

Projektleder KU/SSI	Tariq Halasa
Projektgruppe	Anette Boklund, Matt Denwood, Søren Saxmose Nielsen, Carsten Thure Kirkeby
Fagfællebedømmer	
Kontaktperson i FVST	Sten Mortensen, Kim Sigsgaard

Dato for henvendelse	Dato for svarfrist	Dato for afsendelse	Versionsnummer
27.oktober 2020	05-11-2020	04-11-2020	1

Journalnummer/sagsnummer	FVST	KU	SSI
	2020-14-81-04082	061-0157/20-3680	20/11673

Besvarelse vedr.

- Analyse af sammenhængen mellem afstand og tid ved spredning af COVID-19 blandt minkfarme

Resumé

- Der er analyseret en sammenhæng mellem afstand og tid for SARS-CoV-2 i minkfarme smittet i perioden 1. juni til 3. november 2020. Analysen viser at der er afstand til nærmeste smittede farm er en signifikant risikofaktor for påvisning af SARS-CoV-2. Årsagerne til dette kan ikke fastlægges med sikkerhed. Betydningen af den observerede effekt af afstand varierer med beregningsmetoden.

Baggrund, relevans og perspektiv

- Der ønskes en analyse af sammenhængen i afstand og tid ved spredning af COVID-19 blandt minkfarme. Om muligt bør smittespredning, der kan forklares epidemiologisk med en bestemt smittevej, udelukkes.

I bestillingen spørges der dels til simulering af spredning og kontrolstrategier, dels til analyser af de fylogenetiske clustre og dels af analyser af lokalområder.

Det er ikke muligt at lave en spredningsmodel, da det kræver en del mere tid at lave selv en simpel pathway model med mulighed for at simulere overvågning og kontrolforanstaltninger. Dette skyldes både, at der ikke i forvejen findes en dansk mink-spredningsmodel, og at datagrundlaget er sparsomt.

Det er ikke muligt at fremstille analyser på grundlag af grupper omfattende meget få farme. Dvs. de første farme, Læsø og cluster 5 (indeholder indtil nu kun 5 farme) kan ikke bruges til meningsfulde analyser. Spredningen er så stor, at modeller ikke vil kunne fittes og måske heller ikke konvergere. Den deskriptive statistik fortæller det hele.

En stor mangel med cluster-data er særligt, at der mangler informationer på ca. 2/3 af de smittede besætninger jævnt fordelt over hele perioden. En mindre mangel er, at vi ikke ved, om der er flere stammer på flere farme. Dvs. at vi ikke kan lave meningsfuld statistik, særligt fordi manglerne ikke er tilfældigt fordelte. For farm 1-27 er der sekventeret flere minkprøver pr farm for 23-25 farme. I 5 af disse blev der påvist to forskellige clustre. Efterfølgende er der i langt de fleste farme sekventeret virus fra en mink, dvs. at det ikke er muligt at påvise evt. tilstedeværelse af flere minkvarianter i samme farm.

Konsekvensen af denne bias i data vil være, at parameterestimaterne og usikkerhed kan være forkerte, og retningen af denne fejl vil være ukendt.

Det vil introducere en væsentlig selektionsbias (primært spektrumbias), hvis visse observationer fjernes fra datasættet, eksempelvis introduktioner over større afstande, der kan antages at skyldes humane introduktioner. Dette fordi det er sandsynligt, at en meget stor andel af den øvrige smittespredning er human.

I nedenstående analyser er derfor fokuseret på effekten af afstand til nærmeste smittede farm i forhold til tid. Formålet med analysen er at undersøge, hvordan effekten af afstand til nærmest smittet farm påvirkes af antagelser omkring farmenes infektiøse periode.

Metode, data m.m.

▸ For hvert af de fem clustre påvist ved fylogenetisk analyse blev der foretaget en deskriptiv analyse af afstand til nærmeste smittede minkfarm.

Der bliver brugt overlevelsesanalyse med PCR påvisning af SARS-CoV-2 som "Event" og mistankedato som "time of event". De forklarende variable i analysen er: besætningsstørrelse og afstand til nærmeste smittede farm. Disse variable har tidligere vist høj statistisk signifikans. Start dato for beregning af risikoperioden er sat til den 1. juni 2020, og sidste inkluderede dato er den 03. november 2020. Cox proportional hazard analyser blev kørt i R (R Core Team, 2020) i pakken "coxme" (Therneau, 2020a) for at prædiktere hazard ratio (HR). Forventet tid til påvisning af SARS-CoV-2 i minkfarme (overlevelsestid) blev estimeret i pakken "survival" (Therneau, 2020b). Estimering af den forklarende varians er baseret på pakken "survMisc" (Dardis, 2018).

Vi har kørt 5 forskellige overlevelsesanalyser. I alle scenarier beregnes først afstanden fra en given farm til samtlige smittede farme, inden for en fastlagt periode, og dernæst bruges afstanden til nærmeste smittede farm (minimumsafstand) i overlevelsesanalysen. Afstanden fra en given farm til samtlige smittede farme er beregnet på de 5 forskellige måder:

1. Afstand beregnet i forhold til alle smittede farme uden at tage hensyn til påvisningsdatoen af disse.
2. Afstand beregnet til alle farme, som var smittet tidligere end den givne farm (dvs. individuel dato for hver farm i forhold til, hvornår den blev smittet).
3. Afstand beregnet til alle farme, som var smittet op til 30 dage før den givne smittede farm.
4. Afstand beregnet til alle farme, som var smittet op til 14 dage før den givne smittede farm.
5. Afstand beregnet til alle farme, som var smittet i en periode ± 14 dage i forhold til den givne farm.

Vi præsenterer prædiktionen af ændring i HR for påvisning af SARS-CoV-2 i forhold til afstand, hvor vi som eksempel har brugt en besætningsstørrelse på 10.000 dyr. Vi har brugt en fast besætningsstørrelse for at fjerne effekten heraf. HR kan betragtes som en relativ risiko i forhold til en gennemsnitsfarm i populationen. Dvs. i dette tilfælde en farm med gennemsnitlig afstand til nærmeste smittede farm. HR > 1 repræsenterer en forhøjet risiko, mens HR < 1 repræsenterer en lavere risiko end en gennemsnitsfarm. Populationen inkluderer alle minkfarme i Nordjylland. Farme i regioner Midt- og Sydjylland er ikke inkluderet. Vi har desuden estimeret overlevelsestid som funktion af afstand til nærmeste smittede farm for at undersøge effekten af beregningsmetode for "afstand" på.

Forskellige transformationer af afstande er undersøgt og i modellen, der benyttes, inkluderes et andengradspolynomium for afstanden til den nærmeste smittede farm, dvs. at den estimerede effekt tillades at være ikke-lineær, da denne transformation har vist sig til at give det bedste model fit. Modellernes forklaringsgrad blev sammen med biologisk logisk ræsonnement brugt til at diskutere de fem scenariers validitet.

Resultater

▸ Tabel 1 viser fordeling af afstand til nærmeste minkfarm, hvor SARS-CoV-2 af samme variant er påvist.

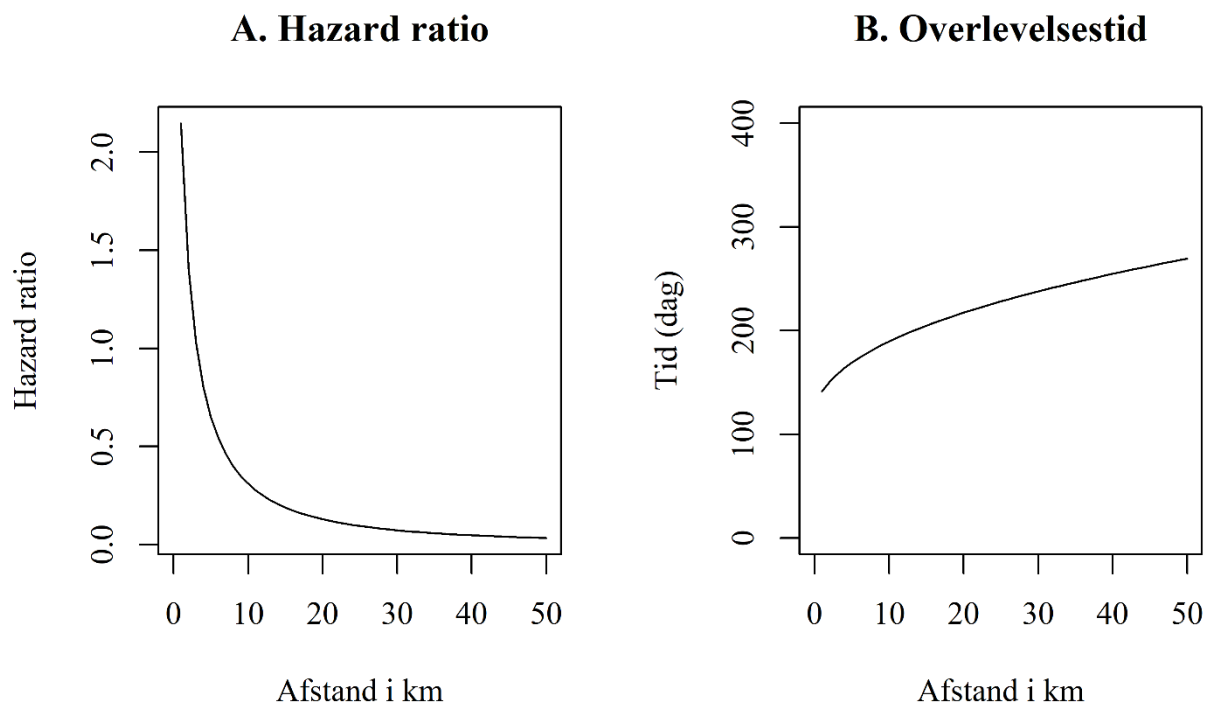
Tabel 1: Fordeling af afstand til nærmeste minkfarm hvor SARS-CoV-2 af samme variant er påvist

Cluster	5%	25%	50%	75%	95%	Antal farme i cluster
1	0.7	0.7	0.7	0.7	0.7	2
2	0.3	0.6	1	4.4	13.4	10
3	0.8	1.2	3.8	4.7	26.1	18
4	0.25	0.77	2.1	3.8	36.4	24
5	0.9	0.9	1	7.7	7.7	5

Resultater af overlevelsesanalyserne

1. Afstand beregnet i forhold til alle smittede farme uden at tage hensyn til påvisningsdatoen

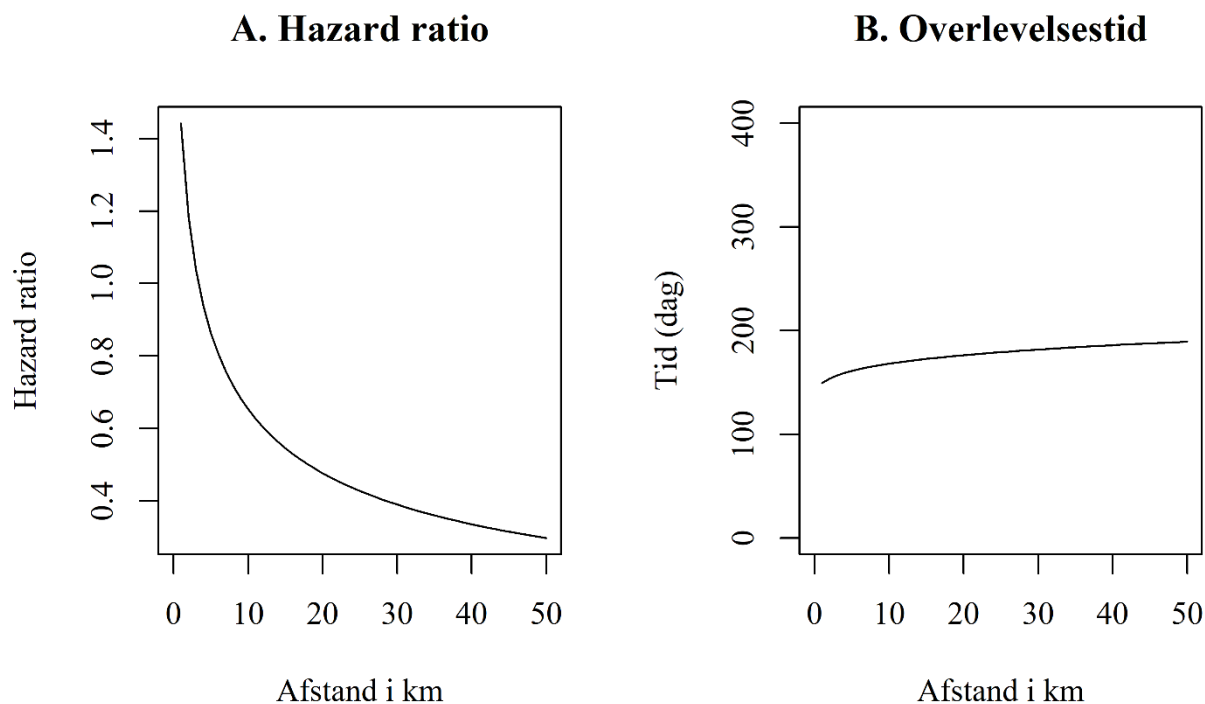
Figur 1A viser sammenhæng mellem afstand til nærmeste smittede farm og HR, hvor afstand er beregnet i forhold til alle smittede farme og uden at tage hensyn til påvisningsdatoen. Figuren viser en skarp reduktion i HR i forhold til afstand, hvor HR falder til under 1 inden for de første 5 km. Medianen (5-95%) af afstand til nærmeste smittede farm er 2.4 (0.09-18.8) km. Figur 1B viser sammenhæng mellem overlevelsestid og afstanden. Den viser en tydelig stigning i overlevelsestid indtil påvisning, når afstand stiger. Modellen forklarer 38% af variationen.



Figur 1: A. Sammenhæng mellem afstand til nærmeste smittede farm og HR. B. Sammenhæng mellem afstand til nærmeste smittede farm og overlevelsestid, når der ikke bliver taget hensyn til påvisningsdatoen (afstand beregnet i forhold til alle smittede farme uden hensyn til påvisningsdatoen).

2. Afstand beregnet til alle farme som har været smittet før en given smittet farm

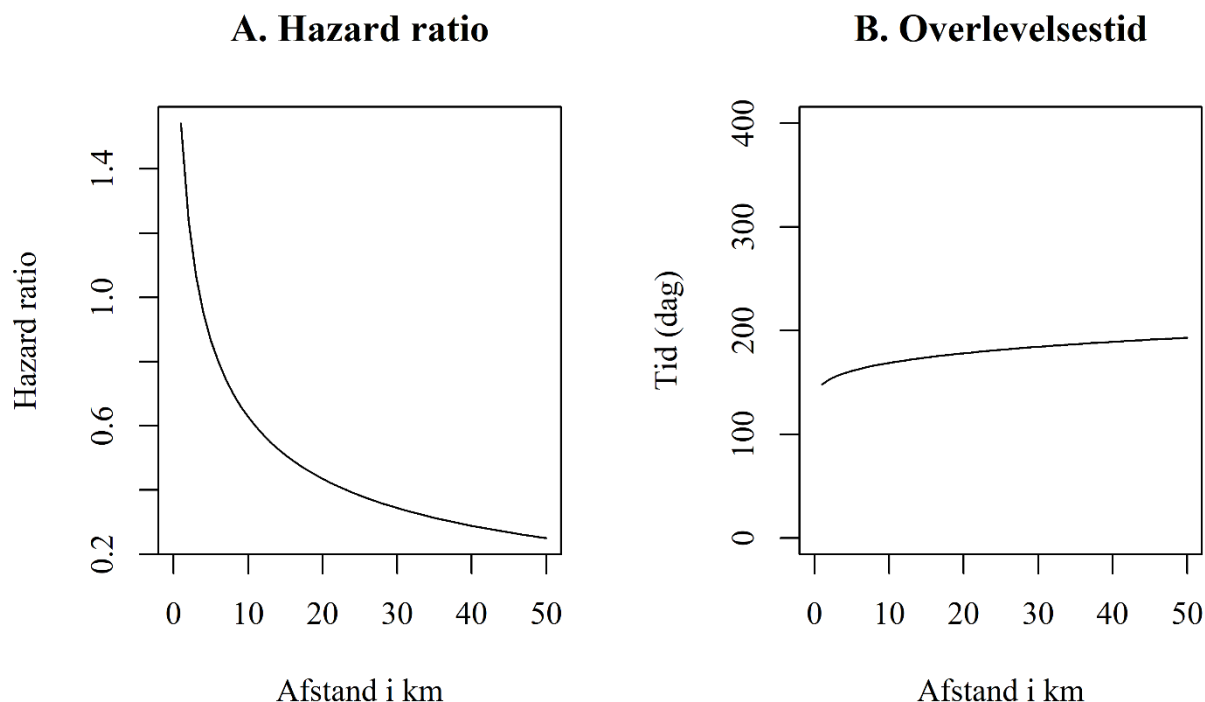
Figur 2A viser forhold mellem afstand til nærmeste smittede farme og HR, hvor afstand er beregnet i forhold til alle farme, som var smittet tidligere end en given farm (dvs. individuel dato for hver farm, i forhold til hvornår den blev smittet). Figuren viser en mindre stejl reduktion i HR i forhold til afstand sammenlignet med figur 1. Denne effekt ses også i sammenhængen mellem overlevelsestid og afstand (fig. 2B), hvor effekten af afstand er mindre end det er observeret i figur 1B. Medianen (5-95%) af afstand til nærmeste smittede farm er 2.8 (0.1-20.7) km. Modellen forklarer 16% af variationen.



Figur 2: A. Sammenhæng mellem afstand til nærmeste smittede farm og HR, B. Sammenhæng mellem afstand til nærmeste smittede farm og overlevelsestid, hvor afstand er beregnet i forhold til alle farme, som var smittet tidligere end en given farm (dvs. individuel dato for hver farm, i forhold til hvornår den blev smittet)

3. Afstand beregnet til alle farme, som var smittet op til 30 dage før en given farm

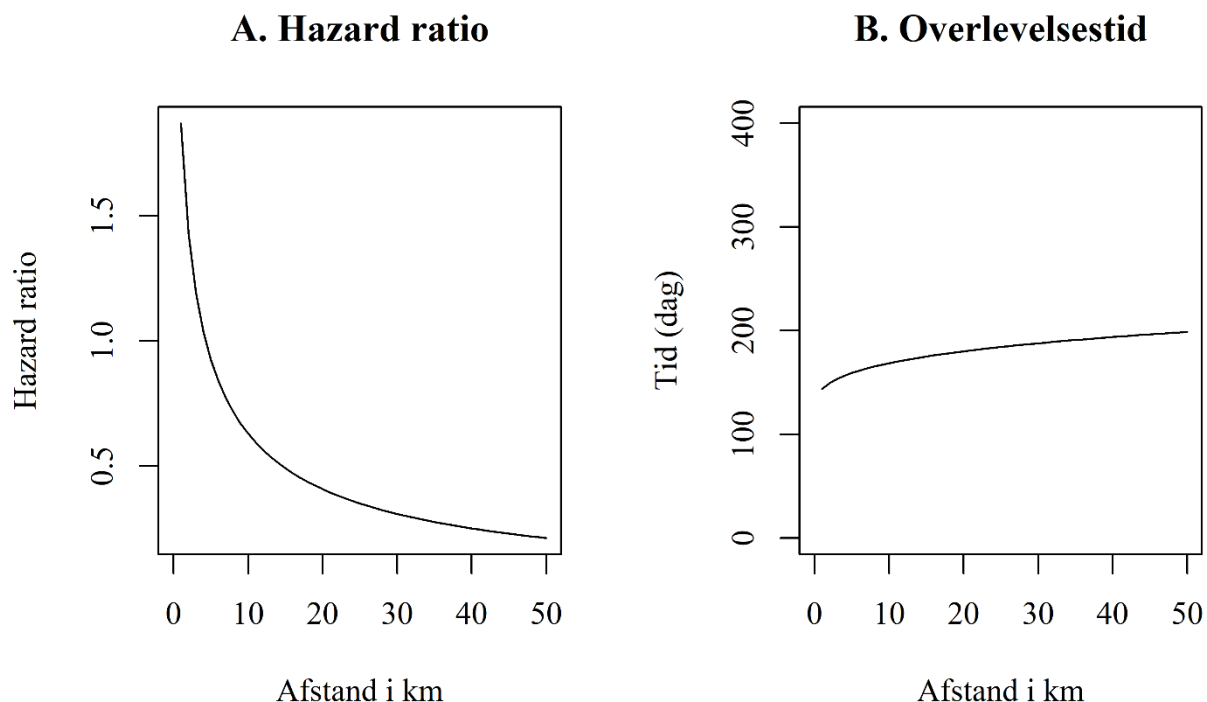
Figur 3A viser sammenhæng mellem afstand til nærmeste smittede farme og HR, hvor afstand er beregnet i forhold til alle farme, der er smittet de sidste 30 dage. Figuren ligner figur 2A, hvilket betyder, at inklusion af alle smittede farme eller inklusion af smittede farme fra de sidste 30 dage, ikke påvirker effekten af den beregnede afstand. Effekten kan ses også i sammenhængen mellem overlevelsestid og afstand (fig. 3B). Medianen (5-95%) af afstand til nærmeste smittede farm er 3.1 (0.11-20.7) km. Modellen forklarer 18% af variationen.



Figur 3: A. Sammenhæng mellem afstand til nærmeste smittede farm og HR, B. Sammenhæng mellem afstand til nærmeste smittede farm og overlevelsestid, hvor afstand er beregnet i forhold til alle farme smittet inden for de sidste 30 dage

4. Afstand beregnet til alle farme, som var smittet op til 14 dage før en given farm

Figur 4A viser sammenhængen mellem afstand til nærmeste smittede farm og HR, hvor afstand er beregnet i forhold til alle smittede farme inden for de sidste 14 dage. Figuren viser en stejlere reduktion i HR i forhold til de foregående scenarier. Sammenhængen mellem overlevelsestid og afstand (fig. 4B) er dog sammenlignelig med figur 2B og 3B. I fig. 4A observerer vi her, at HR falder til under 1 på afstande mellem 5 og 10 km. Medianen (5-95%) af afstand til nærmeste smittede farm er 4.3 (0.14-22) km. Modellen forklarer 26% af variationen.

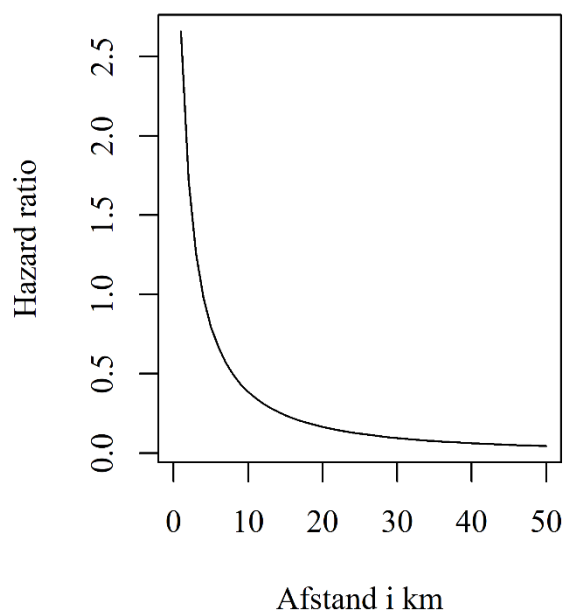


Figur 4: A. Sammenhæng mellem afstand til nærmeste smittede farm og HR, B. Sammenhæng mellem afstand til nærmeste smittede farm og overlevelsestid, hvor afstand er beregnet i forhold til alle farme smittet inden for de sidste 14 dage.

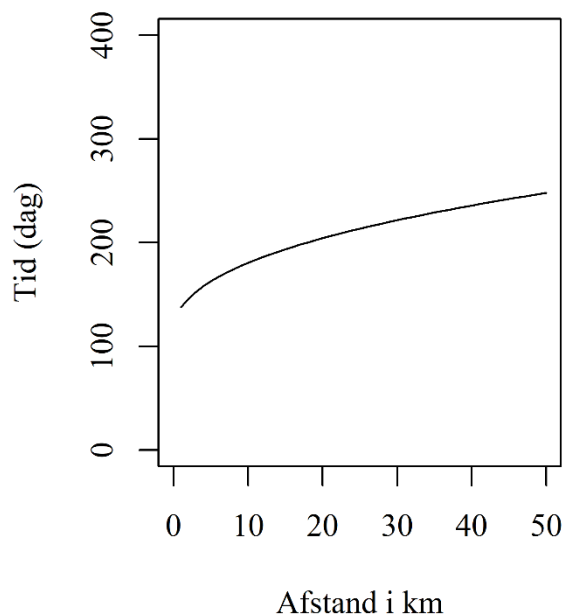
5. Afstand beregnet til alle farme, som var smittet i en periode ± 14 dage fra hver farm

Figur 5A viser sammenhæng mellem afstand til nærmeste smittede farm og HR, hvor afstand er beregnet i forhold til alle smittede farme i en periode ± 14 dage. Der ses et kraftigt fald i HR over de første 5 km. Sammenhængen mellem overlevelsestid og afstanden kan ses i figur 5B. Den viser en tydelig stigning i overlevelsestid med stigning i afstand, og med en stejlere effekt end det der observeret i fig. 2B, 3b, og 4B. Medianen (5-95%) af afstand til nærmeste smittede farm er 3.5 (0.1-20.6) km. Modellen forklarer 46% af variationen.

A. Hazard ratio



B. Overlevelsestid



Figur 5: A. Sammenhæng mellem afstand til nærmeste smittede farm og HR, B. Sammenhæng mellem afstand til nærmeste smittede farm og overlevelsestid, hvor afstand er beregnet i forhold til alle smittede farme i en periode ± 14 dage.

Diskussion

► Fælles for alle analyserne var, at afstand til nærmeste smittede farm var en signifikant variabel, selv om dens effekt varierede. Man kan observere, at effekten af afstand på HR falder hurtigt inden for de første 5 km, men også at det punkt hvor HR falder under 1 (dvs. risikoen for påvisning af SARS-CoV-2 relativt til den gennemsnitlige afstand fra en given farm til nærmeste smittede farm er under 1) ændrer sig, afhængigt af hvilke farme, der inkluderes i afstandsberegningen. Dette kan naturligvis påvirke sammenhængen mellem overlevelsestid og afstand.

De deskriptive analyser inden for hver cluster viser samtidigt at 75% af smittede farme har en anden minkfarm smittet med samme cluster af SARS-CoV-2 inden for en afstand af ca. 4-5 km, dog med undtagelse af cluster 1 og 5, hvor der kun er få observationer. Generelt er den deskriptive analyse af de fem clustre baseret på et mindre antal minkfarme, da der ikke er lavet fylogenetisk analyse af prøver fra alle farme. Desuden er kun få prøver, ofte kun én, anvendt fra hver farm. Vi har valgt at fortolke på baggrund af 75-percentilerne, da disse er mere stabile end 95-percentilerne, mens 95-percentilerne sandsynligvis kan være forårsaget af humane introduktioner over store afstande.

Sammenhæng mellem afstand til nærmeste smittede farm og HR for påvisning af SARS-CoV-2 kan ændre sig væsentligt, afhængigt af hvilke smittede farme, der inkluderes i beregningen, dvs. af antagelserne omkring den periode omkring påvisningstidspunktet, hvori minkfarme antages at være infektiøse.

Introduktion af smitte i de enkelte farme kan ikke umiddelbart antages at være sket i samme rækkefølge som påvisningen af SARS-CoV-2 i farmene, men vil afhænge af såvel minkavlerens og dyrlægens årvågenhed, som af FVST's prioritering af prøveudtagning. Dette udfordrer vores mulighed for at bestemme sammenhængen mellem afstand og tid. Det gør det også svært at bestemme, hvilken beregningsmåde der er den rigtige måde. Tidligere undersøgelser har estimeret, at introduktion af SARS-CoV-2 ofte sker 14-18 dage inden påvisning. Vi vurderer, at scenarie 5 er en realistisk metode til beregning af afstand til nærmeste smittede farm. I scenarie 1 beregnes afstand til alle farme, inklusiv dem som er blevet påvist senere end en specifik farm, hvilket er urealistisk. Scenarier 2, 3 og 4 tager hensyn til mistankedato og inkluderer kun smittede farme hvor påvisning er sket tidligere. I scenarie 2 tages hensyn til alle farme, dette er også urealistisk for mange påviste farme sero-konverterer og vurderes dermed ikke fortsat at udgøre en risiko. Derfor er scenarier 3 og 4 mere realistiske end scenarie 2. Ikke desto mindre, tager scenarier 3 og 4 ikke hensyn til usikkerheden omkring rækkefølgen af påvisning af farmerne. Dette er inkluderet i scenarie 5.

En af de basale regler i standard overlevelsesanalyser er, at alle risikofaktorer er kendt ved analysen start, dvs. dag 1 i analysen. Denne regel er ikke overholdt i de beskrevne analyser. Det var ikke mulig at foretage alternative analyser inden for den givne tidsramme, og det vurderes derfor, at vi ud fra analyserne i sammenhæng med de deskriptive resultater kan konkludere, at afstand er en signifikant risikofaktor. Det er ikke muligt ud fra de tilgængelige data og viden at forklare, hvorfor afstand er en risikofaktor.

Referencer:

R Core Team (2020). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.

Terry M. Therneau (2020a). coxme: Mixed Effects Cox Models. R package version 2.2-16. <https://CRAN.R-project.org/package=coxme>

Therneau T (2020b). `_A Package for Survival Analysis in R_`. R package version 3.2-7, <URL: <https://CRAN.R-project.org/package=survival>>.

Chris Dardis (2018). survMisc: Miscellaneous Functions for Survival Data. R package version 0.5.5. <https://CRAN.R-project.org/package=survMisc>