrlæger fra FVST vurderede hvert enkelt tilfælde med hensyn til evt. overtrædelse og sanktion. De blev bedt om at vurdere ud fra den gældende instruks om vejledning til velfærdskontrol i svinebesætninger, og for hver situation angive, hvad situationen vurderes at have af betydning i forhold til lovgivning og instruks, samt hvilken sanktion, vurderingen skulle medføre.

I tredje fase blev disse vurderinger anvendt i en latent klasse model, hvor sensitivitet og specificitet for hver kontrollant udregnes uden at der defineres en reference kontrollant. Det vil sige, uden at den absolutte sandhed angående en vurdering er kendt.

#### Resultater:

I fase 1 blev der indsamlet 176 film og 127 billeder, disse blev reduceret til 63 cases (D), 74 cases (O) og 29 cases (RB). Der blev afholdt to workshops, ved den første blev O og RB testet af 5 dyrlæger og 2 teknikere, ved den anden blev D og RB testet af 4 dyrlæger og 1 tekniker. Analyser af sensitivitet og specificitet viser, at der generelt er en meget høj sensitivitet og specificitet (>90%) i vurderingerne af video- og billedmateriale. Feedback fra begge workshops var, at man ikke mente materialet tilgodeser behovet for en helhedsvurdering, men udelukkende fokuserer på enkelte situationer, hvilket ikke afspejler den virkelighed kontrollanterne føler de opererer i.

#### Konklusion:

I forhold til andre publicerede estimater af sensitivitet og specificitet af velfærdsvurderinger, som fx halhedsvurderinger eller kødkontrol, så er kontrollanterne særdeles gode til at vurdere situationerne korrekt og der er stor enighed mellem kontrollanterne. Der skal arbejdes videre med muligheder for at integrere materialet i en operationel helhedsvurdering af besætningerne.

### **6. Beskrivelse af projektets formål, evt. hypoteser, og materialer og metoder: Formål:**

Det overordnede formål med projektet er at kunne validere og kvalitetssikre velfærdsregistreringer i svinebesætninger. Projektet skal skabe mulighederne for at danne et objektivt og sammenligneligt billede af både det nuværende, tidligere og fremtidige niveau af dyrevelfærdsmæssige overtrædelser.

#### Materialer og metoder:

Projektet består af tre faser:

I første fase blev der indsamlet video- og billedmateriale, der fremstiller forskellige situationer/scenarier med hensyn til rode- og beskæftigelsesmaterialer (RB), opstaldning (O), herunder sygestier, og dyrene (D), herunder brok, aflivning og syge dyr. Dette materiale blev

indsamlet ved kontrolbesøg foretaget i forbindelse med nulpunktsundersøgelsen for svin, udarbejdet af FVST i 2011 (Houe, 2012) og repræsenterer derfor aktuelle situationer, der forekommer i tilfældigt udvalgte besætninger.

I anden fase blev det indsamlede billedmateriale anvendt ved 2 workshops, hvor teknikere og dyrlæger fra FVST vurderede hvert enkelt tilfælde med hensyn til evt. overtrædelse og sanktion. De blev bedt om at vurdere ud fra den gældende instruks om vejledning til velfærdskontrol i svinebesætninger og for hver situation angive, hvad situationen vurderes at betyde: Ingen anmærkninger, Væsentlig ulempe, Uforsvarlig behandling af dyr, Groft uforsvarlig behandling af dyr, Groft uforsvarlig behandling af dyr med karakter af mishandling. Desuden skulle de angive hvilken sanktion, vurderingen skulle medføre: Ingen anmærkninger; Indskærpelse; Påbud; Politianmeldelse. Dette blev vurderet at være tæt på den vurdering, der bør gennemføres i en konkret situation og dermed lettest genkendelig for kontrollanterne.

I tredje fase blev disse vurderinger anvendt i en latent klasse model, hvor sensitivitet (Se) og specificitet (Sp) for hver kontrollant udregnes uden, at der defineres en reference kontrollant. Det vil sige, uden den absolutte sandhed angående en vurdering er kendt. For at kunne udregne sensitivitet og specificitet blev vurderingerne aggregeret til overtrædelse ja/nej.

Den anvendte model er i store træk den samme, som der er anvendt i (Baadsgaard & Jørgensen 2003), dog med nogle få udvidelser. I korte træk er princippet i latent klasse modellen flg.: Det antages, at alle situationerne har en sand, men ukendt (latent), tilstand, således at der enten er tale om en sand overtrædelse af velfærdsreglerne eller en korrekt håndtering/situation. I modsætning til den klassiske latent klasse model (Hui & Walter 1980) antages det ikke, at kontrollanter er betinget uafhængige givet den sande tilstand. Derimod antages det, at de stammer fra den samme fordeling via en tilfældig effekt model for hhv Se og Sp. Vi har dog antaget, at dyrlæger stammer fra én fordeling, mens teknikere stammer fra en anden fordeling. Denne antagelse blev gjort for at kunne estimere eventuelle systematiske forskelle mellem disse to typer af kontrollanter. Som følge af det relativt lille antal kontrollanter, blev der valgt at anvende en model, hvor variationen mellem kontrollanter (beskrevet ved variansen på den tilfældige effekt) var den samme for de tre grupper af vurderinger (RB, O og D), mens Se og Sp for den enkelte kontrollant kunne variere mellem de tre grupper. For at kunne bestemme parametrene i denne type af latent klasse modeller udelukkende ved hjælp af data er det nødvendigt med mindst 3 kontrollanter af hver type (dyrlæge eller tekniker), der har bedømt de samme billeder.

#### Referencer:

Baadsgaard, N P, and E Jørgensen. 2003. A Bayesian Approach to the Accuracy of Clinical Observations. *Prev Vet Med* 59 (4) (June): 189–206..

Houe, H, 2012. Nulpunktsundersøgelsen – overvågning på baggrund af tilfældig udpegning. I: 2011 – Dyrevelfærd i Danmark, Videncenter for Dyrevelfærd. Fødevarestyrelsen, 18-20

Hui, S L, and S D Walter. 1980. Estimating the Error Rates of Diagnostic Tests. *Biometrics* 36 (1) (March): 167–71..



## 7. Oversigt over projektets samlede resultater:

### Fase 1:

I fase 1 blev der i alt besøgt 11 besætninger i perioden september – oktober 2011 i forbindelse med gennemførelse af kontrolbesøg i nulpunktsundersøgelsen. En oversigt over de 11 besætninger er givet i nedenstående tabel:

Besætning	Sted	Type
01	Sjælland (FVST)	Søer (løs, halm)
02	Sjælland (FVST)	Slagtesvin
03	Fyn (FVST)	Slagtesvin
04	Fyn (FVST)	Slagtesvin
05	Fyn (FVST)	Søer (ej løs) + Klima
06	Sjælland (FVST)	Slagtesvin
07	Fyn (FVST)	Søer (løs fast)
08	Fyn (FVST)	Klima + Slagtesvin
09	Fyn (Privat)	Søer (løs, halm)
10	Fyn (FVST)	Slagtesvin (fuldspalte)
11	Fyn (FVST)	Søer, klima, slagt

Indsamlingen af billede- og videomateriale i besætningerne foregik uafhængigt af kontrolbesøget og kun efter forudgående accept fra landmanden. Besætning 9 er dog besøgt i privat sammenhæng for at skaffe materiale fra den type besætning også. Indsamlingen fokuserede på at skaffe materiale, der fremstiller forskellige situationer/scenarier med hensyn til rode- og beskæftigelsesmaterialer (RB), opstaldning (O), herunder sygestier, og dyrene (D), herunder brok, aflivning og syge dyr, da disse var vurderet at være mest realistiske at illustrere i den givne kontekst. I alt blev der indsamlet 176 film og 127 billeder, disse blev efterfølgende reduceret til 63 cases (D), 74 cases (O) og 29 cases (RB) med en forventet fordeling af velfærdsovertrædelser omkring 50% i hver af de tre grupper.

### Fase 2:

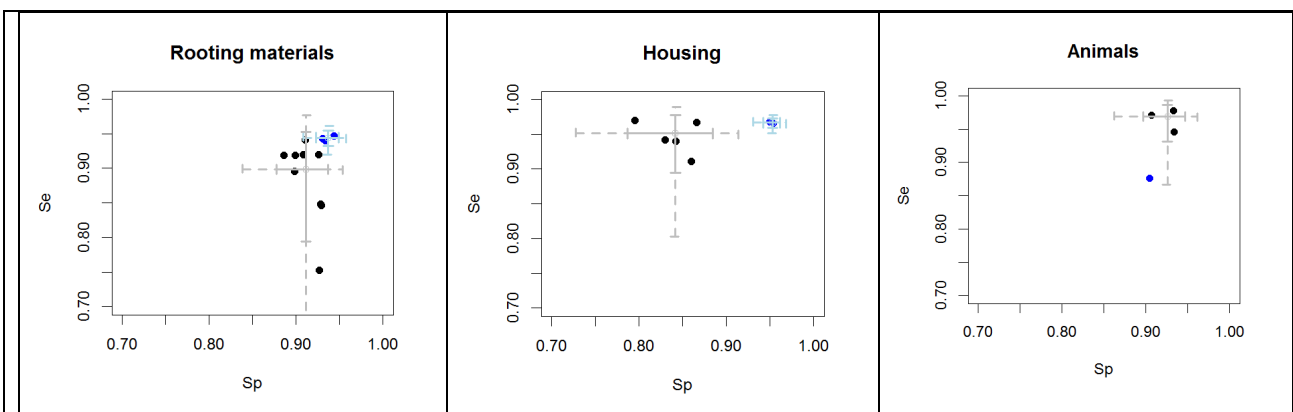
I den oprindelige projektplan indgik udvikling af et databaseværktøj til opbevaring og fremvisning af cases, samt indtastning af vurderinger fra kontrollanterne. I forbindelse med fase 2 blev der undersøgt forskellige muligheder for implementering af et sådant system, men vurderingen var, at der var behov for at teste materialet inden der blev anvendt ressourcer på opbygning af en databasestruktur. Ved testfremvisninger af udvalgte dele af materialet for medarbejdere i FVST (Glostrup), blev der givet udtryk for et behov for yderligere information; fx om tilstanden der blev vist i materialet var generel i besætningen osv. Derfor blev cases tilføjet yderligere information og samlet i 3 Powerpoint filer med tilhørende PDF versioner af afkrydsningsskemaer til vurderinger. Materialet blev testet på 2 workshops. Den første blev afholdt i team svin i veterinærenhed nord i Herning og havde deltagelse af 6 dyrlæger og 2 teknikere, mens den anden blev gennemført i team svin i veterinærenhed syd med deltagelse af 3 dyrlæger og 1 tekniker. De i alt 12 personer (11 kvinder/1 mand) havde 1-16 års erfaring (gns 6 år) med velfærdsvurderinger/kontrolbesøg.

Det viste sig meget hurtigt, at det er krævende og trættende at vurdere materialet, og det var ikke muligt at gennemgå alt i en endags workshop. Derfor gennemgik workshop 1 materialet omkring rode- og beskæftigelsesmaterialer samt opstaldning, mens workshop 2 gennemgik materialet omkring dyrene, samt rode- og beskæftigelsesmaterialer. Det blev antaget, at man derved kunne anvende vurderingerne af rode- og beskæftigelsesmaterialer som en slags baseline eller kalibrering mellem workshops.

Udover indsamling af selve vurderingerne, havde de 2 workshops til formål at få vurderet materialet og processen. Derfor blev forløbet evalueret mundtligt efter afslutningen af hver workshop. Der var generelt en afvisning af metoden og forløbet som en realistisk simulering af en reel arbejdssituation. Specielt var der stor frustration over, at der ikke ved hjælp af materialet kunne gennemføres en helhedsvurdering af de faktiske forhold i besætningen, som de følte var et centralt element i deres arbejde, specielt i forhold til vurdering af om en overtrædelse skulle noteres og medføre en sanktion. De accepterede dog at vurdere som om der er nul-tolerance ved gennemgangen af materialet. Der var stor tilfredshed med materialet, og der var ønske om at det gøres tilgængeligt til anvendelse ved interne kurser og til diskussion af situationer. Deltagerne mente nærmest samstemmende, at dette ville være en mere realistisk og nyttig anvendelse af materialet.

### Fase 3:

Til trods for den udbredte skepsis med metoden viste resultaterne generelt en høj Se og Sp for kontrollanterne. Disse er gengivet i nedenstående figurer:



(de sorte prikker er dyrlæger, de blå teknikere. Stregerne angiver hvor 70% (fast) og 95% (stiplet) af alle kontrollanter fra samme gruppe forventes at ligge indenfor)

Der ses en del mere variation i dyrlægerne i forhold til teknikere. En del af dette må tilskrives det større antal, men overordnet set ser det ud til, at der er systematiske forskelle mellem dyrlæger og teknikere. Endvidere er tendensen at dyrlægerne har både højere Se og Sp når det drejer sig om konkrete vurderinger af dyrene, mens teknikerne har højst Se og Sp når det drejer sig om at vurdere opstaldning og inventar. Det er datamaterialet for spinkelt til at udtale sig om, hvorvidt det er en generel tendens.

Givet den manglende accept af metoden blandt kontrollanterne, kan der stilles spørgsmålstegn ved hvorvidt konklusionerne af analysen kan overføres til den situation, hvor det er hele besætningen, der indgår i vurderingen.

## 8. Beskrivelse af, hvordan resultaterne bidrager til at opfylde projektets formål:

Tanken var, at de indsamlede cases kunne anvendes til at kalibrere kontrollanter. Ved at lade kontrollanterne, der gennemførte nulpunktundersøgelsen gennemgå materialet kunne man estimere Se og Sp for de enkelte kontrollanter, hvilket ville gøre det muligt at udregne den sande forekomst af overtrædelser. Lod man så kontrollanterne fra FVSTs risikobaserede velfærdskontrol gennemgå de samme cases kunne man justere den risikobaserede velfærdskontrol på samme måde og dermed få mulighed for at sammenligne de 2 undersøgelses *sande* niveau af overtrædelser. Med det indsamlede materiale og den statistiske model er formålet opfyldt, under antagelse af at vurderingerne foretaget på video- og billedmaterialet afspejler de vurderinger, der foretages af sammen situation i en besætning. Det var dog tydeligt at kontrollanterne ikke selv troede på at dette var tilfældet. Resultaterne af projektet må derfor nok anses for at være en begyndelse, men der er behov for at videretænke, hvorledes de enkelte situationer kan integreres i en helhedsbetragtning. En stor udfordring i den henseende er, at de nuværende stikprøver på 63 cases (D), 74 cases (O) og 29 cases (RB) reelt må anses for at være noget under hvad de burde. En yderligere integration mod en helhedsbetragtning ville derfor enten reducere den samlede stikprøve betydeligt, eller øge tidsforbruget til vurderinger i en grad, så det ikke er realistisk at få

gennemført.

### **9. Konklusion og perspektivering:**

Konklusionen af de statistiske analyser er, at kontrollanterne er særdeles gode til at vurdere situationerne ens i forhold til undersøgelsens anvendte billedmateriale. Set i forhold til lignende undersøgelser af kliniske eller adfærdsmæssige vurderinger er kombinationen af de høje Se og Sp yderst tilfredsstillende og tyder på en meget lav grad af subjektivitet i vurderingerne, når disse foretages på video- og billedmateriale.

Det må også konkluderes, at der udestår et stykke arbejde endnu, i at få lavet et set-up, hvor materialet kan vurderes i en kontekst, der kan understøtte en helhedsvurdering.

### **10. Redegørelse for hvordan projektet og projektets resultater har været eller forventes offentliggjort:**

Projektet har været præsenteret som mundtligt indlæg og poster på VID konferencen i 2011.

Projektet har været præsenteret som poster på VID konferencen i 2013.

Projektet har været præsenteret som poster på SVEPM konferencen i 2013.

Et manuskript til ANIMAL er under udarbejdelse forventes indsendt ultimo januar 2014.

En populærvidenskabelig artikel til Dyrlægen eller Hyologisk udarbejdes, når resultaterne er publiceret internationalt.

Video- og billedmateriale fremsendes til VID til internt brug i FVST.