



Rapport Sammendrag

12. september 2018
Proj.nr. 2006247
Version 1/EVO

Visionhjælpeværktøj til ind- og udvendig kontrol

- Kravspecifikation
- Risikovurdering
- Referencedata

Indledning Den nuværende kødkontrol på kyllingeslagterierne gennemføres manuelt af dyrlæger og teknisk personale. Slagtelinjernes høje hastighed (11.500 kyllinger/time) gør dette arbejde meget krævende. Planer om endnu højere linjehastigheder gør ikke udfordringen mindre. Potentialet af automatiske hjælpeværktøjer til kødkontrollen er, at kødkontrolpersonalet kun behøver at kontrollere kyllinger, som de automatiske værktøjer udpeger til kontrol. Hvis denne andel udgør færre end for eksempel 10% af kyllingerne, vil det aflaste personalet betydeligt ikke mindst hvis slagtehastigheden øges. De automatiske registreringer vil samtidig kunne give slagterierne mere information om kyllingerne, og potentielt vil tilbagemeldingen til landmændene kunne blive mere detaljeret med input til optimering af produktionen.

Formål Formålet med dette delprojekt, finansieret af Fjerkræafgiftsfonden, er opstilling af kravspecifikation, udarbejdelse af risikoanalyse og fremstilling af referencedata til brug for kalibrering af måleudstyret – både til indvendig og udvendig kontrol. Opgaven er varetaget af Teknologisk Institut i samarbejde med Fødevarerstyrelsens afdeling for kødkontrol.

Indstilling Hjælpeværktøjet, VET-systemet, er baseret på ihfoods "ClassifEYE" system til klassificering og kvalitetskontrol. En prototype af del-systemet til kontrol af organsæt er implementeret i Danpo/Aars, og en funktionsmodel af systemet til den ydre kontrol er opstillet samme sted. Der er etableret referencedata til begge delsystemer, således at systemets evne til at identificere kvalitetsafvigelse kan analyseres. Det anbefales, at der etableres et repræsentativt datasæt opnået med prototype/funktionsmodel-versionen af systemet. Eventuel afhængighed af leverancer, årstid mv. analyseres. Der udtages en stikprøve til verificering af predikterede hhv. ikke identificerede kvalitetsafvigelse med henblik på fastlæggelse af præcision og målesikkerhed som input til endelig indstilling og anbefaling om færdiggørelse af VET-systemet.

Indhold

Denne rapport indeholder et sammendrag af

- Kravspecifikationen
- Risikoanalysen
- Præcision/kvalitet af referencedata
 - Vurdering af organsæt
 - Vurdering af udvendige defekter

For uddybning af hver af de tre emner henvises til de vedlagte rapporter, se referenceliste.

Kravspecifikation

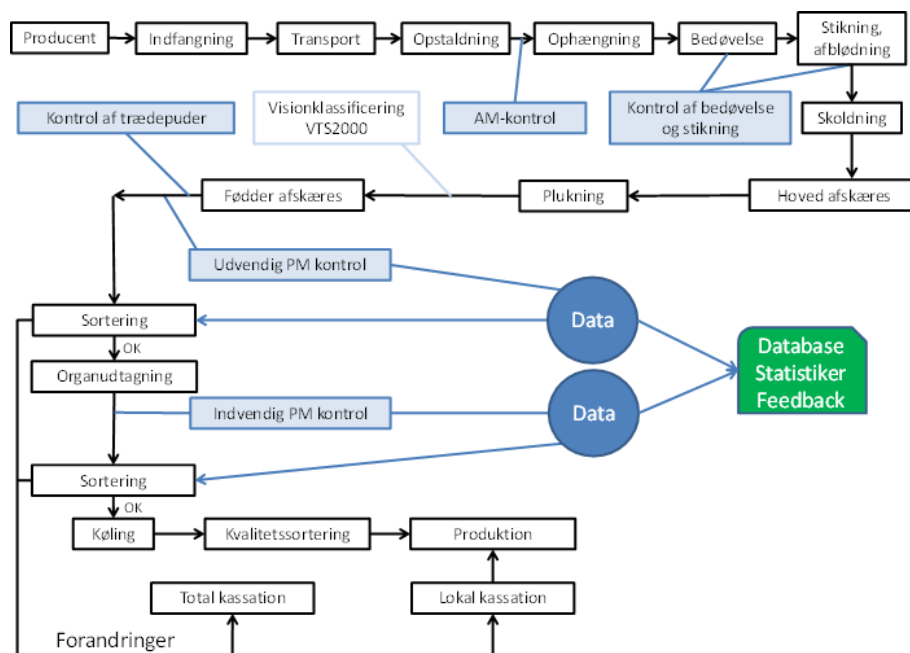
Kravspecifikationen (øverste kravniveau) er vedlagt i den version, som blev vedtaget ultimo 2017. Nedenfor gives et sammendrag af centrale elementer i specifikationen.

Dokumentet specificerer kravene for et udstyr til:

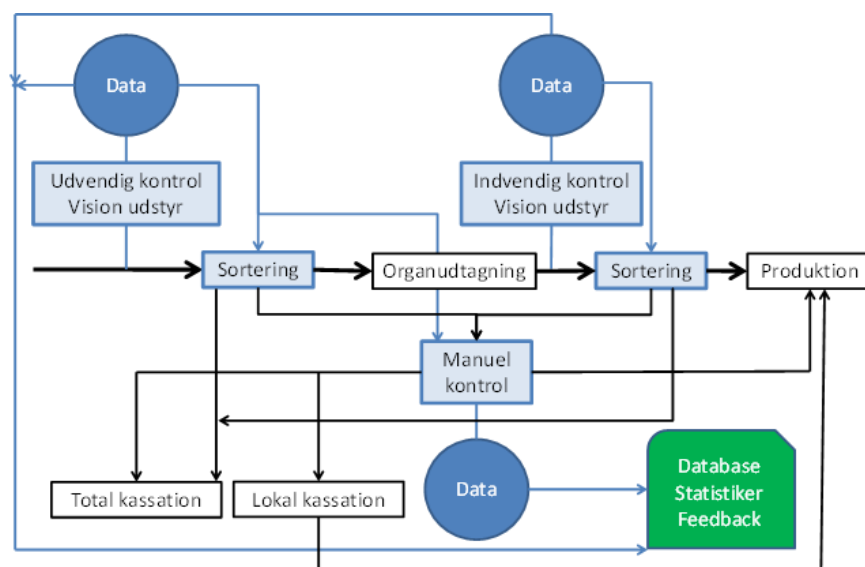
- Automatiseret måling af slagtede kyllingers kvalitetsforandringer (herefter kaldet *forandringer*), som i dag foretages manuelt af i forbindelse med post mortem (PM) kontrollen og
- Sortering af kyllingerne på basis af målingerne

Udstyret skal omfatte forandringer, som indgår i den nuværende udvendige og indvendige PM kontrol. I kravspecifikationen skelnes ikke mellem forandringer fra udvendig og indvendig kontrol, men det angives, hvis en forandring skal findes inden eviscerering (organudtagning). De angivne krav til udstyret inklusive måle-/sorteringsusikkerhed i form af niveau af falsk negative og niveau af falsk positive er for det samlede måle- og sorteringsudstyr inklusiv visionmåling, sortering baseret på måling, manuelle bedømmelser, datahåndtering og endelig sortering. I figur 1 er en skematisk fremstilling af en slagtelinje med angivelse af kontrolpunkter.

Figur 1. Kød-kontrollens placering (markeret med blå) på de danske kyllingeslagterier (HKScan Denmark, Vinderup og Scandi Standard Danpo, Aars)



Figur 2.
Idéskitse til
udstyr omfattende måling
og sortering



Primære output fra måleudstyr

Idéen er, at udstyrets kameraer placeres på de bedst egnede steder på slagtekæden. Et eksempel er vist i figur 2. Måledata skal danne grundlag for sortering af kyllinger inkl. organsæt i minimum tre grupper:

- Kyllinger, som med meget stor sikkerhed kan godkendes.
- Kyllinger, som med meget stor sikkerhed skal totalkasseres.
- Kyllinger, som måleudstyret ikke kan placere i gruppe a eller b. Kyllinger i denne gruppe skal inspiceres manuelt af kødkontrolpersonale, som herefter afgør om kyllingerne skal godkendes, total- eller lokalkasseres.

Der er ikke et specifikt krav til hvor stor en andel af kyllingerne, der af måleudstyret må placeres i gruppe c. Men færre end ca. 5 % af kyllingerne vil være ønskværdigt.

Reference

Referencedata består af billedsæt af kyllinger og organsæt, som er vurderet af tre dyrlæger med hensyn til de forandringer, som i dag indgår i kontrolarbejdet. Det er testet, at billeddata udgør et tilstrækkeligt grundlag for bedømmelsen. Alternativt skulle vurderingen være foretaget på slagtelinjen, hvilket ville have været en betydelig langsommere proces. Se senere.

Måle og sorteringsikkerhed

Udstyrets måle-/sorteringsusikkerhed beskrives ved hjælp af niveauet af falsk negative (NFN) hhv. niveauet af falsk positive (NFP), defineret ved:

- $NFN = \text{antal}(FN) / (\text{antal}(FN) + \text{antal}(SP))$
- $NFP = \text{antal}(FP) / (\text{antal}(FP) + \text{antal}(SN))$

baseret på begreberne *sandt positive (SP)*, *sandt negative (SN)*, *falsk positive (FP)* og *falsk negative (FN)*. *Positiv* betyder, at en kylling har en given forandring, mens *negativ* betyder, at en kylling ikke har den givne forandring.

<i>Måling af forandringer</i>	<p><i>Krav</i></p> <p>Det er et krav, at udstyret kan måle og sortere efter følgende hovedgrupper af forandringer:</p> <ul style="list-style-type: none"> • Afmagring • Misfarvning (inkl. overskoldet) • Bugvattersot • Ledbetændelse/ledforandring • Hudlidelse/Hudbetændelse • Rifter/kradsemærker • Hjertesækbetændelse • Luftsækbetændelse og bughindebetændelse • Akut leverbetændelse • Kronisk leverforandring/-betændelse • Diverse: Forstørret / fyldt kro • Muskelforandringer <ul style="list-style-type: none"> ○ Træbryst ○ Ændringer i rygmuskler ved vingebasis • Røde ikke afblødte • Brud og blødninger • Synlige forureninger <ul style="list-style-type: none"> ○ Gødningsforurening ○ Galdeforurening • Ikke udtagne organer / organrester • Maskinskader fra plukkeren <p>I specifikationen er der for hver forandring angivet, hvilket krav der stilles til NFP og NFN opnået ved hjælp af risikoanalysen.</p>
<i>Kapacitet</i>	<p>Udstyret skal kunne måle og sortere med en hastighed svarende til en linjehastighed på mindst 14.000 kyllinger/time.</p>
<i>Daglig opetid</i>	<p>Udstyret skal være operationelt ved op til maksimal målehastighed i 99,9 % af tiden i en periode over 18 timer pr. dag.</p>
<i>Grænseflade til slagteriets produktionsstyring</i>	<p>Output inkluderer de automatiske målinger og sortering i gruppe a, b og c, indtastede manuelle bedømmelser for gruppe c og den endelige fysiske sortering i godkendte, total- og lokalkasserede. Billeder af kyllinger med forandringer skal kunne gemmes.</p> <p>Udstyret skal kunne modtage input om <i>floknummer</i> fra slagteriets IT-system og levere de registrerede data løbende til slagteriets IT-system til lagring, driftsovervågning og rapportering.</p> <p>Udstyret skal kunne fungere som stand-alone i tilfælde af nedbrud på slagteriets IT-system.</p>
<i>Diverse</i>	<p>Specifikationen omfatter endvidere krav til levetid, service, rengøring. Krav til materialer, ydermål, montering og sikkerhed.</p>

Risikoanalyse

Rapport omhandlende risikoanalyse, version 2, udarbejdet primo 2017, er vedlagt. Selve analysen består af et regneark omhandlende grunddata og antagelser som basis for beregningerne (er ikke vedlagt). Nedenfor gives et overblik over de enkelte elementer i analysen.

- Princip for analysemetode** Når udstyret måler/sorterer forkert, kan det enten være falsk negativ (FN) eller en falsk positiv (FP). Det er valgt at fokusere på konsekvenserne af fejlmåling/-sortering for:
1. Uegnet til konsum pga. produktkvalitet.
Det maksimalt acceptable niveau for falsk negative for hver af de medtagne forandringer beregnes. Det vil sige, hvor ofte må det ske, at en kylling, som er uegnet til konsum pga. produktkvaliteten, når ud til forbrugeren?
 2. Fastlæggelse (estimering) af en kyllingefloks niveau af forskellige forandringer og sikkerheden på denne fastlæggelse. (Prævalenserne for en flok kyllinger).
 3. Slagteriets økonomi.
De økonomiske konsekvenser af givne NFP beregnes. Det vil sige, hvad koster det, når en kylling fejlagtigt kasseres?

- 1. Uegnet til konsum** Analysen er baseret på en lang række antagelser og beslutninger om for eksempel maksimalt acceptabel forekomst af de forskellige forandringer i form af antal falsk negative på et år for ét fremtidigt udstyr. På basis af disse oplysninger beregnes:
- $$\text{Maks_NFN} = \frac{\text{maks_ulempe_forekomst}}{(\text{årlig_produktion} * \text{prævalens} * \text{sandsynlighed})}$$
- hvor *Prævalens* er den gennemsnitlige hyppighed af den givne forandring. *Sandsynlighed* for ulempe er sandsynligheden for at forbrugeren reagerer negativt på den angivne ulempe.
- I tabel 1 nedenfor er anført nogle eksempler på beregnet maksimalt niveau for falsk negative, Maks. NFN, efter implementering af automatisk måling og sortering.

Tabel 1 Eksempler på antagelser og beregnet Maks. NFN

Eksempler på Forandringer	Prævalens %	Ulempe alvor	Sandsynlighed for ulempe	Maks. NFN Frekvens	Antal FN pr. år	Antal FN pr. år, som medfører ulempe
Afmagring	0,1425	3	1	0,00008173	6	6
Misfarvning (inkl. overskoldet)	0,1575	3	1	0,00007395	6	6
Bugvattersot	0,3450	0	0		0	0
Ledbetændelse/ ledforandring	0,1125	1	1	0,00465875	250	250
Hudlidelse/ Hudbetændelse	0,6750	1	1	0,00077646	250	250

2. Forekomst af forandringer i kyllingeflok (prævalens) Udfordringen er her at beregne forekomsten under hensyntagen til udstyrets præcision, det vil sige omfanget af falsk negative, NFN, og falsk positive, NFP. Nedenstående estimater er udarbejdet:

$$\text{Estimat(forekomst)} = KP/N$$

Hvor $KN = ((1 - NFN) * (N - x) - NFN * x) / ((1 - NFP) * (1 - NFN) - NFN * NFP)$
 $KP = (x / (1 - NFN)) - (NFP / (1 - NFN)) * KN$
 x er antal kyllinger målt positiv
 N er antal kyllinger i flokken
 NFP er niveauet af falsk positive
 NFN er niveauet af falsk negative

$$\text{Varians(forekomst)} = (x/N * (1 - x/N + 2 * (pFP - pFN)) - (pFP - pFN) - (pFP - pFN)^2) / N$$

Hvor $pFP = KN * NFP / N$
 $pFN = KP * NFN / N$

I tabel 2 nedenfor er der givet eksempler på estimerede prævalenser samt usikkerheden på bestemmelsen, med de givne antagelser.

Tabel 2 Eksempler på beregnet forekomst (prævalens) af forandringer i en kyllingeflok

Eksempler på Forandringer	Målt forekomst x/N (frekvens)	Niveau af falsk negative NFN	Niveau af falsk positive NFP	Estimeret prævalens	95 % konfidensinterval for estimeret prævalens (+/-)
Afmagring	0,001425	0,000082	0,000059	0,001366	0,000389
Misfarvning (inkl. overskoldet)	0,001575	0,000074	0,000059	0,001516	0,000410
Bugvattersot	0,003450	0,000100	0,000059	0,003392	0,000613
Ledbetændelse/ ledforandring	0,001125	0,004659	0,000059	0,001071	0,000345
Hudlidelse/ Hudbetændelse	0,006750	0,000776	0,000059	0,006697	0,000860

3. Slagteriets økonomi Slagteriets økonomi kan påvirkes af både NFN og NFP. Den økonomiske konsekvens af falsk negative er imidlertid meget vanskelig at kvantificere, det vil sige den økonomiske konsekvens for slagteriet, hvis kyllinger kasseres af forbrugere på grund af afvigende kvalitet. Niveauet af falsk positive (NFP) har konsekvens for slagteriet, idet falsk positive er kyllinger, som kasseres selvom de ikke fejler noget. Det årlige tab som følge af et givet NFP beregnes således for hver forandring:

$$\text{Tab} = (\text{\AA}rlig_produktion - \text{\AA}rlig_produktion * \text{pr\AA}evalens) * \text{NFP} * \text{tab_pr_FP}$$

Det samlede tab bliver summen af tabene for alle forandringer:

$$\sum_{f=1}^n \text{Tab}_i = \sum_{f=1}^n ((\text{\AA}p - \text{\AA}p \times \text{pr}_f) \times \text{NFP}_f \times t)$$

Hvor f er forandringerne
 \AA p er den \AA rlige produktion af kyllinger
 pr er pr\AA evalensen
 t er tabet pr. FP kylling

Et NFP p\AA 0,00012 for hver af de i alt 85 hoved- og undergrupper af afvigelse medf\AA rer 47-48 tusinde fejlagtigt kassererede kyllinger svarende til et tab p\AA ca. 0,5 mio. kr. pr. \AA r. Det skal ses i forhold til, at der med det nuv\AA rende manuelle system \AA rligt i alt kasseres ca. 1 % af alle kyllingerne – alts\AA 477.000 kyllinger af en \AA rproduktion p\AA 47,7 mio. kyllinger. Det vides ikke hvor stor en del af de kasserede i det nuv\AA rende manuelle system, der er falsk positive.

*F\AA lsomheds-
analyse*

Rapporten gennemg\AA r en lang r\AA kke scenarier der viser effekten af for eksempel af at variere niveauerne for falsk negative hhv. positive.

Pr\AA cision/kvalitet af referencedata

Vurdering af organs\AA t

Den veterin\AA re kontrol af organs\AA t best\AA r af kontrol af lever og hjerte, og omfatter fem bed\AA mmelser af leverdefekter: "Liver Necrotic Hepatitis", "Liver Perihepatitis", "Liver Cobble", "Liver White" og "Liver Change". Hjertet bed\AA mmes med hensyn til "Heart Inflammation". 3984 billede-s\AA t er bed\AA mt med hensyn til de seks klasser. I visse tilf\AA lde har det ikke v\AA ret muligt at bed\AA mme klassen og i stedet er \AA n af f\AA lgende klasser anvendt: "Manuel inspektion n\AA dvendig", "Anden sygdom" og hvor billede-s\AA ttet har v\AA ret fejlbeh\AA fttet, er klasserne "Ukendt" og "Nonsens" anvendt. Hver sygdom er inddelt i 4 sv\AA rhedsgrader: No, Low, Medium og High Occurrence.

Form\AA l og metode

Tre dyrl\AA ger har bed\AA mt de samme billeds\AA t med det form\AA l at fastl\AA gge en f\AA lles bed\AA mmelse, som kan anvendes til kalibrering af m\AA leudstyret. Kvaliteten eller pr\AA cisionen af bed\AA mmelsen er defineret ved at bed\AA mmerne er i stand til at holde det samme niveau over tid, det vil sige de samme t\AA rskelv\AA rdier for sv\AA rhedsgraderne: No, Low, Medium og High Occurrence, samt at de er enige om b\AA de tilstedev\AA relse af en afvigelse og graden af en afvigelse.

Variation over tid Det kan ikke med sikkerhed konkluderes, at bedømmelserne varierer over tid. For de hyppigst forekommende afvigelse, *Liver Necrotic Hepatitis*, *Liver Perihepatitis* og *Liver Cobble*, er der en tendens til afvigende bedømmelse af den sidste serie i tidsrækken. Dette kan være en kombination af lavere forekomst af afvigelse og dermed en mere kritisk granskning af data. Endvidere tyder det på, at der er en sammenblanding af *Liver Necrotic Hepatitis* og *Liver Perihepatitis*. En detaljeret undersøgelse viser, at i de fleste tilfælde, hvor der kun er én dyrlæge, der har registreret en afvigelse, da er registreringen angivet som "Low Occurrence".

Grad af enighed Nedenstående tabel viser den gennemsnitlige vurdering og spredning for hver klasse. I få tilfælde bedømme "High Occurrence" af alle dyrlæger. Kolonnen "(DL 1+2+3)/3" angiver relationstallet ("korrelationskoefficienten") mellem de enkelte dyrlæger og gennemsnitsvurderingen. For *Liver Necrotic Hepatitis*, *Liver Perihepatitis* og *Liver Cobble* er relationstallene høje (omkring 0.9) og tyder på en god overensstemmelse mellem dyrlægerne. For de øvrige lever-afvigelse er overensstemmelsen ringe eller ikke eksisterende (*Liver Change*).

Simple Statistics Variable	N= 1907		Relationstal		
	Mean	Std Dev	(DL 1+2+3)/3	Dyrlæge 2	Dyrlæge 3
LiverNecroticHepatitis	0.30	0.75			
Dyrlæge 1	0.31	0.85	0.96	0.82	0.89
Dyrlæge 2	0.26	0.66	0.92		0.81
Dyrlæge 3	0.32	0.85	0.96		
LiverPerihepatitis	0.26	0.56			
Dyrlæge 1	0.34	0.74	0.94	0.76	0.78
Dyrlæge 2	0.20	0.45	0.87		0.70
Dyrlæge 3	0.23	0.64	0.91		
LiverCobble	0.50	0.81			
Dyrlæge 1	0.46	0.94	0.91	0.70	0.77
Dyrlæge 2	0.47	0.69	0.87		0.74
Dyrlæge 3	0.56	1.03	0.93		
LiverWhite	0.15	0.29			
Dyrlæge 1	0.43	0.78	0.96	0.11	0.23
Dyrlæge 2	0.00	0.08	0.26		0.31
Dyrlæge 3	0.02	0.22	0.48		
LiverChange	0.01	0.09			
Dyrlæge 1	0.02	0.21	0.74	0.00	-0.01
Dyrlæge 2	0.00	0.05	0.18		0.00
Dyrlæge 3	0.02	0.18	0.64		
HeartInflammation	0.18	0.38			
Dyrlæge 1	0.43	0.80	0.93	0.52	0.43
Dyrlæge 2	0.06	0.27	0.73		0.52
Dyrlæge 3	0.05	0.28	0.67		

<i>Konklusion</i>	<p>Gennemsnitsbedømmelsen af alle bedømmelser pr. billede foreslås som referencemåling.</p> <p>Dyrlægerne bedømmer ikke på samme niveauer, hvorfor den gennemsnitlige/samlede bedømmelse dækker et interval, der er bredere end normal praksis på slagterierne og derfor egnet som referencemål.</p> <p>Kalibrering rettet mod specifikke afvigelser kan kun udføres for <i>Liver Necrotic Hepatitis</i>, <i>Liver Perihepatitis</i> og <i>Liver Cobble</i>, og eventuelt <i>Heart Inflammation</i> baseret på gennemsnit af dyrlægenes bedømmelser.</p> <p>Referencematerialet er generelt kendetegnet ved få eksempler med høj grad af afvigelse.</p> <p>En højere grad af konsensus med hensyn til vurdering af afvigelser og niveau anbefales.</p> <p><i>Vurdering af udvendige afvigelser</i></p> <p>De vurderede billedeksempler udgør ikke en tilfældig stikprøve, men er udvalgt til at repræsentere de kyllinger, som veterinærkontrollen har nedtaget baseret på mistanke om afvigende kvalitet suppleret med eksempler, som kontrollen har godkendt.</p> <p>Tre bedømmere (dyrlægestuderende) har bedømt 2.257 sæt billeder, billeder, som de alle er enige om har tilstrækkelig kvalitet, så en bedømmelse kan foretages. To af bedømmerne har tillige bedømt yderligere ca. 2000 billedsæt.</p>
<i>Formål og metode</i>	<p>Erfaringerne fra bedømmelse af organsæt resulterede i afholdelse af to workshops, hvor bedømmerne blev trænet dels sammen med en erfarne dyrlæge dels hvor de tre bedømmere gennemførte sammenlignende bedømmelser og diskuterede tærskelværdier. Formålet er derfor at analysere, om træning sikrer tilstrækkelig ensartede bedømmelser.</p>
<i>Overensstemmelse med slagteriets bedømmelser</i>	<p>De nedtagne kyllinger på slagteriet har ikke nødvendigvis afvigende kvalitet. To af bedømmerne fandt således ingen kvalitetsafvigelse i 3 hhv. 10% af disse, medens den tredje bedømmer godkendte næsten 40%. Omvendt var alle tre bedømmere enige om at omkring halvdelen af de kyllinger, som var godkendte på slagteriet, havde kvalitetsafvigelser, men på lavt niveau.</p>
<i>Overensstemmelse bedømmerne imellem</i>	<p>Overordnet er der en acceptabel overensstemmelse mellem bedømmerne, men de har forskellige tærskelværdier, det vil sige at graden af en afvigelse vurderes forskelligt. Det er især afvigelser som <i>Afmagring</i>, <i>Misfarvning</i> og <i>rifter</i>, hvor den ene bedømmer har en højere tærskelværdi, end de to andre bedømmere.</p>

Hudbetændelse og til dels også *bugvattersot* er tilsyneladende vanskelige at bedømme og det anbefales, at der etableres et større sæt træningsdata for disse effekter.

Det vurderes alt andet lige, at en gennemsnitsvurdering kan anvendes som referenceværdi for kalibrering af visionudstyret. Den opnåede præcision ved minimum to bedømmere vurderes at være på et realistisk opnåeligt niveau.

Referencer

- 1 Kravspecifikation udvendig og indvendig kontrol - udkast 6-2
- 2 Rapport Risikoanalyse kødkontrol kyllinger Version2 Marts 2017
- 3 Vurdering af referencedata – organsæt
- 4 Vurdering af referencedata – udvendig kontrol

Tilskud fra Fjerkræafgiftsfonden

År	Projektets samlede tilskudsgrundlag regnskab og budget 1.000 kr.	Tilskud fra Fonden regnskab/bevilget/ansøgt/forventet ansøgt 1.000 kr.	Tilskuddets andel af det samlede tilskudsgrundlag %
2015	498	498	100
2016	464	464	100
2017	419	419	100
2018	322	322	100
I alt	1.703	1.703	100

Deltagere

Kødkontrollen: Ole G. Nielsen, Ole Wenicke

Dyrlægestuderende: Emil Lund, Josefine Johnsson, Katrine Birk Kierulff

DMRI: Chris Claudi-Magnussen, Helle Daugaard Larsen, Eli Vibeke Olsen