

2021

Prædiktionsalgoritme til borgere med KOL



Aalborg Universitet

**Klinisk Videnskab &
Teknologi**

4. Semester

Gruppe 21gr10504

Rasmus Steengaard Holm
Hannah Møldrup Derby
Mille Andersen

Vejledere:

Thomas Kronborg Larsen
Stine Hangaard

1. juni 2021

Titel: Prædiktionsalgoritme til borgere med KOL

Formål: Formålet er at prædiktere, hvilke borgere med KOL der skal tilbydes telemedicin.

Tema: Telemedicin, Prædiktionsalgoritme & KOL

Projektperiode: 01-02-21 til 01-06-21

Projektgruppe: 21gr10504

Deltagere:

Rasmus Steengaard Holm

Hannah Derby Møldrup

Mille Andersen

Vejleder:

Hovedvejleder: Thomas Kronborg Larsen

Bivejleder: Stine Hangaard

Sidetæl: 42

Opslagstal med mellemrum: 64.460

Antal appendix: 7

Afsluttet den: 31-05-21

Resumé

Baggrund

Kronisk Obstruktiv Lungesygdom (KOL) er forbundet med øgede samfundsøkonomiske omkostninger, hvor hospitalsindlæggelser udgør den største omkostning. Derfor tilbydes telemedicin til alle borgere med KOL som en løsning til at forebygge hospitalsindlæggelser. Dog viser studier, at dette ikke gavner samfundsøkonomien. Flere studier foreslår anvendelsen af prædiktionsalgoritmer som en mulighed for at identificere hvilke borgere, der kan være med til at nedbringe samfundsøkonomiske omkostninger ved anvendelse af telemedicin.

Metode

Der udarbejdes en overordnet gruppe og syv stratificerede grupper, hvor der udvikles prædiktionsalgoritmer baseret på både *Multipel Lineær Regression* (MLR) og *Random Forest Regression* (RFR). Prædiktionsalgoritmerne skal prædiktere hvilke borgere, der er i risiko for flest indlæggelsedage, da omkostningerne forbundet med en hospitalsindlæggelse afhænger af indlæggelsestiden. Hver gruppes datasæt præprocesseres og inddeles i et trænings- og testsæt. Prædiktionsalgoritmerne valideres ved en korrelationsanalyse, hvor der beregnes en determinationskoefficient. Prædiktionsevnen for de bedste prædiktionsalgoritmer illustreres i scatterplots.

Resultater

Prædiktionsevnen for prædiktionsalgoritmerne baseret på både MLR og RFR er ikke acceptabel ift. at prædiktere, hvilke borgere der er i risiko for flest indlæggelsedage. Den højeste determinationskoefficient ved MLR er 0,228, hvor den ved RFR er 0,288.

Konklusion

Resultaterne i specialet var ikke acceptabel, hverken i den overordnede gruppe eller de stratificerede grupper, hvorfor de udviklede prædiktionsalgoritmer ikke kan anvendes til at prædiktere, hvilke borgere der er i risiko for flest indlæggelsedage.

Title: Prediction algorithms for citizens with COPD

Aim: The purpose of this study is to predict which citizens with COPD will be offered telemedicine.

Theme: Telemedicine, Prediction algorithm & COPD

Project period: 01-02-21 til 01-06-21

Project group: 21gr10504

Participants:

Rasmus Steengaard Holm

Hannah Derby Møldrup

Mille Andersen

Supervisor:

Thomas Kronborg Larsen

Stine Hangaard

Number of pages: 43

Number of characters including: 64.460

Number of appendixes: 7

Date of completion: 31-05-21

Summary

Background

Chronic Obstructive Pulmonary Disease (COPD) is associated with increased socioeconomic costs, with hospitalizations being the largest. Therefore, telemedicine is offered to all citizens with COPD as a solution to prevent hospitalizations. However, studies show that this does not benefit the economy. Several studies suggest the use of prediction algorithms as an opportunity to identify which citizens can help reduce the socio-economic costs by using telemedicine.

Method

An overall group and 10 stratified groups are prepared, where prediction algorithms are developed based on both Multiple Linear Regression (MLR) and Random Forest Regression (RFR). The algorithms must predict which citizens are at risk for the most bed days, as the costs associated with a hospitalization depend on the amount of bed days. Each group's data set is preprocessed and divided into a training and test set. The prediction algorithms are validated by a correlation analysis, where a coefficient of determination is calculated. The predictive power of the best prediction algorithms is illustrated in scatterplots.

Results

The predictability of the prediction algorithms based on both MLR and RFR is not acceptable in relation to predicting which citizens are at risk for the most bed days. The highest coefficient of determination at MLR is 0.228, where at RFR it is 0.288.

Conclusion

The results in the thesis were not acceptable, neither in the overall group nor the stratified groups, so the developed prediction algorithms cannot be used to predict which citizens are at risk for the most hospitalization days.

Forord

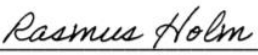
Dette projekt er et kandidatspeciale udarbejdet i foråret 2021 af gruppe 21gr10504 på 4. semester på kandidatuddannelsen i Klinisk Videnskab og Teknologi på Aalborg Universitet. Specialegruppen består af to sygeplejersker og én fysioterapeut.

Forud for projektet blev der lavet en struktureret litteratursøgning for at afdække eksisterende litteratur på området. Denne er yderligere beskrevet i Appendix 1, 2 og 3.

Specialets formål er at udvikle en prædiktionsalgoritme, som kan prædiktere hvilke borgere med KOL, der er forbundet med flest indlæggelsesdage, hvorved telemedicin til disse borgere vil kunne gavne samfundsøkonomien.

Tak til TeleCare Nord for at stille data til rådighed og til Flemming Witt-Udsen for at være behjælpelig med forklaring af datasættets variabler.

Tak til vores hovedvejleder, Thomas Kronborg Larsen, og vores bivejleder, Stine Hangaard, for konstruktiv vejledning og et godt samarbejde.


Rasmus Steengaard Holm


Hannah Møldrup Derby


Mille Andersen

Læsevejledning

I specialet anvendes *Vancouver* som referencesystem, hvorved referencer i teksten angives med tal i parentes. Såfremt en forfatter nævnes i en sætning, placeres referencen efter forfatterens navn.

Nævnes forfatterens navn ikke, placeres referencen til sidst i en sætning eller efter et afsnit.

Referencer, som refererer til foregående sætning angives *før* punktum, mens referencer, som refererer til hele det foregående afsnit angives *efter* punktum.

Specialet består af 41 sider delt ind i 6 hovedafsnit med underafsnit og 7 appendix i særskilt dokument. Hovedafsnit indikeres med et nummer i kronologisk rækkefølge, hvor underafsnit indikeres med to eller tre numre. Det første nummer refererer til hovedafsnittet og det andet nummer refererer til rækkefølgen, afsnittene kommer i (eksempelvis *1.4 Telemedicin* og *4.2.1 Prædiktionsalgoritmer*)

Tal mellem et og ti angives med bogstaver, mens de resterende tal angives med tal. Procenter, år, appendix, figurer, tabeller og beløb angives alle med tal.

I specialet anvendes forkortelser, som skrives ud første gang, de nævnes efterfulgt af forkortelsen i parentes. Herefter skrives kun forkortelserne.

Variabler fra datasættet, som nævnes i teksten, angives i fraser ("...") og med stort forbogstav. Det samme gør de grupper, som stratificeres ud fra udvalgte variabler og overskrifter, som skrives i teksten. Engelske termer skrives i kursiv. Også meta-, tabel- og figurtekster skrives i kursiv.

Tabelnumre består af to tal adskilt af et punktum, hvor det første refererer til det hovedafsnit, tabellen forefindes under, og det andet tal refererer til det nummer i rækkefølgen, tabellen er i under hovedafsnittet (eksempelvis *Tabel 3.1*).

Indledning

I 2020 skulle alle danske borgere med KOL tilbydes telemedicin (1), selvom det ikke er alle borgere, der opnår en effekt ved telemedicin. Et studie af Lilholt et al. (2) viser, at der ikke opnås en statistisk signifikant forskel i sundhedsrelateret livskvalitet, hos de borgere som anvender telemedicin, sammenlignet med almindelig behandling. Et andet studie af Witt-Udsen et al. (2) viser, at der heller ikke ses forbedringer af de samfundsmæssige omkostninger, der er forbundet med KOL, hvis telemedicin gives til alle borgere med KOL. Derfor har flere studier forsøgt at prædiktere hvilke borgere med KOL, der er i risiko for exacerbationer, for at finde hvilke borgere der vil få størst gavn af telemedicin. Her viser et systematisk review af Sanchez Morillo et al. (3), at prædiktionsalgoritmerne ikke opnår tilstrækkelige resultater til at kunne blive anvendt i klinisk praksis. Flere studier giver udtryk for, at problematikken med prædiktionsalgoritmer er, at borgere med KOL er en kompleks befolkningsgruppe, som derfor optræder heterogen. (4–6) Eksempelvis oplever borgere med KOL i grad 3 og 4 flere exacerbationer end borgere med KOL i grad 1 og 2, hvorfor disse borgere er i større risiko for exacerbationer. Derudover er der andre faktorer, som kan være med til at forværre KOL, bl.a. komorbiditeter, rygning og alder, som gør at udvalgte borgere kan få mere gavn af telemedicin.

Vi stiller os derfor undrende over for anbefalingerne omkring at tilbyde telemedicin til alle borgere med KOL, når der ikke ses en øget livskvalitet eller samfundsøkonomisk effekt af telemedicin hos alle borgere med KOL (2,7,8). Dette kan både være en ulempe for borgerne med KOL, som skal bruge deres tid og ressourcer på en behandling, som ikke har den ønskede effekt, men også sundhedspersonalet der skal forholde sig til de indsamlede oplysninger. Der ses derfor en mulighed i at målrette indsatsen til, hvilke borgere med KOL der skal tilbydes telemedicin.

Formålet med dette speciale er derfor at undersøge, om en prædiktionsalgoritme kan prædiktere, hvilke borgere med KOL der bør tilbydes telemedicin

Indholdsfortegnelse

1. Problemanalyse	1
1.1 Epidemiologisk redegørelse af KOL	1
1.2 Diagnosticering af Kronisk Obstruktiv Lungesygdom	2
1.2.1 Behandlingsmuligheder for KOL	2
1.3 Samfundsomkostninger forbundet med KOL	4
1.4 Det respiratoriske system hos borgere med KOL	4
1.5 Borgerens livssituation påvirkes af KOL	6
1.6 Konsekvenser af exacerbationer	7
1.7 Telemedicin	8
1.8 Prædiktionsalgoritmer	10
1.9 Problemafgrensning	11
1.10 Problemformulering	12
2. Metode	13
2.1 Datasæt fra TeleCare Nord	14
2.2 Præprocessering af datasættet	14
2.3 Stratificering af datasættet	15
2.4 Træning af model	16
2.4.1 Multipel Lineær Regression	16
2.4.2 Random Forrest Regression	16
2.5 Test af modeller	18
2.6 Ethiske overvejelser	18
3. Resultater	19
3.1 Præprocessering af datasættet	19
3.2 Stratificering af datasættet	21
3.3 Træning af modeller	21
3.4 Test af modeller	24
3.5 Scatterplots	26
4. Diskussion	29
4.1 Prædiktionsalgoritmer	29
4.2 Stratificering og præprocessering	31
4.3 Telemedicin	32
5. Konklusion	34
6. Perspektivering	35
7. Referenceliste - Vancouver	36

1. Problemanalyse

I problemanalysen belyses problematikken omkring at give telemedicin til alle borgere med KOL. Afsnittet påbegyndes med en epidemiologisk redegørelse af KOL efterfulgt af diagnosticering af KOL, samt hvilken behandling der tilbydes. Herefter redegøres for de samfundsmæssige omkostninger. Det respiratoriske system og hvordan dette forringes ved KOL vil blive belyst, efterfulgt af et afsnit omhandlende borgerens liv med KOL. Der redegøres for konsekvenserne ved exacerbationer og hvordan telemedicin kan være med til at forebygge dette. Til sidst forklares, hvordan prædiktionsalgoritmer ses som en mulig løsning til at prædiktere hvilke borgere med KOL, der skal tilbydes telemedicin.

1.1 Epidemiologisk redegørelse af KOL

Kronisk Obstruktiv Lungesygdom (KOL) er en kronisk lungesygdom, som ifølge *World Health Organisation* ramte 251 millioner mennesker på verdensplan i 2016, og det estimeres, at ca. 5% af alle dødsfald skyldes KOL (9). Prævalensen af KOL forventes fortsat at stige således, at sygdommen i 2030 vil være årsag til 7,8% af alle dødstilfælde på verdensplan, hvilket gør KOL til den 3. hyppigste dødsårsag (10,11). Ifølge Chen et al. (12) forventes borgere med KOL at leve 6-10 år kortere end den generelle befolkning, og ved en akut forværring af KOL, er den forventede levetid herefter mellem 3,4 år og 9,3 år alt efter sværhedsgraden af sygdommen (13).

Udover at være en international udfordring, er KOL også en sygdom, der har store konsekvenser i Danmark, idet prævalensen her er blandt den højeste i Vesteuropa. En dansk kvinde har 3,3 gange større risiko for at dø af KOL, og en dansk mand har 1,7 gange højere risiko for at dø af KOL, sammenlignet med resten af Europa. Det estimeres, at der i Danmark er 430.000 personer med nedsat lungefunktion, der kan klassificeres som KOL, og at der årligt dør omkring 3500 personer af sygdommen. (11) Der ses en sammenhæng mellem prævalensen af KOL og alderen på borgerne, hvor den største prævalens ses hos borgere i aldersgruppen 75-84 år, mens den mindste prævalens ses hos borgere i aldersgruppen 45-65 år. (14) Det kan formodes, at prævalensen af borgere med KOL vil stige i fremtiden, eftersom Danmarks Statistik forventer en stigning i antallet af ældre (15). Det forventes i år 2030, at 22,4% af befolkningen er over 65 år sammenlignet med 19,6% i 2019 (15). Yderligere viser befolkningsfremskrivningen fra 2020, at befolkningen på 80 år og ældre, udgør mere end 7% af befolkningen i 2030, hvor den i 2020 udgjorde 4,7% (16).

1.2 Diagnosticering af Kronisk Obstruktiv Lungesygdom

Til diagnosticering af KOL har organisationen *Global Initiative for Chronic Obstructive Lung Disease* (GOLD) udviklet et værktøj, som benyttes til at klassificere borgere med KOL for bedre at kunne tilpasse behandlingen til den enkelte borger. Med klassifikationsværktøjet inddeles KOL i graderne 1-4 baseret på spiometri, hvor der angives en værdi for *Forced Expired Volume* (FEV1), og borgerens symptomer inddeles i A-D, som angives ved undersøgelse af tidligere exacerbationer, indlæggelser og generelle tilstande. Således vil 1A være forbundet med bedste FEV1 og færrest symptomer, og 4D vil være forbundet med værste FEV1 og flest symptomer. (17,18) Se nedenstående figur 1.1.

GOLD Grad	FEV1	GOLD A-D		
GOLD 1	≥ 80	≥ 2 eller ≥ 1 symptomer for moderat eller svær exacerbation som fører til hospitalsindlæggelse	C	D
GOLD 2	50-79	0 eller 1 symptomer for moderat eller svær exacerbation som <u>ikke</u> fører til hospitalsindlæggelse	A	B
GOLD 3	30-49		*mMRC 0-1 **CAT < 10	mMRC ≥ 2 CAT ≥ 10
GOLD 4	< 30			

*mMRC en forkortelse for Dyspnø skema
** CAT en forkortelse for et vurderingsskema

Figur 1.1: Oversigt over klassifikationen af KOL, med inspiration fra GOLD (19).

1.2.1 Behandlingsmuligheder for KOL

Det primære mål ved behandlingen af KOL er at lindre symptomerne, forebygge exacerbationer, behandle komorbiditeter og udføre lungerehabilitering for at vedligeholde funktionsniveauet (20). Behandlingen af KOL kan inddeles i en stabil og ustabil fase. Den stabile fase indeholder non-farmakologisk og farmakologisk behandling, hvor den ustabile fase indebærer behandling på et hospital som følge af akutte exacerbationer (21).

Den non-farmakologiske behandling af KOL berører rygning, motion, kost og vaccination. Den vigtigste del af behandlingen er rygestop, da dette er hovedårsagen til KOL (21). Et rygestop vil mindske progression af sygdommen markant, idet den accelererede nedbrydning af lungevævet forårsaget af rygningen bremses (7). Et studie af Decramer et al. (7) viser, at dødeligheden kan reduceres med op til 18%, hvis borgeren med KOL stopper med at ryge. En forbedring af konditionen gennem motion kan optimere musklernes iltoptagelse fra blodet, hvorfor kroppen bliver bedre til at udnytte den tilgængelige iltmængde. Dette er relevant, da borgerens nedsatte lungefunktion minimerer blodets iltindhold. Kosten kan også forebygge forværringen af KOL, da undervægt

nedsætter immunforsvaret og heraf evnen til at modstå infektioner, som kan forværre progressionen af KOL. Et studie viser bl.a., at mortaliteten er lavere hos borgere med KOL, der har en BMI over 25 og at der ses en større mortalitet ved undervægt. (21) Desuden anbefales det, at borgeren undgår eksponering af miljømæssige og arbejdsrelaterede risikofaktorer og får en årlig influenzavaccine (7).

Den farmakologiske behandling til borgere med KOL tilrettelægges altid til den enkelte borger, da det er individuelt, hvilken behandling der er mest effektiv. I den stabile fase består behandlingen af inhalationspræparater (21), da disse mindsker slimproduktionen i luftvejene og gør, at den glatte muskulatur i bronkierne dilaterer, hvorved dyspnøen mindskes (22). Der gives oftest bronkodilaterende beta2-agonister eller antikolinergika som initierende behandling af dyspnø. Ved svær grad af KOL, hvor borgeren har oplevet flere exacerbationer, kan der suppleres med inhalationsglukokortikoid. I de sværeste tilfælde af KOL, hvor der opstår respirationssvigt, kan der tilbydes ilt i hjemmet (21,23).

Behandlingen i den ustabile fase foregår på hospitalet, hvor borgeren med KOL først behandles med ilt gennem en maske eller et næsekateter, så der opretholdes en iltmætning i blodet på 90-92%. Kuldioxidindholdet i blodet observeres nøje, da en forhøjet mængde af kuldioxid indikerer en nedsat iltoptagelse og dermed er et advarselstegn på risiko for øget mortalitet. Der gives inhalationsmedicin for at lette borgerens respiration, og der suppleres med glukokortikoider, som f.eks. prednisolon enten i oral eller intravenøs form, for at mindske inflammationen. Der kan yderligere suppleres med antibiotika, hvis der mistænkes en bakteriel infektion, hvilket undersøges nærmere via en blodprøve. (21) Er ovenstående behandling ikke effektiv, kan der gives en behandling kaldet *Noninvasive Ventilation* (NIV). Her påføres borgeren med KOL en tætsluttende maske, som blæser luft ned i lungerne med varierende tryk. Ved NIV-behandlingen blæses størst tryk ved en inspiration, for at luften kommer ud til alveolerne, og lavere tryk under eksspirationen. Det lave blæsetryk ved eksspirationen bidrager til, at bronkierne og bronkiolerne undgår at klappe sammen, hvorved eksspirationen bliver mere effektiv i udvekslingen af ilt og kuldioxid. Vurderes NIV-behandlingen ikke effektiv, kan borgeren med KOL lægges i respirator, såfremt det vurderes, at borgerens tilstand hurtigt kan forbedres, ellers kan behandlingen med respirator være ineffektiv. (21)

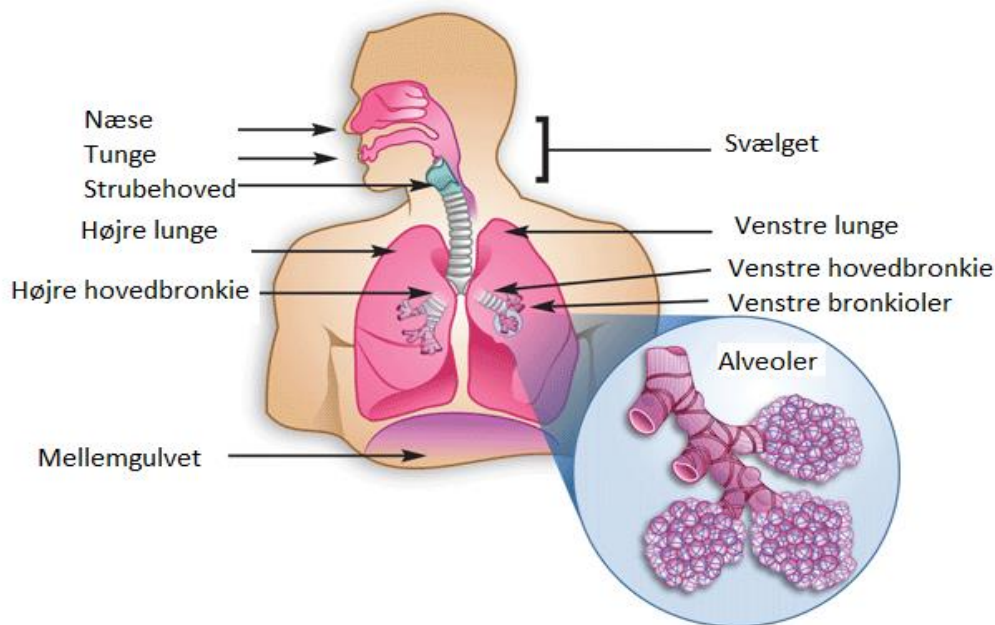
1.3 Samfundsomkostninger forbundet med KOL

Den store befolkningsgruppe, som i Danmark diagnosticeres med KOL, er forbundet med øgede samfundsøkonomiske omkostninger sammenlignet med baggrundsbefolkningen (11). Et dansk studie af Bilde et al. (10) har vist, at det især er hospitalsindlæggelser, der er forbundet med en øget samfundsøkonomisk omkostning i forbindelse med behandling af borgere med KOL. Resultaterne fra studiet af Bilde et al. (10) viser, at borgere med KOL i 2002 udgjorde en ekstra udgift på 1,9 mia. kr. sammenlignet med danske borgere uden KOL. Dette svarer til 6% af det danske forbrug af hospitalsvæsen og primær sektor, selvom denne gruppe kun udgør 2,6% af hele den danske befolkning. (10) Samtidig ses en statistisk signifikant sammenhæng mellem, at borgere med KOL hyppigere får behandling i eget hjem og på sygehus, får offentlig forsørgelse samt hyppigere anvender medicin og sundhedsforsikring sammenlignet med borgere uden KOL. Disse parametre gør, at borgere med KOL i Danmark er forbundet med en øget samfundsøkonomisk omkostning på ca. 65.000 kr. årligt pr. borger, der diagnosticeres med KOL (24).

En undersøgelse viser, at det især er borgere med KOL i grad 4, som er lig med flest udgifter og hospitalsindlæggelser (25). Indlæggelsestiden er derfor vigtig i forklaringen på, hvorfor hospitalsindlæggelser er en stor økonomisk post hos borgere med KOL. Et systematisk review af Rehman et al. (26) viser, at forskellen på omkostningerne for hospitalsindlæggelser mellem borgere med KOL i grad 1 og 4 er hhv. 9.786 kr. til 63.001 kr. gennemsnitligt i Europa pr. år. Et svensk studie af Andersson et al. (27) viser, hvordan omkostningerne i løbet af 4,5 måneder fordelte sig alt efter borgerens sværhedsgrad af KOL. Borgere med KOL i grad 1 var forbundet med omkostninger af exacerbationer på 143kr, hvor borgere med KOL i grad 4 var forbundet med omkostninger på 25.973kr. (27).

1.4 Det respiratoriske system hos borgere med KOL

Respirationssystemet er essentielt for, at menneskekroppens celler kan fungere, idet systemets primære funktion er at tilføre ilt via luftvejene til blodbanen og videre ud i kroppens celler samt at udskille kuldioxid den modsatte vej. Det respiratoriske system består af de øvre luftveje, som udgøres af næsen, mundhulen, svælget, struben og den øverste del af luftrøret, og de indre luftveje, som udgøres af den nederste del af luftrøret, som deler sig i to hovedbronkier, der føres til hver deres lunge. Bronkierne deler sig i mindre bronkier og bronkioler for til sidst at ende i alveoler, som består af mange små luftblærer, der danner et stort overfladeareal. Det er her, udvekslingen af ilt og kuldioxid foregår. (22) Opbygningen af luftvejene ses på figur 1.2.



Figur 1.2: Viser en oversigt over den øvre og nedre luftvej, hvor den nederste del af luftvejen deler sig i to hovedbronkier, til hver sin lunge. Den blå cirkel fremhæver alveolerne, som indikeres ved små cirkler, som danner et stort overfladeareal. Billedet er taget fra Wikipedia, hvor teksten er blevet ændret til dansk (28).

Under normale forhold er alveolernes samlede overfladeareal ca. 70 m², hvilket skaber gode betingelser for udvekslingen af ilt og kuldioxid, dvs. diffusionskapaciteten hos en rask person er stor. Ved en inspiration fyldes lungerne med luft, og ilt diffunderer, dvs. spredes passivt, fra alveolerne over i blodbanen, mens kuldioxid diffunderer den modsatte vej og udskilles ved eksspiration til omgivelserne. (22)

KOL er karakteriseret ved en irreversibel forsnævring af luftvejene og nedbrydning af lungevævet, hvilket medfører nedsat lungefunktion og tiltagende respirationsbesvær (29). Ved KOL sker der en kronisk inflammation i bronkierne. Dette øger slimproduktion og fortykkelse af slimhinderne i bronkierne, hvilket øger luftmodstanden, som medfører besværet respiration, kaldet dyspnø. Den kroniske inflammation kaldes kronisk bronkitis og påvirker særligt eksspirationen, da lungernes elasticitet nedsættes. (21) Under normale forhold gør lungevævets elasticitet, at lungerne har let ved at blive udvidet og trække sig sammen igen, og samtidig bidrager elasticiteten til at holde de nedre luftveje åbne. Ved KOL tabes elasticiteten, hvilket gør lungerne slappe, og det bliver vanskeligt at komme af med luften, fordi de indre luftvejes diameter bliver mindre, hvorved modstanden øges. Den kroniske bronkitis, der forekommer i lungerne ved KOL, kan også medføre nedbrydning af alveolevæggene. Herved opstår der uensartede hulrum, hvilket mindsker overfladearealet og

dermed påvirker udvekslingen af ilt og kuldioxid. Tilstanden kaldes emfysem. (29,30) Ved et mindsket overfladeareal, bliver udvekslingen af ilt og kuldioxid utilstrækkelig. Dette betyder, at der ophobes kuldioxid i blodet, hvilket på længere sigt kan øge risikoen for komplikationer grundet ændringer i syre-base-balancen (22). Disse komplikationer kan være en respiratorisk forgiftning, som respiratorisk acidose, der forårsager ændringer i syre-base-balancen, som påvirker cellernes funktion, hvilket i værste tilfælde kan føre til død. (31)

1.5 Borgerens livssituation påvirkes af KOL

KOL påvirker ikke kun en borger fysisk. Det rammer også sociale og personlige forhold, som kan ændre en borgers livssituation. KOL rammer både mænd og kvinder, og oftest efter 40 årsalderen (19), hvor der ses en generel påvirkning af borgerens livssituation på den sociale, følelsesmæssige og økonomiske front, som bl.a. har indflydelse på livskvaliteten (32).

Et narrativt review af Miravittles et al. (33) fremhæver symptomer ved KOL, som ofte kan give problemer som angst og depression, hvor borgeren bliver mere isoleret fra sin omgangskreds. Nogle borgere oplever, at de er en byrde for deres venner og familie, hvor andre isolerer sig, fordi de oplever angst, når deres vejtrækning bliver for besværet (33). Angst og depression ses oftest hos den yngre befolkningsgruppe med KOL og især hos kvinder (19). Borgeren kan samtidig opleve at følelser og attitude kan blive mere negativt præget. Dette kan skyldes de begrænsninger, som KOL kan medføre, eksempelvis problematikker ved at gå på arbejde eller dyrke en fritidsaktivitet, hvilket kan medføre en mindsket motivation og en følelse af ligegyldighed. Disse følelser kan præge en borgers personlighed negativt, hvilket kan medføre en identitetskrise og en nedsat livskvalitet. (32) Ifølge et studie af Sigurgeirsdottir et al. (34) kan de personlige problemstillinger vise en accept omkring at skulle bede andre om hjælp eller ved de begrænsninger som KOL medfører (34). Disse ændringer i en borgers livssituation kan medføre en påvirkning af sociale og følelsesmæssige aspekter, som identificerer en borger, hvorved der kan opstå en nedsat livskvalitet (32,33).

Et narrativ review af Cruz et al. (32) beskriver, hvordan den socioøkonomiske status også kan påvirke en borgers livssituation. Grundet en øget sværhedsgrad af KOL, kan en borger være nødsaget til at nedsætte deres arbejdstid for at passe på helbredet eller ændre deres levemåde, fordi de har et mindre rådighedsbeløb (32). Her ses der en forskel mellem højt og lavt stillede borgere rent økonomisk, hvor den lavt stillede, ikke nødvendigvis har mulighed for at gå ned i arbejdstid (32). Dette skaber forskellige problemstillinger for højt og lavt stillede borgere, hvor det kan hænde, at den lavt stillede borger går på kompromis med egen sundhed. Et svensk studie af Borné et al. (35) viser, at en borger med et lavt stillede arbejde, lav indkomst eller der er skilt, er i større risiko for en

hospitalsindlæggelse, hvilket kan hænge sammen med en lav *self-management*. En lav *self-management* kan betyde, at en borger har svært ved at tage ansvar for egen sygdom. (36) Det er derfor varierende problemstillinger, KOL kan medføre hos den enkelte borger, hvor der er en sammenhæng mellem alder, sværhedsgraden af KOL og socioøkonomisk status, hvilket skaber en stor variation i befolkningsgruppen med KOL. Der ses derfor en stor samlet befolkningsgruppe med KOL, men som adskiller sig fra hinanden på flere punkter.

1.6 Konsekvenser af exacerbationer

Der findes flere forskellige definitioner af exacerbationer, da det kan defineres på baggrund af forskellige årsager og sværhedsgrader af KOL (37), hvorfor organisationen GOLD anvender en forholdsvis bred definition: *“COPD exacerbations are defined as an acute worsening of respiratory symptoms that result in additional therapy”* (19). En exacerbation kan forstås som en akut forværring af symptomer, der kræver yderligere eller anden behandling alt efter sværhedsgraden af forværringen (38). I milde tilfælde af exacerbationer kan behandlingen bestå af en ændring i medicineringen, hvor det i svære tilfælde er nødvendigt med hospitalsindlæggelse (21). Akutte og hyppige exacerbationer, hvor hyppige defineres som to eller flere exacerbationer årligt (19), forårsager en øget progression af KOL (38). Derfor er det i dette speciale valgt at definere exacerbationer som akutte forværringer med hospitalsindlæggelse til følge, da det er dette, som adskiller sig fra den almindelige behandling af KOL.

En stor del af sundhedsudgifterne, som er forbundet med KOL, kommer som følge af akutte exacerbationer (19,25,26,38,39), hvilket gør sig gældende i både USA, Asien og Europa (25). De sundhedsudgifter som koster samfundet mest ved behandling af akutte exacerbationer, er akuthjælp, ekstra primærpleje, hospitalsindlæggelse og intensivbehandling. Derudover skal borgeren efter en akut exacerbation ofte have efterbehandling for at imødekomme den fysiske, sociale og følelsesmæssige svækkelse, personen har fået ved exacerbationen. (19,38) Den kroniske forringelse af borgerens sundhedsmæssige tilstand er med til at øge risikoen for yderligere exacerbationer og hermed risikoen for hospitalsindlæggelse (40–42). Et studie viser, at en ringere sundhedstilstand også øger risikoen for flere indlæggelsesdage i forbindelse med hospitalsindlæggelse, hvilket yderligere øger risikoen for genindlæggelse (43).

Grundet den store udgift forbundet med exacerbationer er det generelt vigtigt at opdage og behandle exacerbationerne i et tidligt stadie (38). Ubehandlede exacerbationer forringer borgerens livskvalitet, forlænger behandlingsforløbet og er medvirkende til en hurtigere forringelse af

lungefunktionen, hvilket påvirker sundhedsudgifterne (44,45). Et systematisk review af Toy et al. (39) viser, at flere borgere med KOL ikke rapporterer deres exacerbationer, hvorved der ofte allerede er sket en kronisk forringelse af den enkeltes helbred, når de kommer til kontrol af deres KOL (38,46,47). Dette kan skyldes borgerens socioøkonomiske status eller, at borgeren med KOL ikke er bekendt med, hvilke symptomer der skal handles på (19,20).

Uopdagede exacerbationer kan således få store økonomiske konsekvenser, idet hospitalsindlæggelser ifølge Rehman et al. (26) er omdrejningspunktet for de største økonomiske udgifter i forbindelse med exacerbationer. De fremviser en tabel over den samlede omkostning af exacerbationer i Danmark svarende til 65.691 kr. pr. person, hvor hospitalsindlæggelser udgør 38.361 kr. (25) Dette vil sige, at over halvdelen af udgifterne til behandling af exacerbationer går til hospitalsindlæggelser. En indlæggelsesdag for en borger med KOL koster i gennemsnit 4100 kr. (48), og en borger med KOL har gennemsnitligt 10 indlæggelsesdage, hvor en borger uden kroniske sygdomme gennemsnitligt har 5 indlæggelsesdage. (49) En borger med KOL er derfor forbundet med dobbelt så mange indlæggelsesdage, som en borger uden kronisk sygdom ved en hospitalsindlæggelse grundet exacerbationer.

Organisationen GOLD oplyser bl.a. at for at håndtere KOL, skal forebyggelse og behandling af exacerbationer være et centralt mål (19), hvorfor det findes relevant at lave indsatser, som har til formål at forebygge og opspore exacerbationer. En tidlig opsporing af exacerbationer skal bidrage til øget livskvalitet og forebyggelse af hospitalsindlæggelser, hvorved de økonomiske udgifter forbundet med denne post evt. kan reduceres i fremtiden.

1.7 Telemedicin

Der er kommet stort fokus på telemedicin som en eventuel løsning til øget livskvalitet og reducere af de samfundsøkonomiske omkostninger, der er forbundet med KOL, ved at opspore tidlige exacerbationer (39). I Danmark består telemedicin til borgere med KOL bl.a. af telemonitorering, som består af en tablet samt måleudstyr til at måle iltmætning, blodtryk, puls og vægt, som automatisk sendes til kommunen eller hospitalet. Sundhedspersonalet følger borgerens værdier hver uge, og ved ændret tilstand kontaktes borgeren mhp. revurdering af behandlingen. (50) Et studie af Sood et al. (51) har defineret telemedicin som:

“Telemedicine being a subset of telehealth, uses communications networks for delivery of healthcare services and medical education from one geographical location to another, primarily to address challenges like uneven distribution and shortage of infrastructural and human resources” (51).

Denne definition af telemedicin er udarbejdet ud fra 104 peer-reviewed artikler (51) og omfavner en bred definition, men vurderes relevant for dette speciale. Tidlig opsporing af exacerbationer skal forhindre mulige hospitalsindlæggelser, som er den største samfundsøkonomiske omkostning forbundet med KOL (26). Samtidig kan borgeren blive undervist i KOL og herved øge sin egen *self-management* gennem telemedicin. Borgeren med KOL bliver derfor undervist i, hvilke symptomer der skal observeres ved en exacerbation, da borgerne ikke nødvendigvis er opmærksomme på, at de ikke rapporterer exacerbationer og hvor stor risikoen er for hospitalsindlæggelser ved ikke-rapporterede exacerbationer. (52) Muligheden for at opspore tidligere exacerbationer og øge *self-management* bakkes op af et cochrane review, som viser, at telemedicin kan reducere risikoen for hospitalsindlæggelser signifikant (53). På trods af dette viser litteraturen ingen effekt på de samfundsøkonomiske omkostninger ved anvendelsen af telemedicin (8,54,55). Dog blev der i 2016 alligevel blevet besluttet, i et samarbejde mellem Regeringen, Kommunernes Landsforening og Danske Regioner, at telemedicin skulle udbredes nationalt til alle borgere med KOL ved udgangen af 2020 (1).

Witt Udsen et al. (54) stiller spørgsmålstegn ved, om den manglende økonomiske effekt kan skyldes, at gruppen af borgere med KOL måles for bredt, hvor sværhedsgraden af KOL, kan have en betydning. Studiet viser, at der hos borgere med KOL i grad 3 kan opnås *Quality Adjusted Life Years* (QALY) og spares omkring 7000 kr. årligt pr. borger ved anvendelse af telemedicin. Selvom resultaterne ikke er statistisk signifikante, indikerer de, at der kan opnås besparelser, og at disse besparelser primært kan opnås ift. hospitalsindlæggelser. (54) De foreslår, at fremtidige studier undersøger hvilke borgere med KOL, der opnår størst økonomisk effekt (54). Fordelen ved at kigge på dette er, at telemedicin evt. er givet til flere yngre borgere med KOL, som vil have lettere ved at anvende en tablet og udføre telemonitorering, hvorfor datasættet evt. ikke afspejler den ønskede målgruppe. Ældre borgere er oftest forbundet med en højere sværhedsgrad af KOL og deraf flere exacerbationer, hvorfor det er disse borgere, det vil være relevante at undersøge. Her viser et studie af Reed et al. (56), at det oftest er de yngre borgere og personer, der har let adgang til teknologiske midler, som vælger telemedicin. Et studie af Vianello et al. (57) konkluderer, at telemedicin til håndtering af exacerbationer ikke skal generaliseres til alle borgere med KOL, men at forskere skal bestræbe sig på at identificere hvilke borgere med KOL, der kan få størst gavn af telemedicin.

Et tysk studie af Achelrod, Schreyögg & Stargardt (5) viser, at en gruppe af borgere i grad 3 og 4 opnåede en højere økonomisk besparelse samt færre og kortere hospitalsindlæggelser sammenlignet med borgere med KOL i grad 1 og 2, ved anvendelse af telemedicin. Et studie af Haesum et al. (6), som fokuserer på telemedicin til borgere med KOL i grad 3 og 4, viser en samfundsøkonomisk forbedring i denne gruppe. Dette tyder på, at det er borgere med KOL i bestemte sværhedsgrader, der vil få størst økonomisk effekt ved behandling med telemedicin (6), hvorfor det findes anvendeligt at prædiktere hvilke borgere med KOL, der skal tilbydes telemedicin.

1.8 Prædiktionsalgoritmer

En prædiktionsalgoritme bygger på historisk data, hvor der anvendes statistiske algoritmer og *machine learning* til at forudsige mønstre i et datasæt og sandsynligheden for et fremtidigt resultat (58). For at kunne anvende en prædiktionsalgoritme i sundhedssektoren, skal den have et klinisk formål samt opnå en god validitet i at prædiktere et sundhedsmæssigt resultat (59). Dette kan gøres ved at inddele det samlede datasæt i et trænings- og testsæt. Her udføres der en validering på træningssættet, og efterfølgende på testsættet, hvilket kan gøres på forskellige måder. (59)

Prædiktionsalgoritmer kan udvikles forskelligt, og flere forskere har forsøgt at udvikle prædiktionsalgoritmer til at prædiktere risikoen for udvikling af exacerbationer hos borgere med KOL (60–64). De forskellige prædiktionsalgoritmer resulterer i forskellige nøjagtigheder samt sensitivitet og specificitet, hvilket bl.a. hænger sammen med den anvendte model, anvendelsen af forskelligt data og forskellige definitioner af exacerbationer (60–64).

Et systematisk review af Sanchez Morillo et al. (3) har identificeret flere studier, hvor prædiktionsalgoritmer anvendes til at prædiktere exacerbationer hos borgere med KOL. To studier af Fernandez-Granero et al. (61) og Van der Heijden et al. (62) opnår eksempelvis en forholdsvis god nøjagtighed og sensitivitet i to forskellige prædiktionsmodeller. Fernández-Granero et al. (61) anvender en *probabilistic neural network* og opnår en nøjagtighed på 0,88 og en sensitivitet på 0,81. Van der Heijden et al. (62) anvender en *bayesian network* og opnår en nøjagtighed på 0,81 og en sensitivitet på 0,88. Dog indikerer et systematisk review af Sanchez Morillo et al. (3), at de opnåede resultater ikke kan opretholde et godt klinisk niveau for at kunne anvendes i praksis, hvorfor prædiktionsalgoritmerne skal forbedres yderligere.

En mulig løsning, for at kunne optimere en prædiktionsalgoritme, er at finde hvilke borgere der skal have tilbudt telemedicin, da dette ikke er blevet undersøgt. Ifølge Riis et al. (4) bør prædiktive algoritmer tilpasses de individuelle forskelle, der ses blandt borgere med KOL. Det kunne eksempelvis være ved at stratificere borgere med KOL, for at finde ud af om dette kan være med til at styrke prædiktionsalgoritmerne. Witt-Udsen et al. (54) skriver, at udviklingen af exacerbationer afhænger af sværhedsgraden af KOL, hvilket støttes op af organisationen GOLD (19) og Anzueto et al. (38) som belyser, at borgere med KOL i grad 3 og 4 har en signifikant højere risiko for at udvikle exacerbationer end borgere med KOL i grad 1 og 2. Derfor bør det overvejes, om det kan svække prædiktionsalgoritmernes prædiktionssevne at inddrage alle borgere med KOL, da gruppen dækker over en bred vifte af faktorer, som kan skabe mulige bias. Fordelen ved at prøve at udvikle prædiktionsalgoritmer på stratificerede grupper er, at det kan gøre det lettere for sundhedsprofessionelle at udvælge de borgere med KOL, som kan få størst effekt af telemedicin.

Telemedicin kan derfor være med til at forhindre exacerbationer, men ikke hos alle borgere med KOL, hvorfor prædiktionsalgoritmer kan hjælpe med at identificere, hvilke borgere med KOL der skal tilbydes telemedicin, ud fra en stratificeret gruppe.

1.9 Problemafgrænsning

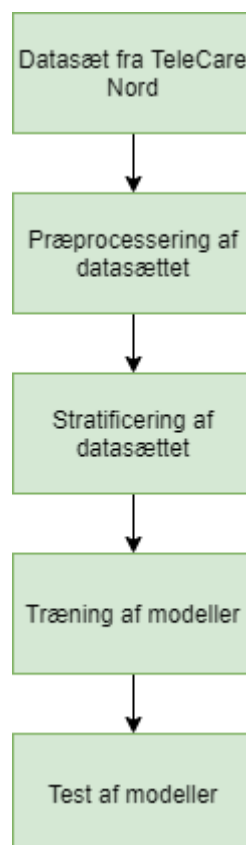
KOL er en irreversibel sygdom, som er forbundet med samfundsøkonomiske omkostninger og nedsat livskvalitet for den enkelte borger. Telemedicin kan være en løsning til at forhindre exacerbationer, der mindsker livskvaliteten og øger samfundsomkostningerne. Udfordringen ved telemedicin er, at ikke alle borgere med KOL opnår en øget livskvalitet eller bidrager til mindskede samfundsøkonomiske omkostninger. Flere studier har lagt fokus på at udvikle prædiktionsalgoritmer, som skal identificere, hvilke borgere der er i størst risiko for exacerbationer. Prædiktionsalgoritmerne opnår ikke tilstrækkelige resultater til at kunne anvendes i klinisk praksis, hvorfor der ses en fordel i at målrette, hvilke borgere telemedicin skal gives til. Der vil blive taget udgangspunkt i at udvikle en prædiktionsalgoritme, som har et samfundsøkonomisk perspektiv, hvorfor der vil blive prædikeret ud fra indlæggelsesdage. Indlæggelsesdage vælges, fordi den største omkostningspost hos borgere med KOL er hospitalsindlæggelser, hvor denne omkostningspost afhænger af antal indlæggelsesdage. Der vil blive undersøgt forskellige stratificerede grupper som vurderes relevante inden for borgere med KOL ud fra litteraturen. Formålet med dette speciale er derfor at finde, hvilke borgere med KOL der skal tilbydes telemedicin, ved at prædiktere ud fra stratificerede grupper.

1.10 Problemformulering

Hvordan kan en prædiktionsalgoritme prædiktere hvilke borgere med KOL, der er i risiko for flest indlæggelsesdage?

2. Metode

Til besvarelse af specialets problemformulering tages der udgangspunkt i et retrospektivt studiedesign med en kvantitativ metodisk tilgang. Metodeafsnittet giver derfor først et overblik over datasættet fra TeleCare Nord, som specialet tager udgangspunkt i. Herefter beskrives, hvordan data præprocesseres og stratificeres, hvorefter der redegøres for træning af de to prædiktionsmodeller, Multipel Lineær Regression og Random Forest Regression. Efterfølgende beskrives hvordan prædiktionsalgoritmerne valideres i "Test af modeller". Til sidst belyses etiske overvejelser. Nedenstående figur 2.1 illustrerer de forskellige trin i den metodiske tilgang til udviklingen af prædiktionsalgoritmerne.



Figur 2.1: Illustration over den metodiske tilgang til udviklingen af prædiktionsalgoritmer. Datasættet præprocesseres og stratificeres i grupper ud fra udvalgte variabler, hvorefter prædiktionsmodellerne fremgår i "Træning af modeller" efterfulgt af en validering af prædiktionsalgoritmerne i "Træning af modeller".

2.1 Datasæt fra TeleCare Nord

Data, som anvendes i dette speciale, stilles til rådighed af forskningsprojektet TeleCare Nord. I dag er TeleCare Nord en telemedicinsk servicefunktion i Nordjylland, der har udført forskellige projekter med telemedicin. (65) Der er bl.a. udført et projekt med telemedicin til borgere med KOL, hvor der i perioden 2012-2015 blev indsamlet data fra 1225 borgere med KOL i Region Nordjylland, fordelt over 11 nordjyske kommuner. Der blev indsamlet data ved baseline, løbende via telemonitorering, til årlig opfølgning og til opfølgning sidst i projektperioden. Ved baseline blev der indsamlet sociodemografiske data, eksempelvis rygestatus, uddannelse og sværhedsgrad af KOL. Telemonitoreringen bestod primært af en opfølgning på borgerens helbred bl.a. via måling af vitale værdier (66).

Til dette speciale stilles der 56 variabler til rådighed, hvoraf 52 variabler er indsamlet ved baseline og derfor udgør mulige prædiktorer, mens fire variabler er outcomes, som er indsamlet ved projektets afslutning, se appendix 4. Outcomet "Antal indlæggelsesdage" vælges som outcome, fordi det i dette speciale er varigheden af hospitalsindlæggelser, der er vurderet som den bedste indikator for omkostninger i forbindelse med behandlingen af KOL.

2.2 Præprocessering af datasættet

Specialets datasæt indeholder 52 prædiktorer og ét outcome, dvs. 53 variabler i alt, som skal præprocesseres, inden de kan anvendes til udvikling af prædiktionsalgoritmerne. De 53 variabler består af 37 kontinuerlige variabler, 12 nominale variabler og 4 ordinale variabel, se appendix 4. Præprocesseringen foregår i SPSS 26 og Excel.

Der beregnes, hvor mange procent manglende data de enkelte variabler har. Alle variabler med mere end 20% manglende data ekskluderes, da dette er en uacceptabel andel af manglende data (67). Der laves efterfølgende et estimat for de manglende data i de resterende variabler vha. *imputation model*. Dette gøres i SPSS hvor der vælges en automatisk imputationsmetode, hvorved der enten vælges en *Markov Chain Monte Carlo* metode hvis der ikke er noget mønster i, hvor der er manglende data. Er mønsteret derimod monotont i hvilke data der mangler, vælges en imputationsmetode der kan tage højde for dette (68). Det monotone mønster opstår, hvis der er mange patienter, der er udgået fra studiet, hvorefter de vil udgå fra den videre dataopsamling. Alle borgere uden indlæggelsesdage fjernes, da disse ikke findes relevante for specialets formål.

2.3 Stratificering af datasættet

Efter præprocesseringen stratificeres data i syv grupper ud fra udvalgte variabler, hvorved der dannes et separat datasæt for hver gruppe. Dette gøres for at skabe mindre grupper at udvikle prædiktionsalgoritmerne ud fra og på denne måde undersøge, hvordan stratificering påvirker resultaterne. Sammen med hele datasættet er der i alt syv datasæt, som skal anvendes til udvikling og validering af hver deres prædiktionsalgoritme baseret på hhv. MLR og RFR. Stratificeringen udføres mhp. at undersøge, om prædiktionsalgoritmnernes prædiktionssevne kan forbedres ved sammenligning med hele datasættet. I tabel 2.1 ses de syv grupper.

Stratificering af data		
<i>Hele datasættet</i>	Den overordnede gruppe	
<i>Sværhedsgrad af KOL</i>	Grad 3	Grad 4
<i>Rygestatus</i>	Rygere	Ikke rygere
<i>Alder</i>	71-80 år	≥ 81 år

Tabel 2.1: Oversigt over de syv grupper som anvendes. Her ses, hvilke variabler de syv grupper stratificeres ud fra. Disse er markeret med kursiv i den første kolonne.

Den overordnede gruppe anvendes mhp. at sammenligne resultaterne for denne gruppe med resultaterne for hver af de seks grupper for at undersøge, om stratificeringen af data har betydning for resultaterne. De seks grupper er stratificeret ud fra variablerne "Sværhedsgrad af KOL", "Rygestatus" og "Alder", som indgår i "Den overordnet gruppe" og alle har betydning for udviklingen af exacerbationer og hermed risikoen for hospitalsindlæggelser (7,17,38,54). Sværhedsgraden af KOL hænger sammen med, hvor fremskreden sygdommen er, og således også risikoen for exacerbationer, hvorfor det kun er grad 3 og 4 der undersøges (38,54). Ofte hænger sværhedsgraden af KOL sammen med borgerens alder, da de skal have haft sygdommen i en længere periode for at have udviklet KOL i grad 3 og 4 (17). Det forventes derfor, at risikoen for exacerbationer stiger med alderen, hvorfor der anvendes aldersgrupperne 71 til 80 år og over 80 år. Rygning øger risikoen for exacerbationer og progressionen af KOL, hvorfor der undersøges hvorvidt rygning påvirker resultaterne for prædiktionsalgoritmerne (7).

Efter stratificeringen af data deles hver af de syv datasæt op i et trænings- og et testsæt.

Træningssættet udgør 70% af data, og testsættet udgør 30%.

2.4 Træning af model

I dette afsnit beskrives de to prædiktionsmodeller Multipel Lineær Regression og Random Forrest Regression. De to prædiktionsmodeller skal bruges til at udvikle prædiktionsalgoritmer på træningssættet på de 7 grupper.

2.4.1 Multipel Lineær Regression

Multipel Lineær Regression (MLR) er en statistisk metode til at prædiktere et afhængigt kontinuert outcome baseret på flere uafhængige variabler. Med denne metode kigges på det lineære forhold mellem outcomet og de uafhængige variabler. (69)

Udvælgelsen af de variabler der er bedst til at prædiktere antallet af indlæggelsesdage foregår i SPSS med *forward selection* på træningssættet. Ved *forward selection* beregnes, hvilke variabler og hvilken kombination af disse der bedst kan prædiktere outcomet (70). Dette gøres ved at tilføje variabler i den rækkefølgede prædikterer bedst, indtil tilføjelse af yderligere variabler ikke længere forbedrer prædiktionsalgoritmen (71,72).

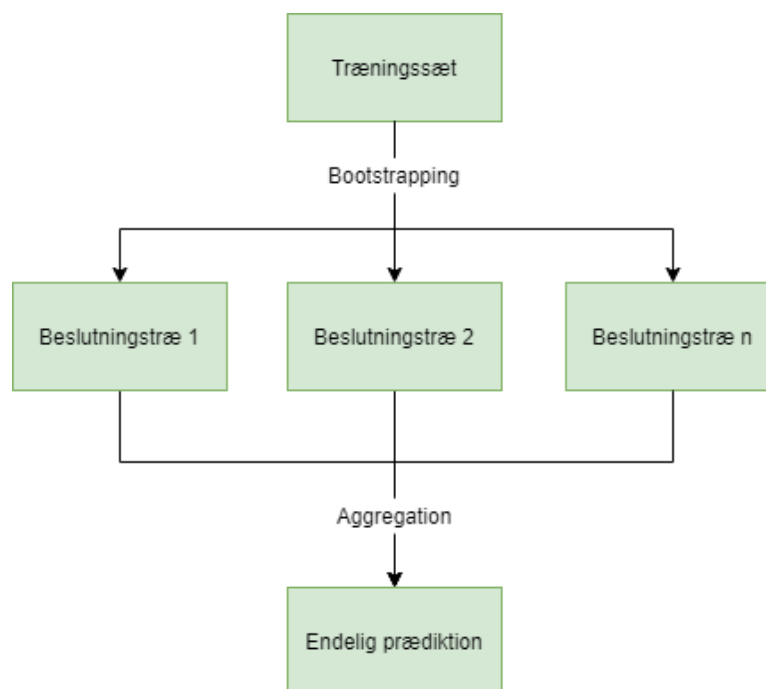
Efter *forward selection* vurderes ud fra variablernes determinationskoefficient (R^2), hvorvidt alle de tilføjede variabler skal medtages i prædiktionsalgoritmen. Determinationskoefficienten benyttes som en *goodness of fit test*, hvor den belyser sammenhængen mellem outcomet og de uafhængige variabler (73). En determinationskoefficient på 1 angiver en perfekt korrelation, og værdien 0 angiver ingen korrelation (74). Her kigges på, hvor meget determinationskoefficienten stiger ved tilføjelse af de enkelte variabler. Der findes ikke en bestemt værdi determinationskoefficienten bør stige med (75). Der er i dette speciale anvendt en stigning på 0,02. Når intervallet mellem to variabler stiger med mindre end 0,02, kan der argumenteres for, at prædiktionsalgoritmen ikke opnår en forbedring, der er stor nok til at retfærdiggøre tilføjelse af flere variabler. Dette kan mindske prædiktionsalgoritmens styrke, hvorfor de efterfølgende variabler ekskluderes. (76)

2.4.2 Random Forrest Regression

Random Forrest Regression (RFR) er en metode, som er baseret på beslutningstræer. Der anvendes i dette speciale regression ved *Random Forest*, da datasættets outcome er kontinuert. RFR bygger på en *ensemble*-metode, hvor modellen prædikterer ud fra flere randomiserede beslutningstræer,

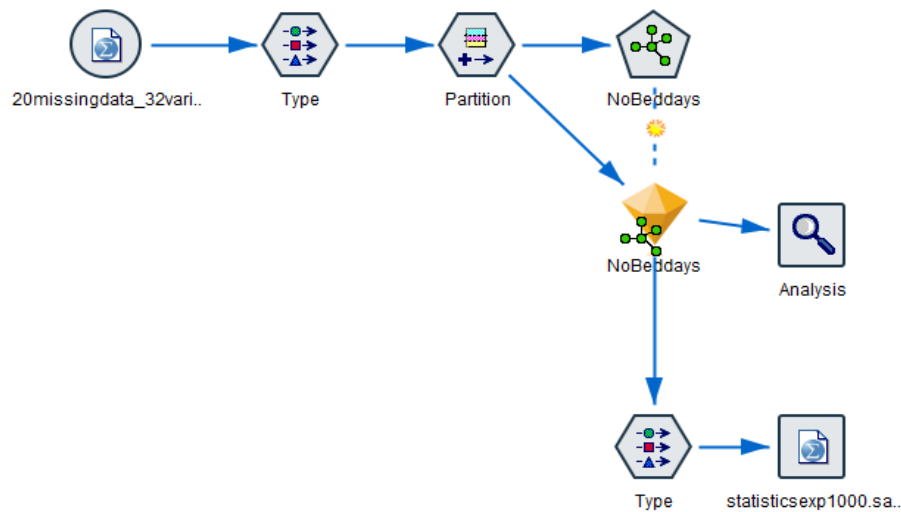
hvilket resulterer i en robust prædiktions. (77) I dette speciale anvendes 1000 beslutningstræer for at mindske risikoen for tilfældige fejl (78). For at undgå, at prædiktionsalgoritmen bliver for specifik, vælges en *node size* på 5 (78).

Algoritmen er opbygget efter *bootstrapping* og *aggregation*. *Bootstrapping* anvendes til at udvikle beslutningstræer, hvilket gøres ved at vælge et tilfældigt antal borgere til hvert beslutningstræ samt medtage et tilfældigt antal variabler, hvor der ændres på rækkefølgen af disse for hvert beslutningstræ. *Aggregation* udleder det endelige resultat for hver borger ved at tage det gennemsnitlige resultat for de kontinuerte variabler, og ved de kategoriske variabler beregnes *den hyppigste værdi* fra de dannede beslutningstræer (77). Beslutningsmetoden for RFR vises i nedenstående figur 2.2. (79)



Figur 2.2: Illustration over hvordan beslutninger træffes i RFR ved anvendelse af bootstrapping og aggregation.

Udviklingen af RFR foregår i SPSS Modeler, hvor der beregnes en *predictor importance*, som angiver, hvor vigtige de tilføjede variabler er ift. RFR. Hver variabel, som er tilføjet i prædiktionsalgoritmen, tildeles et tal mellem 0 og 1, hvor summen af disse tal giver 1 (80). Nedenstående figur 2.3 viser, hvordan modellen er opstillet i SPSS Modeler.



Figur 2.3: Opbygningen af prædiktionsalgoritmen i SPSS Modeler.

2.5 Test af modeller

Valideringen af prædiktionsalgoritmerne baseret på MLR og RFR udføres på testsættet mhp. at vurdere deres prædiktionssevne, overfitting samt generaliserbarheden af resultaterne. Valideringen foregår i SPSS og SPSS-Modeler ved først at prædiktere antal indlæggelsesdage ud fra testsættet vha. de udviklede prædiktionsalgoritmer for hver af de syv grupper. Herefter udføres en korrelationsanalyse, hvor der beregnes en determinationskoefficient, som belyser sammenhængen mellem de faktiske værdier og de prædikterede værdier for antal indlæggelsesdage for hver gruppe. Resultaterne fra korrelationsanalysen visualiseres i et scatterplot for "Den overordnede gruppe" og de tre stratificerede grupper som opnår bedst resultat på deres prædiktionsmodel. Der beregnes en regressionslinje, som går igennem de estimerede gennemsnit for hver af de prædikterede og de faktiske værdier. Derefter beregnes konfidensgrænser, som viser, hvor sikkert estimatet af gennemsnittet er, hvilket vurderes ud fra intervalstørrelsen mellem regressionslinjen og konfidensgrænserne (81).

2.6 Etiske overvejelser

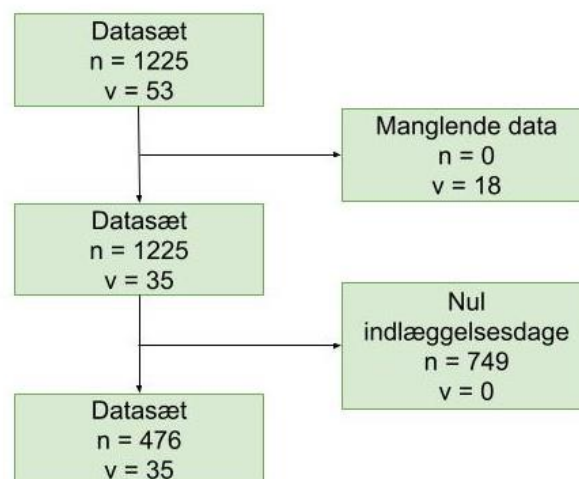
I dette speciale anvendes der retrospektivt data fra TeleCare Nord-projektet, hvorfor de inkluderede borgere i TeleCare Nord-projektet allerede har givet samtykke til, at data kan anvendes til videre forskning (54). Derved er der ikke anmeldelsespligt hos Videnskabetisk Komite. Det inddragede data indeholder ikke personfølsomme data og anonymiseres, så borgerne ikke kan identificeres, hvormed der ligeledes ikke er anmeldelsespligt hos Datatilsynet. Specialet er udført i henhold til Helsinki-deklarationen.

3. Resultater

I resultatafsnittet fremvises udvælgelsen af, hvilke borgere og variabler der medtages i datasættet, samt hvilke variabler der frasorteres grundet manglende data. Efterfølgende vises en oversigt over de syv grupper og hvor mange borgere og variabler hvert datasæt består af. Der vises en oversigt over de variabler, som medtages i de tre bedste prædiktionsalgoritmer og variablernes betydning. Herefter vises de opnåede determinationskoefficienter for hver af de udviklede prædiktionsalgoritmer. Til sidst illustreres "Den overordnede gruppe" og de tre bedste prædiktionsalgoritmer i hver sit scatterplot.

3.1 Præprocessering af datasættet

Figur 3.1 viser en oversigt over, hvor mange borgere og variabler der frasorteres i præprocesseringen for, at datasættet indeholder de mest relevante borgere og variabler. Der frasorteres variabler med mere end 20% manglende data og de borgere, som har nul indlæggelsesdage.



Figur 3.1: Flowdiagrammet viser, hvor mange variabler (v) der fjernes pga. mere end 20% manglende data samt, hvor mange borgere (n) der fjernes pga. nul indlæggelsesdage.

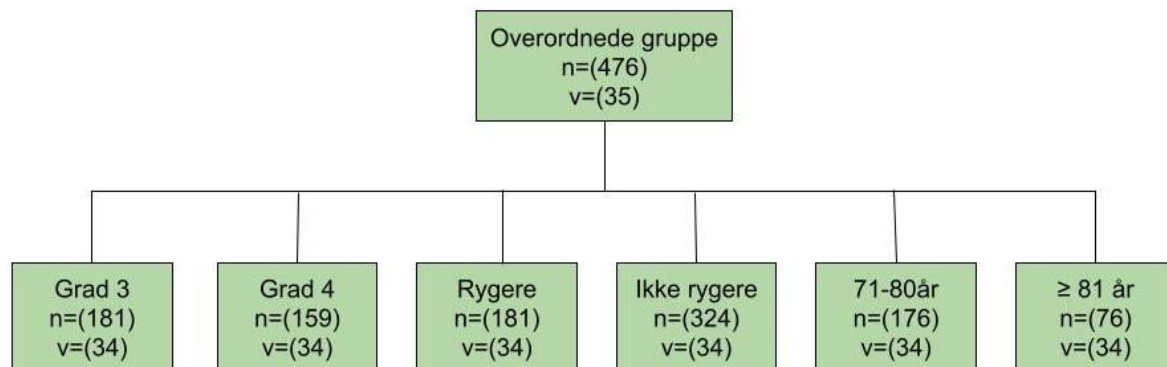
Datasættet indeholder derfor 476 borgere og 35 variabler for "Den overordnede gruppe", som stratificeringen skal foregå ud fra. De 35 variabler består af 23 kontinuerlige variabler, 11 nominale variabler og 1 ordinal variabel, se appendix 5. Nedenstående tabel 3.1 viser en oversigt over, hvilke variabler der ekskluderes på baggrund af mere end 20% manglende data.

Ekskluderede variabler (n=18)	
<i>Variabler</i>	<i>Manglende data (%)</i>
FEV1FVC	52,7%
Antal piber om dagen	39,2%
Antal cigarer om dagen	38,9%
Antal cerutter om dagen	38,8%
Antal cigaretter om dagen	38,8%
Dyspnøscore	38,4%
Rygeår	37,5%
FVC	37,1%
RVCL	33,3%
Ryger dummy	32,4%
Iltbehandling	32,3%
FEV1L	29,6%
BMI	25,5%
SF36-mental score	24,7%
SF36-fysisk score	24,7%
Højde	22,3%
Antal komorbiditeter	21,9%
Uddannelse	21,3%

Tabel 3.1: Oversigt over ekskluderede variabler med mere end 20% manglende data. Alle variabler er målt ved baseline.

3.2 Stratificering af datasættet

De seks stratificerede grupper tager udgangspunkt i datasættet for “Den overordnede gruppe”, hvorved nedenstående figur 3.2 oplyser, hvor mange borgere og variable der indgår i hvert datasæt.



Figur 3.2 Viser en oversigt over, hvor mange variable (v) og borgere (n) der indgår i de syv datasæt.

3.3 Træning af modeller

Der er udviklet prædiktionsalgoritmer på hver af de syv grupperes træningssæt, hvorved der er udviklet 14 prædiktionsalgoritmer i alt. De tre bedste prædiktionsalgoritmer er opnået ved RFR på grupperne “Grad 4”, “Rygere” og “≥ 81 år”. Der fremvises i de nedenstående tabeller 3.2-3.4 hvilke variable de tre bedste prædiktionsalgoritmer er bygget op efter og hvor stor en betydning den enkelte variabel har i prædiktionsalgoritmen for træningssættet for både MLR og RFR. Der vises både resultater for MLR og RFR for at kunne se, hvilke variable der er tilføjet, og hvilken betydning variableerne har for hver af prædiktionsalgoritmerne.

"Grad 4"					
MLR (Forward Selection)		RFR (Predictor Importance)			
Antibiotika	(0,096)	Vægt	(0,17)	År med KOL	(0,07)
Sygepleje tid	(0,151)	Kommunale indsatser	(0,15)	Diastolisk blodtryk	(0,03)
Antal lægebesøg	(0,211)	FEV1	(0,14)	Livskvalitet	(0,02)
FEV1	(0,245)	Antal KOL præparater	(0,12)	Antal indlæggelsesdage	(0,02)
Hypertension	(0,287)	Systolisk blodtryk	(0,11)	Puls	(0,01)
		Antibiotika	(0,08)	Antal lægebesøg	(0,01)
		Alder	(0,07)		

Tabel 3.2: Viser en oversigt over, hvilke variabler der medtages i prædiktionsalgoritmerne baseret på hhv. MLR og RFR på træningssættet for gruppen "Grad 4". Ved RFR er betydningen af den enkelte variabel angivet ved predictor importance, hvor der for MLR er angivet determinationskoefficienten for hele prædiktionsalgoritmen. Alle de medtagne variabler i prædiktionsalgoritmerne er målt ved baseline.

"Rygere"					
MLR (Forward Selection)		RFR (Predictor Importance)			
Antibiotika	(0,059)	Indlæggelsesdage	(0,18)	Diastolisk blodtryk	(0,04)
Vægt	(0,094)	Livskvalitet	(0,13)	Ambulant besøg	(0,03)
Hospitalsindlæggelse	(0,122)	Sygepleje tid	(0,11)	Hjemmepleje tid	(0,03)
		Vægt	(0,08)	Alder	(0,02)
		Lægebesøg	(0,07)	Systolisk blodtryk	(0,02)
		FEV1	(0,07)	Antibiotika	(0,02)
		Træningstid	(0,06)	Hospitalsindlæggelser	(0,01)
		Antal KOL præparater	(0,06)	Kommunale indsatser	(0,01)
		År med KOL	(0,05)	Sværhedsgrad af KOL	(0,01)

Tabel 3.3: Viser en oversigt over, hvilke variabler der medtages i prædiktionsalgoritmerne baseret på hhv. MLR og RFR på træningssættet for gruppen "Rygere". Ved RFR er betydningen af den enkelte variabel angivet ved predictor importance, hvor der for MLR er angivet determinationskoefficienten for hele prædiktionsalgoritmen. Alle de medtagne variabler i prædiktionsalgoritmerne er målt ved baseline.

"Alder ≥ 81 "					
MLR (Forward Selection)		RFR (Predictor Importance)			
Sygepleje tid	(0,471)	Vægt	(0,15)	Hjemmepleje tid	(0,06)
Hypertension	(0,544)	Antibiotika	(0,14)	FEV1	(0,04)
Indlæggelsesdage	(0,595)	Systolisk blodtryk	(0,13)	Sværhedsgrad af KOL	(0,03)
		Hypertension	(0,12)	År med KOL	(0,02)
		Antal KOL præparater	(0,11)	Puls	(0,02)
		Lægebesøg	(0,10)	Diastolisk blodtryk	(0,01)
		Sygepleje tid	(0,07)		

Tabel 3.4: Viser en oversigt over, hvilke variabler der medtages i prædiktionsalgoritmerne baseret på hhv. MLR og RFR på træningssættet for gruppen " ≥ 81 år". Ved RFR er betydningen af den enkelte variabel angivet ved predictor importance, hvor der for MLR er angivet determinationskoefficienten for hele prædiktionsalgoritmen. Alle de medtagne variabler i prædiktionsalgoritmerne er målt ved baseline.

Der henvises til appendix 6 og 7, hvor det ses, hvilke variabler der er tilføjet og hvilken betydning disse har, i de resterende otte prædiktionsalgoritmer baseret på hhv. MLR og RFR. Appendix 6 viser en oversigt over hvilke og hvor mange variabler, der medtages i prædiktionsalgoritmerne samt ligningen for de udviklede prædiktionsalgoritmer baseret på MLR. I appendix 7 ses, hvilke og hvor mange variabler der medtages i prædiktionsalgoritmen baseret på RFR.

3.4 Test af modeller

Nedenstående tabel 3.5 viser en oversigt over de syv grupperes determinationskoefficienter for prædiktionsalgoritmerne. Der vises både determinationskoefficienter for trænings- og testsættet, baseret på hhv. MLR og RFR. Der vil blive markeret med fed i tabellen, hvilke tre der opnår de højeste determinationskoefficienter for tre af prædiktionsalgoritmerne.

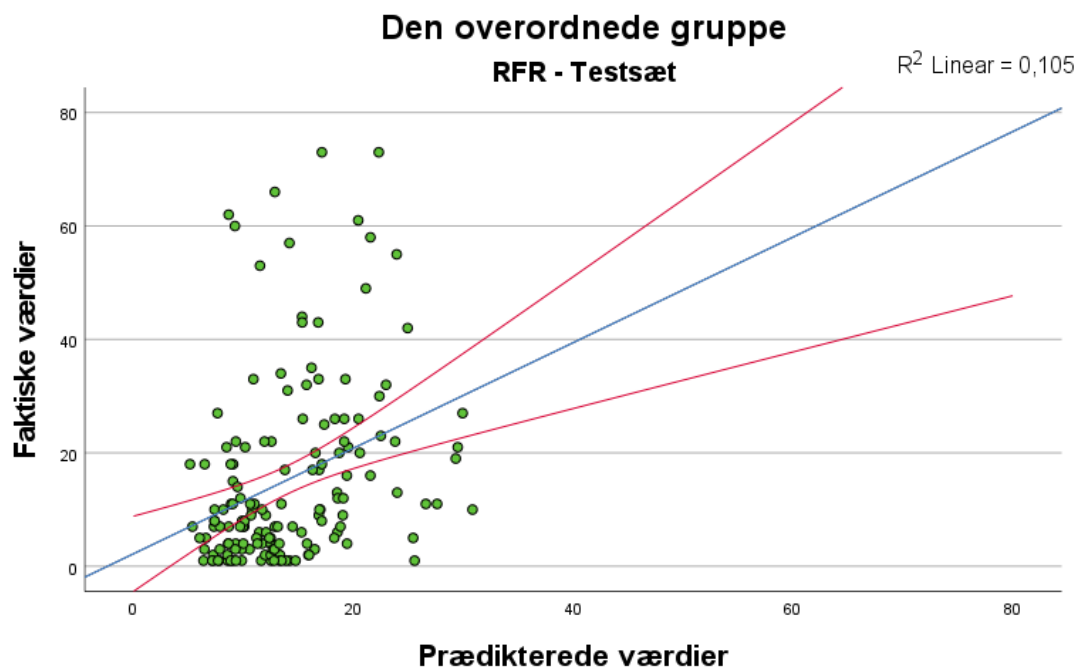
Determinationskoefficienter for prædiktionsalgoritmerne				
	MLR (R^2)		RFR (R^2)	
	Træning (70%)	Test (30%)	Træning (70%)	Test (30%)
Den overordnede gruppe (n= 476)	0,156	0,045	0,692	0,105
Grad 3 (n=181)	0,176	0,018	0,658	0,097
Grad 4 (n=159)	0,287	0,003	0,698	0,184
Rygere (n=152)	0,341	0,059	0,637	0,288
Ikke-rygere (n=324)	0,122	0,026	0,674	0,043
71-80 år (n=176)	0,138	0,206	0,667	0,139
≥ 81 år (n=76)	0,594	0,228	0,554	0,276

Tabel 3.5: Viser de syv grupper hvor der i første kolonne indikeres under hvert gruppenavn, hvor mange borgere der er i hvert datasæt med (n=x). Der er angivet, hvilke determinationskoefficienter grupperne opnår for hhv. MLR og RFR på trænings- og testsættet. I tabellen er de 3 bedste resultater markeret med fed.

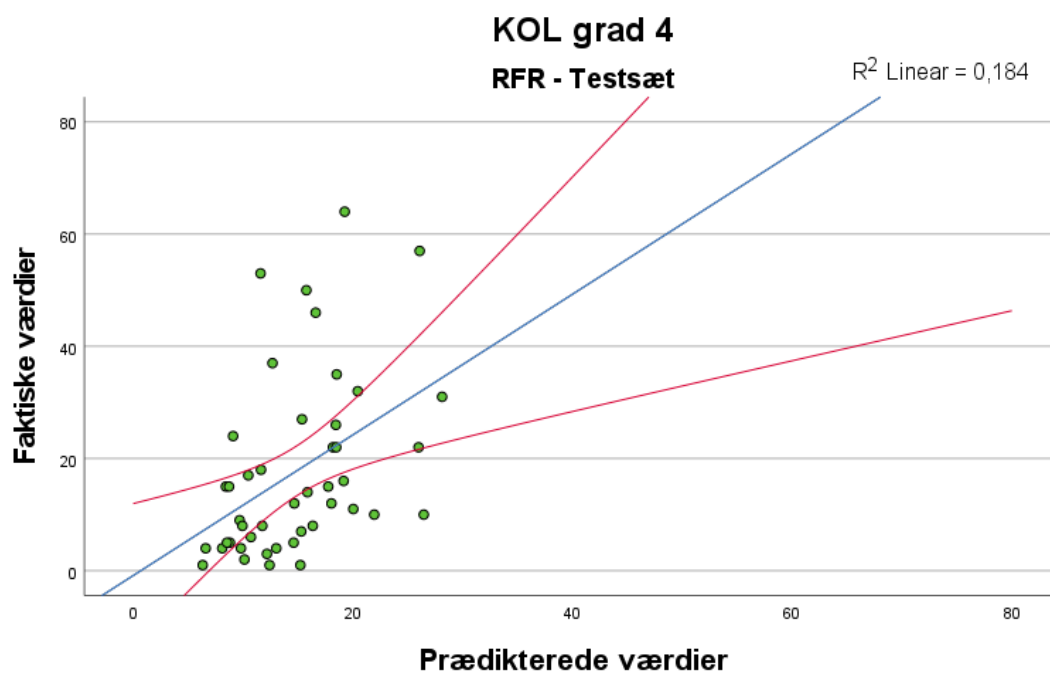
Tabel 3.5 giver en oversigt over, hvilke determinationskoefficienter de syv grupper opnår på både MLR og RFR. De tre bedste determinationskoefficienter på testsættet opnås på prædiktionsalgoritmer, baseret på RFR, for grupperne “Grad 4”, “Rygere” og “≥ 81 år”. Disse tre grupper opnår en determinationskoefficient på hhv. 0,184, 0,288 og 0,278.

3.5 Scatterplots

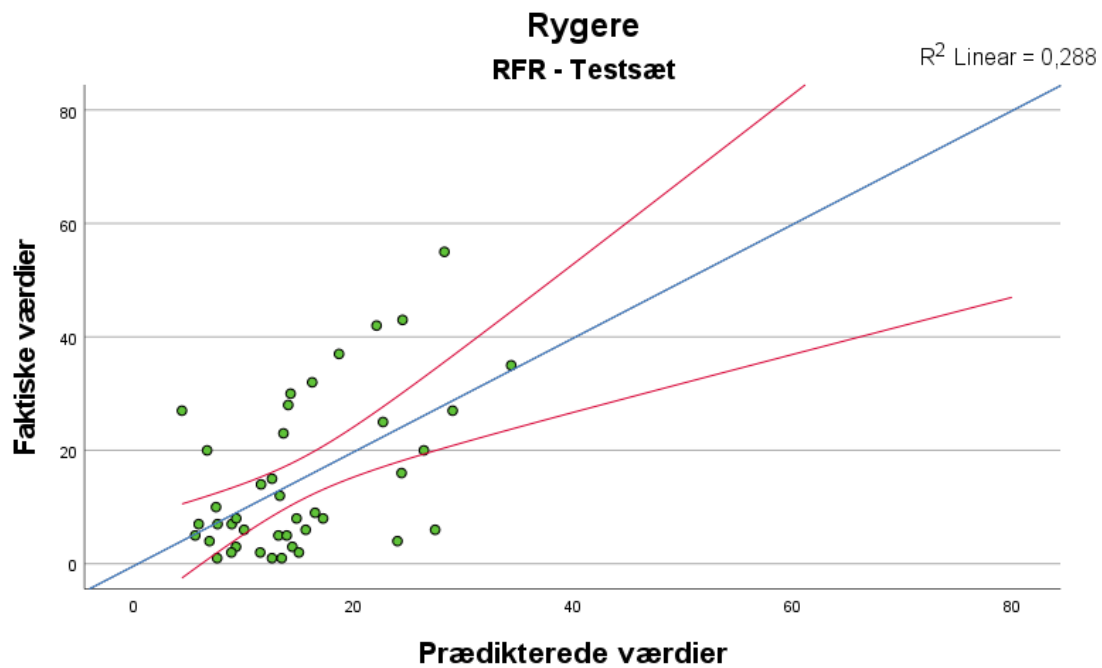
Der fremvises scatterplots over "Den overordnede gruppe" og de tre bedste prædiktionsalgoritmer, som er udviklet på RFR, se figur 3.3-3.6. Scatterplottene viser, hvordan de nævnte prædiktionsalgoritmer har prædikeret på testsættet. De grønne punkter viser forholdet mellem de faktiske værdier på y-aksen og de prædikterede værdier på x-aksen. På alle scatterplots indtegnes en regressionslinje (blå) og konfidensgrænser (røde). Determinationskoefficienten for sammenhængen mellem de faktiske og prædikterede værdier angives i højre hjørne på scatterplottene.



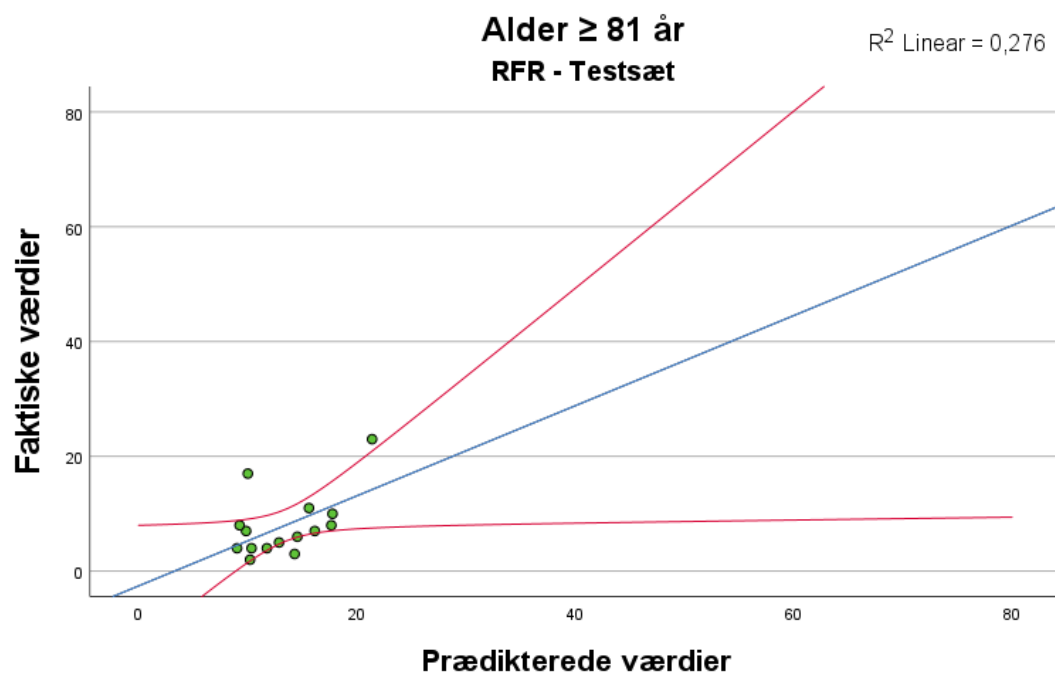
Figur 3.3: Scatterplot over prædiktionsalgoritmen baseret på RFR for "Den overordnede gruppe".



Figur 3.4: Scatterplot over prædiktionsalgoritmen baseret på RFR for "Grad 4".



Figur 3.5: Scatterplot over prædiktionsalgoritmen baseret på RFR for "Rygere".



Figur 3.6: Scatterplot over prædiktionsalgoritmen baseret på RFR for " $\geq$ 81 år".

4. Diskussion

I diskussionsafsnittet opridses specialets formål og resultater, hvor efter afsnittet inddeles i "Prædiktionsalgoritmer", "Stratificering og præprocessering" og "Telemedicin".

"Prædiktionsalgoritmer" tager udgangspunkt i, hvilken prædiktionsmodel der var bedst, samt resultaterne diskuteres med ekstern litteratur. I "Stratificering og præprocessering" vil der blive diskuteret, hvad en anden metode og valg i præprocessering, kunne have medført i specialet. Til sidst belyses dilemmaet omkring kun at give telemedicin til nogle borgere med KOL.

Formålet med dette speciale var at undersøge, hvorvidt en prædiktionsalgoritme kan prædiktere, hvilke borgere der bør tilbydes telemedicin baseret på antal indlæggelsesdage. Resultaterne viser, at prædiktionsalgoritmerne ikke kan anvendes til at prædiktere, hvilke borgere med KOL der er i risiko for flest indlæggelsesdage. Der ses dog tendenser til, at RFR prædikterer bedre end MLR og at prædiktionsalgoritmerne under- og overestimerer i de tre scatterplots, samt opnår de værst tilredte borgere i hver stratificering de bedste determinationskoefficienter.

4.1 Prædiktionsalgoritmer

Gennem en vurdering af determinationskoefficienterne for MLR og RFR, ses der en tendens til, at RFR prædikterer bedre end MLR. Det kan derfor diskuteres, hvorfor prædiktionsalgoritmerne baseret på RFR er bedre til at prædiktere, hvilke borgere der i risiko for flest antal indlæggelsesdage. MLR er simpel og transparent i sin beslutningsmetode, hvorfor det er en hyppigt anvendt prædiktionsmodel (81). Dog viser et studie af Duncan et al. (82), som tester seks forskellige algoritmer inden for prædiktation af sundhedsudgifter, at netop modellens simpelhed kan ligge til grund for dårlige resultater. I studiet opnår prædiktionsalgoritmen baseret på MLR, ligesom i dette speciale, det laveste resultat med en determinationskoefficient på 0,17 (82). Duncan et al. (82) mener, at det kræver et dybere forklaringsmønster, i form af en mere avanceret prædiktionsmodel, at kunne prædiktere det valgte outcome. Dette kan være forklaringen på, hvorfor RFR prædikterer bedre, da denne prædiktionsmodel er mere robust og prædikterer på et mere kompliceret grundlag. Således kan det argumenteres for, at MLR ikke er den rette model til prædiktation af outcomet, hvorfor det kan overvejes, om andre prædiktionsmodeller i tråd med RFR ville have været bedre at teste.

Det vurderes, at andre studier som prædikterer exacerbationer, kan sammenlignes med dette speciale, i forhold til at prædiktere en gruppe, som er i risiko for hospitalsindlæggelser. Det kan derfor diskuteres, hvorfor disse studier opnår bedre prædiktationsevne. Mohktar et al. (40) anvender

en anden form for beslutningstræ, kaldet CART, hvor der måles på et binært outcome modsat dette speciale, der har et kontinuert outcome. Et binært outcome forenkler ofte et svar, hvor den statistiske evne reduceres, da sammenhængen mellem en prædiktør og resultatet ikke kan forklares. Dette skaber usikkerhed omkring de værdier, der ligger tæt på skæringspunktet ved et binært outcome. (84) I specialet ses der usikkerhed ved de nedre og øvre observationer. I specialet ses der at scatterplottene under- og overestimerer i prædiktionen, hvilket indikeres ved konfidensgrænserne, som viser et større interval ved de lave og høje observationer. Dette kan forklares ved, at der er for få observationer på borgere med lave og høje indlæggelsesdage for at kunne prædiktere et kontinuert outcome. Omvendt kritiserer Mohktar et al. (40) sig selv og andre studier for at inddrage et for lille sample, hvor et større sample vil forbedre robustheden i algoritmen (40). Mohktar et al. (40) inddrager kun 21 borgere med KOL i grad 2, 3 og 4, hvor der i specialet inddrages mellem 80-400 borgere i de enkelte datasæt og kun grad 3 og 4. Ved inddragelse af flere borgere i specialet øges prædiktionsalgoritmens robusthed eller prædiktionssevne dog ikke. Det kan formodes at en *node size* som er tilpasset hvert datasæt ville have optimeret prædiktionen. Et studie af Couronné et al. (85) har testet forskellige indstillinger af RFR på 243 forskellige datasæt, hvor der testes *node sizes* på 2, 5, 10 og 20. Studiet konkluderer at de forskellige tilpasninger af RFR har en beskeden betydning for præcisionen af algoritmen og at der ikke kan forventes store forbedringer ved at finjustere parametrene. Dog finder studiet også en tendens til at større *node sizes* fungerer bedre. Det kan derfor overvejes, om der i specialet bør have været anvendt en højere *node size* og om denne burde have været justeret individuelt på de forskellige grupper. Dette underbygges af et studie af Ishwaran et al. (86), der beskriver, hvordan *node size* bør tilpasses den enkelte gruppe, da et større datasæt skal have en større *node size* for at sikre en god prædiktionssevne af algoritmerne.

Den ringe prædiktionssevne kan også påvirkes af en korrelation mellem prædiktionsvariablerne. Det kan derfor overvejes, om der i præprocesseringen skulle have været undersøgt korrelationen mellem variablerne. Der er eksempelvis i de tre bedste prædiktionsalgoritmer baseret på RFR tilføjet "Vægt" som én af de bedste prædiktionsvariabler. En borgers vægt har bl.a. fysiologisk sammenhæng med puls og fysisk aktivitet, hvor overvægt eksempelvis kan indikeres med forhøjet hvilepuls og inaktivitet (87). I prædiktionsalgoritmen for "Grad 4" prioriteres variabelen "Vægt" højere end "Puls" og "Træningstid". I prædiktionsalgoritmerne for "Rygere" og "≥81 år" ses, at variabelen "Vægt" prioriteres højere end "Blodtryk". Her findes der også en fysiologisk sammenhæng med variabelen "Vægt", idet forhøjet blodtryk kan forklares ved overvægt (87). Det kan derfor formodes, at disse variabler evt. har en korrelation med variabel "Vægt", som påvirker prædiktionsalgoritmen.

Dette kan skabe usikkerhed i en prædiktionsalgoritme, hvis en variabel afhænger af en anden variabel, hvorved der ikke vides, hvilken variabel der påvirker outcomet.

I et systematisk review af Sanchez-Morillo et al. (3) vurderes, at nøjagtig prædiktions af exacerbationer ikke er muligt endnu, hvorfor prædiktionsalgoritmerne endnu ikke er anvendelige i praksis. Der skal derfor flere undersøgelser til for at optimere prædiktionsalgoritmerne.

4.2 Stratificering og præprocessering

I specialet tager metoden udgangspunkt i en stratificering af borgere med KOL, for at finde de borgere som skal tilbydes telemedicin. Stratificeringen er valgt ud fra variabler, som vurderes at kunne påvirke antallet af indlæggelsesdage. Der er ingen tidligere studier, som specialets forfattere er bekendt med, som har forsøgt at udvikle en prædiktionsalgoritme der kan udvælge, hvilke borger med KOL der bør tilbydes telemedicin.

Det kan diskuteres, om andre metoder end algoritmer ville være bedre til at udvælge borgere der skal tilbydes telemedicin. På den ene side er fokuset i dette speciale på de borgere, der anvender telemedicin, som er værst tilredt, da der søges borgere i risiko for flest indlæggelsesdage. Dette hænger godt sammen med de kriterier en borger skal opfylde, for at få ordineret telemedicin ud fra Sundhedsstyrelsens rapport 2017 (88). De kriterier, som eksempelvis skal opfyldes, er at være klassificeret i gruppe D og/eller have en FEV1 under 50% (89). Dette er i god tråd med specialets stratificering, som tager udgangspunkt i de værste grader af KOL. På den anden side viser et studie af Kichloo et al. (90), at de borgere, der vurderer deres helbred som værst, ikke anvender telemedicin. Det samme gør sig gældende i et studie af Reed et al. (56) som viser, at det oftest er yngre og dem med flair for teknologi, som vælger en telemedicinsk behandling. Det kan dermed formodes, at nogle af de ældste og mest kritiske borgere med KOL, som er dem, der er forbundet med størst risiko for indlæggelsesdage, fravælger telemedicin. Derfor kunne det overvejes, om en kvalitativ metode og et andet outcome var bedre til at identificere borgere, der skal tilbydes telemedicin. Et studie af Kampmeijer et al. (91) vurderer, at der er mange ældre, som fravælger telemedicin på grund af manglende motivation og støtte fra deres omgangskreds. Dette er bl.a. også et af kriterierne for at give telemedicin, da det er vigtigt for, at behandlingen lykkes (89). Her kunne det overvejes, om der i klinisk praksis skulle lægges fokus på et interview eller indsatser omkring den manglende motivation, støtte eller teknologiske kunnen, som gør, at disse borgere fravælger telemedicin. Et studie af Hunting et al. (92) viser bl.a. en måling af motivation, kunne give et godt indblik i, om telemedicin ville have en virkning. Outcomet kunne også ændres, til eks. livskvalitet eller en motivationsfaktor, som kunne omfavne problemstillingen omkring borgere, der fravælger

telemedicin. Der ses derfor anden relevant metode til at undersøge den gruppe af borgere med KOL, der bør anvende telemedicin, ift. at være forbundet med en samfundsøkonomisk omkostning. En læge kan også vurdere, hvorvidt telemedicin kan have en god effekt på borgeren med KOL (89), hvorfor ordinationen af telemedicin afhænger af, hvad telemedicin skal bidrage med. Det kan derfor diskuteres, om præprocessering af udvælgelsen af borgere med KOL kunne have været bedre, hvis der blev taget udgangspunkt i samme værdier som ved lægens vurdering. Der er i dette speciale valgt borgere, der vurderes som de mest kritiske borgere med KOL, da de er forbundet med flest indlæggelsesdage. Omvendt har et studiet af Mcdowell et al. (8) taget udgangspunkt i borgere med KOL i grad 2, 3 og 4, hvor der ikke opnås en effekt på de samfundsøkonomiske omkostninger forbundet med hospitalsindlæggelser. Dette kan hænge sammen med, at grad 1 og 2 ikke er forbundet med mange hospitalsindlæggelser, hvor der tages højde for det, i dette speciale. Dog kan telemedicin blive tilbudt af forskellige grunde, udover det økonomiske aspekt kan det også være med til at øge *self-management* gennem patientundervisning, grundet demografiske forhold eller for at forebygge exacerbationer. Derfor kunne nogle borgere med KOL i grad 3 tilbydes telemedicin på baggrund af demografiske forhold, i stedet for risikoen for indlæggelsesdage. Gruppen af borgere med KOL er en stor heterogen gruppe, hvilket flere studier indikerer (5,6,54). Det kræver derfor mere forskning for at kunne præcisere hvilke borgere der skal udvælges, for at måle det ønskede formål.

4.3 Telemedicin

Specialet tager udgangspunkt i, hvilke borgere der skal tilbydes telemedicin, hvor fokuset er på de borgere, der er forbundet med flest indlæggelsesdage. Dette undersøges, da der ingen økonomisk effekt ses ved anvendelse af telemedicin til alle borgere med KOL. Det kan diskuteres, om alle borgere med KOL bør tilbydes telemedicin, hvis der ingen økonomisk effekt er. Et perspektiv omkring, hvad den manglende økonomiske effekt kan skyldes, er, at den i tidligere studier er blevet undersøgt for tidligt ift. implementeringen af telemedicin. Ifølge Witt-Udsen et al. (54) kan den manglende effekt på økonomiske omkostninger skyldes den hurtige implementering af telemedicin, hvor sundhedspersonalet først hen mod projektets afslutning blev bedre til at tyde borgernes behov og reaktioner ift. telemonitoreringen. Stoddart et al. (55) beskriver i relation til dette, at tiden brugt på de nye arbejdsopgaver mindskes i takt med, at både borgere og personale bliver bedre til at anvende det nye udstyr og får proceduren på plads. Telemedicin gives dog ikke kun for at opnå en økonomisk effekt, men også hvis borgerens sundhed kan forbedres ved brug af telemedicin. Der kan anvendes telemonitorering pga. demografi, angst eller hvis KOL er den dominerende sygdom ved komorbiditeter (93). Der kan gives telemedicinsk behandling som patientundervisning eller

videokonsultationer ved ambulante konsultationer (89). Et nyere studie af Barbosa et al. (94) viser eksempelvis, at telemedicin kan bidrage med øget *self-management* ved at give patientundervisning og styrke borgerens kendskab til egen sygdom. Telemedicin har samtidig vist sig at være en god behandlingsmulighed ift. situationen med COVID-19. Her viser et studie af Barbosa et al. (94), at telemedicin har været med til at bevare sikkerheden for patienterne og sundhedspersonalet, da man undgår fysisk kontakt og derfor mindsker risikoen for smittespredning. Der kan derfor skabes et etisk dilemma, hvis telemedicin kun skal gives til nogle borgere med KOL, hvilket ikke er hensigten. Specialet tager dog udgangspunkt i et samfundsøkonomisk perspektiv, der viser at det ikke er muligt at opnå en økonomisk effekt ved at give telemedicin til alle borgere, men hvor hensigten er at anvende prædiktionsalgoritmer til at identificere de borgere med risiko for flest indlæggelsesdage.

5. Konklusion

Formålet med dette speciale var at undersøge, hvorvidt det er muligt at anvende prædiktionsalgoritmer til at prædiktere, hvilke borgere der er i risiko for flest indlæggelsesdage og derfor kan have samfundsøkonomiske fordele ved at anvende telemedicin. I dette speciale konkluderes det, at de syv udviklede prædiktionsalgoritmer ikke kunne prædiktere hvilke borgere med KOL, der er i risiko for flest indlæggelsesdage. Ingen af de syv udviklede prædiktionsalgoritmer opnåede et acceptabelt niveau for determinationskoefficienten. Derved kan det også konkluderes, at stratificering af datasættet ikke havde den ønskede effekt, da den ikke kunne bidrage med væsentlige forbedringer af prædiktionsalgoritmernes præstationer.

Resultaterne i specialet viste, at de bedste prædiktioner blev foretaget af RFR algoritmen, og de grupper, hvor de bedste prædiktioner blev foretaget, var borgere med KOL i grad 4, rygere og aldersgruppen over 81 år. Der ses derfor en tendens til, at det er borgere med KOL i kritisk tilstand, der bedst kan anvendes til at prædiktere et økonomisk outcome.

For at kunne målrette indsatsen af telemedicin, kræves yderligere undersøgelser inden for, hvilke borgere der er i risiko for flest indlæggelsesdage, end hvis der tages udgangspunkt i alle borgere med KOL. Det vurderes, at den anvendte stratificering af borgere med KOL ikke var en løsning til at finde de borgere med KOL, som havde flest indlæggelsesdage. Resultaterne viste dog tendenser til, at valget af prædiktionsmodel har en betydning for prædiktionsevnen, hvor RFR blev vurderet bedst. Befolkningsgruppen med KOL er en stor heterogen gruppe, hvor fremtidige studier bør være opmærksomme på, hvilke borgere der skal måles på for at finde en økonomisk effekt.

6. Perspektivering

Der blev i specialet prædikeret ud fra stratificerede grupper, hvilket er efterspurgt i tidligere forskning (4–6,54). På trods af dette var resultaterne i specialet ikke acceptable, hvorfor der kræves yderligere forbedringer, såfremt en prædiktionsalgoritme skal kunne anvendes i klinisk praksis. Fremtidige studier bør derfor overveje, hvilke borgere med KOL der inddrages i en gruppe, afhængigt af studiets formål. Der kunne eksempelvis være fokus på om gruppen skal stratificeres ud fra flere variabler eller prøve en anden metode for at finde de borgere med KOL, der har flest indlæggelsesdage. Her er der i specialet foreslået en kvalitativ metode, som retter fokus på de ældre kritiske syge, som fravælger telemedicin.

Der kan undersøges andre prædiktionsmodeller, der ville være relevante at anvende mhp. at tilstræbe en bedre prædiktionssevne. Et studie har bl.a. opnået acceptable resultater ved en *Support Vector Model* og *Neural Network* (95). En prædiktionsalgoritme, som kan prædiktere en gruppe af borgere med KOL i risiko for flest indlæggelsesdage, kan både gavne den enkelte borger og den praktiserende læge. Samtidig kan den praktiserende læge lettere udvælge på baggrund af resultatet fra prædiktionsalgoritmen, hvilke borgere der har gavn af telemedicin. Disse fordele gør det relevant at undersøge udviklingen af prædiktionsalgoritmer yderligere for at kunne vurdere, om der i fremtiden kan opnås en samfundsøkonomisk gevinst ved at give telemedicin til en bestemt gruppe af borgere med KOL.

7. Referenceliste - Vancouver

1. Houmøller Rasmussen H. National udbredelse Telemedicin KOL [Internet]. Region Nordjylland. 2020 [henvist 16. februar 2021]. Tilgængelig hos: <https://rn.dk/sundhed/til-sundhedsfaglige-og-samarbejdspartnere/telecare-nord/national-udbredelse-telemedicin-kol>
2. Lilholt PH, Witt Udsen F, Ehlers L, Hejlesen OK. Telehealthcare for patients suffering from chronic obstructive pulmonary disease: effects on health-related quality of life: results from the Danish "TeleCare North" cluster-randomised trial. *BMJ Open*. 9. maj 2017;7(5).
3. Sanchez-Morillo D, Fernandez-Granero MA, Leon-Jimenez A. Use of predictive algorithms in-home monitoring of chronic obstructive pulmonary disease and asthma: A systematic review. *Chron Respir Dis*. august 2016;13(3):264–83.
4. Riis HC, Jensen MH, Cichosz SL, Hejlesen OK. Prediction of exacerbation onset in chronic obstructive pulmonary disease patients. *Journal of Medical Engineering & Technology*. 2. januar 2016;40(1):1–7.
5. Achelrod D, Schreyögg J, Stargardt T. Health-economic evaluation of home telemonitoring for COPD in Germany: evidence from a large population-based cohort. *Eur J Health Econ*. september 2017;18(7):869–82.
6. Haesum LKE, Soerensen N, Dinesen B, Nielsen C, Grann O, Hejlesen O, m.fl. Cost-Utility Analysis of a Telerehabilitation Program: A Case Study of COPD Patients. *Telemedicine and e-Health*. november 2012;18(9):688–92.
7. Decramer M, Janssens W, Miravittles M. Chronic obstructive pulmonary disease. *The Lancet*. april 2012;379(9823):1341–51.
8. McDowell JE, McClean S, FitzGibbon F, Tate S. A randomised clinical trial of the effectiveness of home-based health care with telemonitoring in patients with COPD. *J Telemed Telecare*. marts 2015;21(2):80–7.
9. World Health Organization. Chronic obstructive pulmonary disease (COPD) [Internet]. WHO; 2017. Tilgængelig hos: [https://www.who.int/news-room/fact-sheets/detail/chronic-obstructive-pulmonary-disease-\(copd\)](https://www.who.int/news-room/fact-sheets/detail/chronic-obstructive-pulmonary-disease-(copd))
10. Bilde L, Rud Svenning A, Dollerup J, Baekke Borgeskov H, Lange P. The cost of treating patients with COPD in Denmark—a population study of COPD patients compared with non-COPD controls. *Respir Med*. marts 2007;101(3):539–46.
11. Lange P, Tøttenborg SS, Sorknæs AD, Andersen JS, Søgaaard M, Nielsen H, m.fl. Danish Register of chronic obstructive pulmonary disease. *Clin Epidemiol*. 2016;8:673–8.
12. Chen C-Z, Shih C-Y, Hsiue T-R, Tsai S-H, Liao X-M, Yu C-H, m.fl. Life expectancy (LE) and loss-of-LE for patients with chronic obstructive pulmonary disease. *Respiratory Medicine*. oktober 2020;172:106132.
13. van Hirtum PV, Sprooten RTM, van Noord JA, van Vliet M, de Kruif MD. Long term survival after admission for COPD exacerbation: A comparison with the general population. *Respiratory Medicine*. april 2018;137:77–82.
14. Anecchino C, Rossi E, Fanizza C, De Rosa M, Tognoni G, Romero M, m.fl. Prevalence of chronic obstructive pulmonary disease and pattern of comorbidities in a general population. *Int J Chron Obstruct Pulmon Dis*. 2007;2(4):567–74.

15. Klintefelt A. Befolkningsfremskrivninger 2019-2060. Danmarks Statistik. 2019;2.
16. Klintefelt A, Harbo L. Befolkningsfremskrivninger 2020-2060. Danmarks Statistik. 2020;3.
17. Agarwal AK, Raja A, Brown BD. Chronic Obstructive Pulmonary Disease. I: StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing; 2020 [henvist 28. december 2020]. Tilgængelig hos: <http://www.ncbi.nlm.nih.gov/books/NBK559281/>
18. Hatipoğlu U. Chronic obstructive pulmonary disease: More than meets the eye. Ann Thorac Med. marts 2018;13(1):1–6.
19. Decker R, Wisconsin F, Hadfield R, Hess M. Global strategy for the diagnosis management, and prevention of chronic obstructive pulmonary disease - 2019 report [Internet]. Spanien: GOLD; 2019 [henvist 16. februar 2021] s. 155. Tilgængelig hos: <https://goldcopd.org/wp-content/uploads/2018/11/GOLD-2019-v1.7-FINAL-14Nov2018-WMS.pdf>
20. Lange P, Nielsen LP, Plaschke P. Diagnostik af astma og KOL [Internet]. 2020 [henvist 23. februar 2021]. Tilgængelig hos: <https://pro.medicin.dk/Sygdomme/Sygdom/318292>
21. Viborg AL. Respirationsorganerne. I: Sygdomslære - Hånden på hjertet. 2. kbh.: Munksgaard; 2015. s. 111–45.
22. Nielsen OF, Bojsen-Møller M. Respiration. I: Anatomi & Fysiologi - Hånden på hjertet. 1. udg. Kbh.: MUNKSGAARD; 2015. s. 121–31.
23. Hillas G, Papaporfyriou A, Dimakou K, Papaioannou AI. Pharmacological treatment of stable COPD: need for a simplified approach. Postgraduate Medicine. 17. februar 2020;132(2):126–31.
24. Løkke A, Hilberg O, Tønnesen P, Ibsen R, Kjellberg J, Jennum P. Direct and indirect economic and health consequences of COPD in Denmark: A national register-based study: 1998-2010. BMJ Open. 6. januar 2014;4(1):e004069.
25. Anees ur Rehman, Ahmad Hassali MA, Muhammad SA, Shah S, Abbas S, Hyder Ali IAB, m.fl. The economic burden of chronic obstructive pulmonary disease (COPD) in the USA, Europe, and Asia: results from a systematic review of the literature. Expert Review of Pharmacoeconomics & Outcomes Research. 1. november 2020;20(6):661–72.
26. Rehman A ur, Hassali MAA, Muhammad SA, Harun SN, Shah S, Abbas S. The economic burden of chronic obstructive pulmonary disease (COPD) in Europe: results from a systematic review of the literature. Eur J Health Econ. marts 2020;21(2):181–94.
27. Andersson F, Borg S, Jansson S-A, Jonsson A-C, Ericsson Å, Prütz C, m.fl. The costs of exacerbations in chronic obstructive pulmonary disease (COPD). Respiratory Medicine. september 2002;96(9):700–8.
28. File:Human Lungs.gif - Wikipedia [Internet]. [henvist 25. maj 2021]. Tilgængelig hos: https://commons.wikimedia.org/wiki/File:Human_Lungs.gif
29. Nielsen OB. Respirationssystemet. I: Fysiologi og anatomi - Det levende menneske. 1. udgave, 2. oplag. Kbh.: Munksgaard Danmark; 2010. s. 269–98.
30. Laursen LC. Lungesygdomme. I: Basisbog i Sygdomslære. 2. udgave, 3. oplag. Kbh: Munksgaard Danmark; 2011. s. 271–9.
31. Bruno CM, Valenti M. Acid-Base Disorders in Patients with Chronic Obstructive Pulmonary Disease: A Pathophysiological Review. Journal of Biomedicine and Biotechnology. 2012;2012:1–8.

32. Cruz J, Marques A, Figueiredo D. Impacts of COPD on family carers and supportive interventions: a narrative review. *Health Soc Care Community*. januar 2017;25(1):11–25.
33. Miravittles M, Ribera A. Understanding the impact of symptoms on the burden of COPD. *Respir Res*. december 2017;18(1):67.
34. Sigurgeirsdottir J, Halldorsdottir S, Arnardottir RH, Gudmundsson G, Bjornsson EH. COPD patients' experiences, self-reported needs, and needs-driven strategies to cope with self-management. *COPD*. maj 2019;Volume 14:1033–43.
35. Borné Y, Ashraf W, Zaigham S, Frantz S. Socioeconomic circumstances and incidence of chronic obstructive pulmonary disease (COPD) in an urban population in Sweden. *COPD: Journal of Chronic Obstructive Pulmonary Disease*. 2. januar 2019;16(1):51–7.
36. Disler RT, Gallagher RD, Davidson PM. Factors influencing self-management in chronic obstructive pulmonary disease: An integrative review. *International Journal of Nursing Studies*. februar 2012;49(2):230–42.
37. Burge S, Wedzicha JA. COPD exacerbations: definitions and classifications. *European Respiratory Journal*. 1. juni 2003;21(Supplement 41):46S – 53s.
38. Anzueto A. Impact of exacerbations on COPD. *European Respiratory Review*. 1. juni 2010;19(116):113–8.
39. Toy EL, Gallagher KF, Stanley EL, Swensen AR, Duh MS. The Economic Impact of Exacerbations of Chronic Obstructive Pulmonary Disease and Exacerbation Definition: A Review. *COPD: Journal of Chronic Obstructive Pulmonary Disease*. maj 2010;7(3):214–28.
40. Mohktar MS, Redmond SJ, Antoniadis NC, Rochford PD, Pretto JJ, Basilakis J, m.fl. Predicting the risk of exacerbation in patients with chronic obstructive pulmonary disease using home telehealth measurement data. *Artificial Intelligence in Medicine*. januar 2015;63(1):51–9.
41. Chandra D, Tsai C-L, Camargo CA. Acute Exacerbations of COPD: Delay in Presentation and the Risk of Hospitalization. *COPD: Journal of Chronic Obstructive Pulmonary Disease*. januar 2009;6(2):95–103.
42. Wilkinson TMA, Donaldson GC, Hurst JR, Seemungal TAR, Wedzicha JA. Early Therapy Improves Outcomes of Exacerbations of Chronic Obstructive Pulmonary Disease. *Am J Respir Crit Care Med*. 15. juni 2004;169(12):1298–303.
43. Rinne ST, Graves MC, Bastian LA, Lindenauer PK, Wong ES, Hebert PL, m.fl. Association between length of stay and readmission for COPD. *Am J Manag Care*. 1. august 2017;23(8):253–8.
44. Xu W, Collet JP, Shapiro S, Lin Y, Yang T, Wang C, m.fl. Negative impacts of unreported COPD exacerbations on health-related quality of life at 1 year. *European Respiratory Journal*. 1. maj 2010;35(5):1022–30.
45. Jones PW, Lamarca R, Chuecos F, Singh D, Agustí A, Bateman ED, m.fl. Characterisation and impact of reported and unreported exacerbations: results from ATTAIN. *Eur Respir J*. november 2014;44(5):1156–65.
46. Seemungal TAR, Donaldson GC, Paul EA, Bestall JC, Jeffries DJ, Wedzicha JA. Effect of Exacerbation on Quality of Life in Patients with Chronic Obstructive Pulmonary Disease. *Am J Respir Crit Care Med*. maj 1998;157(5):1418–22.
47. Pavord I, Jones P, Burgel PR, Rabe KF, Rabe KF. Exacerbations of COPD. *COPD*. februar 2016;21.

48. Bilde L, Rud Svenning A, DSI - Institut for Sundhedsvæsen. Omkostninger ved behandling af patienter med kronisk obstruktiv lungesygdom (KOL). København: DSI Institut for Sundhedsvæsen; 2004.
49. Sundhedsstyrelsen. Brugen af sundhedsvæsenet for borgere med kronisk sygdom [Internet]. Kbh.; 2019 dec [henvist 1. april 2021] s. 18. Tilgængelig hos: <https://sundhedsdatastyrelsen.dk/da/tal-og-analyser/analyser-og-rapporter/sygdomme/kronisk-sygdom#:~:text='Brugen%20af%20sundhedsv%C3%A6senet%20for%20borgere,stadig%20hyppigere%20kontakt%20med%20sundhedsv%C3%A6senet.>
50. Region Nordjylland. Til patienter og pårørende - KOL [Internet]. Region Nordjylland. 2020 [henvist 9. marts 2021]. Tilgængelig hos: <https://rn.dk/sundhed/til-sundhedsfaglige-og-samarbejdspartnere/telecare-nord/telemedicin-kol/patienter-og-paaroerende>
51. Sood S, Mbarika V, Jugoo S, Dookhy R, Doarn CR, Prakash N, m.fl. What Is Telemedicine? A Collection of 104 Peer-Reviewed Perspectives and Theoretical Underpinnings. *Telemedicine and e-Health*. oktober 2007;13(5):573–90.
52. Murphy LA, Harrington P, Taylor SJ, Teljeur C, Smith SM, Pinnock H, m.fl. Clinical-effectiveness of self-management interventions in chronic obstructive pulmonary disease: An overview of reviews. *Chron Respir Dis*. august 2017;14(3):276–88.
53. McLean S, Nurmatov U, Liu JL, Pagliari C, Car J, Sheikh A. Telehealthcare for chronic obstructive pulmonary disease: Cochrane Review and meta-analysis. *Br J Gen Pract*. november 2012;62(604):739–49.
54. Witt Udsen F, Lilholt PH, Hejlesen O, Ehlers L. Cost-effectiveness of telehealthcare to patients with chronic obstructive pulmonary disease: results from the Danish 'TeleCare North' cluster-randomised trial. *BMJ Open* [Internet]. maj 2017 [henvist 4. februar 2021];7(5). Tilgængelig hos: <https://bmjopen.bmj.com/lookup/doi/10.1136/bmjopen-2016-014616>
55. Stoddart A, van der Pol M, Pinnock H, Hanley J, McCloughan L, Todd A, m.fl. Telemonitoring for chronic obstructive pulmonary disease: a cost and cost-utility analysis of a randomised controlled trial. *J Telemed Telecare*. marts 2015;21(2):108–18.
56. Reed ME, Huang J, Graetz I, Lee C, Muelly E, Kennedy C, m.fl. Patient Characteristics Associated With Choosing a Telemedicine Visit vs Office Visit With the Same Primary Care Clinicians. *JAMA Netw Open*. 17. juni 2020;3(6):e205873.
57. Vianello A, Fusello M, Gubian L, Rinaldo C, Dario C, Concas A, m.fl. Home telemonitoring for patients with acute exacerbation of chronic obstructive pulmonary disease: a randomized controlled trial. *BMC Pulm Med*. december 2016;16(1):157.
58. Ogunleye J. The Concepts of Predictive Analytics. *International journal of knowledge, Innovation and Entrepreneurship*. 2014;2(2):9.
59. Grant SW, Collins GS, Nashef SAM. Statistical Primer: developing and validating a risk prediction model†. *European Journal of Cardio-Thoracic Surgery*. 1. august 2018;54(2):203–8.
60. Shah SA, Velardo C, Farmer A, Tarassenko L. Exacerbations in Chronic Obstructive Pulmonary Disease: Identification and Prediction Using a Digital Health System. *J Med Internet Res* [Internet]. 7. marts 2017 [henvist 22. december 2020];19(3). Tilgængelig hos: <http://www.jmir.org/2017/3/e69/>
61. Fernández-Granero MA, Sánchez-Morillo D, León-Jiménez A, Crespo LF. Automatic prediction of chronic obstructive pulmonary disease exacerbations through home telemonitoring of

- symptoms. *Biomed Mater Eng.* 2014;24(6):3825–32.
62. van der Heijden M, Lucas PJF, Lijnse B, Heijdra YF, Schermer TRJ. An autonomous mobile system for the management of COPD. *J Biomed Inform.* juni 2013;46(3):458–69.
63. Kronborg T, Mark L, Cichosz SL, Secher PH, Hejlesen O. Population exacerbation incidence contains predictive information of acute exacerbations in patients with chronic obstructive pulmonary disease in telecare. *Int J Med Inform.* marts 2018;111:72–6.
64. Kronborg T, Hangaard S, Cichosz SL, Hejlesen O. A two-layer probabilistic model to predict COPD exacerbations for patients in telehealth. *Comput Biol Med.* januar 2021;128:104108.
65. Houmøller Rasmussen H, Thaarup Høegh R, Mylund L, Adamsen. Om TeleCare Nord [Internet]. Region Nordjylland. 2021 [henvist 31. marts 2021]. Tilgængelig hos: <https://rn.dk/sundhed/til-sundhedsfaglige-og-samarbejdspartnere/telecare-nord/om-telecare-nord>
66. Dorte Stigaard, Bente Graversen. Telecare Nord Afslutningsrapport - telemedicinsk storskalaundersøgelse i Nordjylland [Internet]. Nordjylland: Nordjyske Kommuner, Regionssygehede, PLO Nordjylland, Lægernes kvalitetsorganisation NordKAP, Lungeforening og Aalborg Universitet; 2015 jun [henvist 22. december 2020] s. 56. Tilgængelig hos: https://digst.dk/media/11275/telecare_nord_afslutningsrapport_18112015_final_-1.pdf
67. Roberts MB, Sullivan MC, Winchester SB. Examining solutions to missing data in longitudinal nursing research: ROBERTS et al. *J Spec Pediatr Nurs* [Internet]. april 2017 [henvist 13. april 2021];22(2). Tilgængelig hos: <http://doi.wiley.com/10.1111/jspn.12179>
68. IBM. IBM SPSS Missing Values 22 [Internet]. IBM; [henvist 31. marts 2021]. Tilgængelig hos: http://www.sussex.ac.uk/its/pdfs/SPSS_Missing_Values_22.pdf
69. Krzywinski M, Altman N. Multiple linear regression. *Nat Methods.* december 2015;12(12):1103–4.
70. Aggarwal N, Vass AA, Minardi HA, Ward R, Garfield C, Cybyk B. People with dementia and their relatives: personal experiences of Alzheimer's and of the provision of care. *JPsychiatrMentHealth Nurs.* 2003;10(2):187–97.
71. Meyer-Baese A, Schmid V. Feature Selection and Extraction. I: *Pattern Recognition and Signal Analysis in Medical Imaging* [Internet]. Elsevier; 2014 [henvist 17. marts 2021]. s. 21–69. Tilgængelig hos: <https://linkinghub.elsevier.com/retrieve/pii/B9780124095458000029>
72. Walczak B, Massart DL. Calibration in Wavelet Domain. I: *Data Handling in Science and Technology* [Internet]. Elsevier; 2000 [henvist 17. marts 2021]. s. 323–49. Tilgængelig hos: <https://linkinghub.elsevier.com/retrieve/pii/S0922348700800404>
73. Duncan IG. Healthcare risk adjustment and predictive modeling. Winsted, Conn: ACTEX Publications; 2011. 338 s.
74. Zhao J, Gu S, McDermaid A. Predicting outcomes of chronic kidney disease from EMR data based on Random Forest Regression. *Mathematical Biosciences.* april 2019;310:24–30.
75. Heinze G, Wallisch C, Dunkler D. Variable selection - A review and recommendations for the practicing statistician. *Biom J.* maj 2018;60(3):431–49.
76. PennState Eberly College of Science. Best Subsets Regression, Adjusted R-Sq, Mallows Cp | STAT 462 [Internet]. Applied Regression Analysis. 2018 [henvist 7. april 2021]. Tilgængelig hos: <https://online.stat.psu.edu/stat462/node/197/>
77. Liaw A, Wiener M. Classification and Regression by RandomForest. *Forest.* 30. november

2001;23.

78. Breiman L. Random Forest. Machine Learning. 2001;45(1):5–32.
79. Denisko D, Hoffman MM. Classification and interaction in random forests. Proc Natl Acad Sci U S A. 20. februar 2018;115(8):1690–2.
80. SPSS Modeler. IBM Docs [Internet]. 2014 [henvist 14. april 2021]. Tilgængelig hos: www.ibm.com/docs/en/spss-modeler/18.1.0
81. Ditlevsen S, Sørensen H. Sandsynlighedsregning og statistik(SS) [Internet]. Kbh.: Københavns Universitet; 2018 [henvist 14. april 2021] s. 164. Report No.: 5. Tilgængelig hos: ISBN 978-87-7078-882-3
82. Duncan I, Loginov M, Ludkovski M. Testing Alternative Regression Frameworks for Predictive Modeling of Health Care Costs. North American Actuarial Journal. 2. januar 2016;20(1):65–87.
83. Uddin S, Khan A, Hossain ME, Moni MA. Comparing different supervised machine learning algorithms for disease prediction. BMC Med Inform Decis Mak. 21. december 2019;19(1):281.
84. Altman DG, Royston P. The cost of dichotomising continuous variables. BMJ. 6. maj 2006;332(7549):1080.1.
85. Couronné R, Probst P, Boulesteix A-L. Random forest versus logistic regression: a large-scale benchmark experiment. BMC Bioinformatics. 17. juli 2018;19(1):270.
86. Ishwaran H, Malley JD. Synthetic learning machines. BioData Min. 2014;7(1):28.
87. Falkenberg Nielsen O, Bojsen-Møller MJ. Anatomi og fysiologi. Kbh.: Munksgaard; 2012.
88. Sundhedsstyrelsen. Telemedicinsk understøttelse af behandlingstilbud til mennesker med KOL - Anbefalinger for målgruppe, sundhedsfagligt indhold samt ansvar og samarbejde. Kbh.; 2017. Report No.: 1.1.
89. TeleCare Nord KOL - sundhed.dk [Internet]. [henvist 27. maj 2021]. Tilgængelig hos: <https://www.sundhed.dk/sundhedsfaglig/information-til-praksis/nordjylland/almen-praksis/patientforloeb/forloebbeskrivelser/r-luftveje/telecarenord/>
90. Kichloo A, Albosta M, Dettloff K, Wani F, El-Amir Z, Singh J, m.fl. Telemedicine, the current COVID-19 pandemic and the future: a narrative review and perspectives moving forward in the USA. Fam Med Com Health. august 2020;8(3):e000530.
91. Kampmeijer R, Pavlova M, Tambor M, Golinowska S, Groot W. The use of e-health and m-health tools in health promotion and primary prevention among older adults: a systematic literature review. BMC Health Serv Res. august 2016;16(S5):290.
92. Hunting G, Shahid N, Sahakyan Y, Fan I, Moneyppenny CR, Stanimirovic A, m.fl. A multi-level qualitative analysis of Telehomecare in Ontario: challenges and opportunities. BMC Health Serv Res. juni 2015;15(1):544.
93. Ordination af hjemmemonitorering [Internet]. [henvist 28. maj 2021]. Tilgængelig hos: https://www.rm.dk/sundhed/faginfo/center-for-telemedicin/projekter-og-indsatser/telemedicin-til-kol/udbredelse-i-midtjylland/projektspor/Sundhedsfagligt_indhold_og_organisering/Sundhedsfaglig_indsats/ordination_af_hjemmemonitorering/
94. Barbosa MT, Sousa CS, Morais-Almeida M, Simões MJ, Mendes P. Telemedicine in COPD: An Overview by Topics. COPD: Journal of Chronic Obstructive Pulmonary Disease. 2. september

2020;17(5):601–17.

95. Iadanza E, Mudura V, Melillo P, Gherardelli M. An automatic system supporting clinical decision for chronic obstructive pulmonary disease. *Health Technol. marts* 2020;10(2):487–98.