

Dokumentet er delvist forældet.

Grundet ændringer i projektets design er dele af dette dokument forældet.

Det er ikke længere ambitionen at gennemføre afprøvning af Beslutningsstøtten på faktiske sager i dette projekt. Dermed er enhver beskrivelse af Pilottest 2 og Lodtrækningsforsøg forældet.

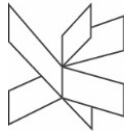
Baggrunden for ændringerne er beskrevet i "Notat om ændringer i projekt Underretninger i Fokus", som er tilgængelig på Trygfondens Børneforskningscenters hjemmeside <https://childresearch.au.dk/udsatte-boern-unge-og-familier/projekter/underretninger-i-fokus> og UC viden <https://www.ucviden.dk/da/projects/underretninger-i-fokus>.

Projektleder
Line Berg

Juni 2022



TRYGFONDENS
BØRNEFORSKNINGS-CENTER
INSTITUT FOR ØKONOMI
AARHUS UNIVERSITET



VIA University
College

TrygFonden

Egenskaber ved den statistiske model i forskningsprojektet *Underretninger i fokus*

Første udgave: 10. marts 2021*

**Dette dokument opdateres ikke længere. Modeludviklingen dokumenteres mere detaljeret i en videnskabelig artikel, som er under udarbejdelse, og som vil blive fagfællebedømt. Når den videnskabelige artikel er udgivet, vil vi indsætte en henvisning til den på projektets hjemmesider.*

Michael Rosholm
TrygFondens Børneforsknings-
center og Institut for Økonomi,
Aarhus Universitet
Email: rom@econ.au.dk

Simon Bodilsen
TrygFondens Børneforsknings-
center og Institut for Økonomi,
Aarhus Universitet

Sanne Dalgaard Toft
TrygFondens Børneforsknings-
center, Aarhus Universitet



INDHOLDSFORTEGNELSE

1	Introduktion.....	3
1.1	Baggrund.....	5
2	Beslutningsstøtten.....	6
2.1	Data.....	6
2.2	Valg af outcome.....	8
2.3	Den statistiske model.....	9
3	Modellens egenskaber.....	10
3.1	Forudsigelse af anbringelser.....	10
3.2	Ekstern validering af modellen.....	13
4	Diskussion.....	16
5	Konklusion.....	16
5.1	Analyser under udarbejdelse.....	17
6	Referencer.....	18

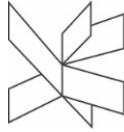
1 Introduktion

Udsatte børn og unges trivsel har de senere år opnået stor politisk og offentlig bevågenhed. Det kan man blandt andet se med Barnets Reform i 2011, Overgrebspakken i 2013 og senest Regeringens annoncering af Barnets Lov i januar 2021. Denne øgede bevågenhed har blandt betydet en markant stigning i antallet af underretninger med bekymringer vedrørende børn og unges trivsel og udvikling.

For socialrådgiverne er det imidlertid ikke en let opgave at identificere, hvilke børn, der er i størst risiko for at komme til at mistrives, og flere studier peger på, at der er plads til forbedringer, hvad angår de beslutninger, som træffes i de kommunale forvaltninger (se fx Bartelink et al, 2015; Drake m.fl., 2020; Taylor, 2021). En af udfordringerne er, at underretningerne ofte indeholder få konkrete oplysninger, og at det derfor kan være vanskeligt at vurdere barnets reelle risiko for mistrivsel. En anden udfordring er, at der ikke er klar evidens for eller enighed om, hvordan forskellige risiko- og beskyttelsesfaktorer skal vægtes i forhold til hinanden bl.a. med baggrund i, at børns udvikling ikke kan forudsiges med stor præcision (Villumsen, 2015; Rutter, 2012a, b). Disse udfordringer overlader naturligvis en høj grad af skøn til socialrådgiverne, hvilket både øger risikoen for fejl, men også risikoen for inkonsistens i de beslutninger, der træffes på tværs af sager og på tværs af socialrådgivere. I en dansk kontekst viser Ejrnæs m.fl. (2020) i et vignet-studie, at fagprofessionelle vurderer risikoen ved forskellige risikofaktorer vidt forskelligt. Studiet omhandler ikke underretninger om børn, men afspejler problematikken mere generelt. Internationale studier viser, at denne inkonsistens ikke kun er en dansk udfordring (Arad-Davidzon og Benbenishty, 2008; Arad-Davidzon og Benbenishty, 2010; Benbenishty m.fl., 2015; Enosh og Bayer-Topilsky, 2015; Rossi et al, 1999; Schuerman m.fl., 1999).

I forskningsprojektet Underretninger i Fokus undersøger vi, hvorvidt det er muligt at etablere et beslutningsunderstøttende redskab, som, ved at trække på data kommunerne allerede indsamler, kan informere socialrådgiverne om risikoen for, at barnet eller den unge mistrives i en grad, så det kan begrunde offentlig intervention, når socialrådgiverne sidder med en konkret underretning.

Redskabet er et statistisk baseret værktøj, som kan vurdere risikoen for mistrivsel blandt børn og unge. Dette redskab kalder vi Beslutningsstøtten. Det er vores håb, at vi med Beslutningsstøtten kan være med til at kvalificere det grundlag, hvorpå socialrådgiverne træffer deres beslutning, ved at tilbyde dem en mere standardiseret information på tværs af underretningssager og socialrådgivere. Det er således hverken planen eller forventningen, at Beslutningsstøtten vil reducere socialrådgivernes tidsforbrug i forbindelse med den



enkelte beslutning, ligesom Beslutningsstøtten heller ikke kommer med eller kobles til konkrete handleanvisninger. Håbet er derimod, at det vil føre til et bedre, mere sikkert og mere ensartet grundlag for de beslutninger, som træffes.

Grundene til at overveje at indføre en sådan model er mange, se for eksempel Søbjerg m.fl. (2020); 1) den kan bidrage til at reducere tilfældig variabilitet i de fagprofessionelles afgørelser, 2) den kan lede til bedre afgørelser (øget præcision), og 3) endelig kan den bidrage til at analysere og reducere eventuel bias i de fagprofessionelles afgørelser. Coulthard m.fl. (2020) argumenterer for, at sådanne modeller derfor udgør en etisk forbedring – blandt andet i relation til emner som transparens, naturligvis i det omfang datagrundlaget bag risikovurderingen præsenteres. Taylor (2021) argumenterer som konsekvens heraf for, at statistiske beslutningsunderstøttende værktøjer bør anvendes komplementært med de fagprofessionelles øvrige beslutningsgrundlag, når der træffes beslutninger i socialt arbejde.

Der er naturligvis også en lang række grunde til at være kritisk i forbindelse med at indføre statistiske redskaber som grundlag for beslutninger foretaget af fagprofessionelle. Disse diskuteres grundigt af O'Neil (2016). Her males et billede af et dystopisk samfund, hvor beslutninger vedrørende ansættelse, lånoptagelse, forsikringspriser mv. i stigende grad foretages af ikke-transparente algoritmer. Det medfører en lang række risici som for eksempel 1) risikoen for at videreføre og forstærke eksisterende biases og uligheder, hvilket kan føre til yderligere institutionalisering af diskrimination mod bestemte grupper, 2) ikke-transparens, 3) manglende konfidentialitet, og 4) tab af autonomi/medbestemmelse/den menneskelige faktor.

Det er derfor også vigtigt for os at påpege, at modellen er udviklet som led i et forskningsprojekt, som netop skal belyse, hvorvidt en model som Beslutningsstøtten kan bidrage til at forbedre de beslutninger, som træffes - uden samtidig at introducere en lang række andre problemer.

I denne artikel beskriver vi den prædiktive risikomodel, der ligger bag Beslutningsstøtten. Modellen er i skrivende stund fortsat under udarbejdelse, og denne artikel præsenterer således alene vores foreløbige arbejde. Artiklen vil derfor løbende blive opdateret i takt med, at arbejdet skrider frem. Formålet med artiklen er at skabe åbenhed omkring arbejdet med modeludviklingen, at præsentere modellens egenskaber og samtidig forsøge at adressere nogle af de udfordringer, der er i udviklingen af en prædiktiv risikomodel i forhold til bias/fairness/diskrimination.

Modeludviklingen dokumenteres mere detaljeret i en videnskabelig artikel, som er under udarbejdelse, og som vil blive fagfællebedømt. Når den videnskabelige artikel er udgivet, vil vi indsætte en henvisning til den her.



Du kan læse mere om forskningsprojektet og projektets øvrige elementer, herunder vores etiske overvejelser, projektets juridiske grundlag m.v. på [projektsiden hos TrygFondens Børneforskningscenter](#) eller på [UC Viden](#). Denne artikel bør læses inden for konteksten af det samlede forskningsprojekt inklusive de juridiske analyser og vurderinger samt de etiske overvejelser.

Artiklen er struktureret som følger: I resten af introduktionen beskriver vi baggrunden for forskningsprojektet. I afsnit 2 beskriver vi arbejdet med modellen, herunder datagrundlag, valg af outcome og den statistiske model. I afsnit 3 præsenterer vi modellens egenskaber, mens analyser af bias/fairness/diskrimination diskuteres i afsnit 4. Artiklen afsluttes med en beskrivelse af det fremtidige arbejde med udviklingen af den prædiktive risikomodell.

1.1 Baggrund

I 2018 modtog kommunerne mere end 120.000 underretninger. Dette tal har i en årrække været stigende. Bare tre år tidligere, i 2015, modtog kommunerne til sammenligning omkring 93.000 underretninger. I denne periode har der således været en stigning på cirka 30 procent – og det ser kun ud til, at denne udvikling fortsætter.

Når en kommune modtager en underretning, skal den inden for 24 timer vurdere, hvorvidt ”*barnets eller den unges sundhed eller udvikling er i fare, og om der derfor er behov for at iværksætte akutte foranstaltninger...*” (Serviceloven §155, stk. 2). Hvis kommunen vurderer, at der *ikke* er behov for at iværksætte akutte foranstaltninger, skal den vurdere, om der er behov for at igangsætte en tidlig forebyggende indsats (Serviceloven §11, stk. 3), eller om det må antages, at et barn eller en ung har behov for særlig støtte, og at der derfor skal igangsættes en børnefaglig undersøgelse (Serviceloven §50, stk. 1). Undersøgelsen skal afdække barnets, familiens og netværkets ressourcer og problemer (ibid., stk. 4) og skal danne grundlag for en beslutning om, hvorvidt der skal iværksættes foranstaltninger, og i så fald af hvilken type (ibid., stk. 6). Den børnefaglige undersøgelse skal afsluttes senest 4 måneder efter, at kommunen bliver opmærksom på, at et barn eller en ung kan have behov for særlig støtte (ibid., stk. 7).

Der er stor forskel på de underretninger kommunerne modtager. Underretningerne kan for eksempel komme fra forskellige kilder; herunder blandt andet fra barnets dagtilbud eller skole, tandlægen, politiet, barnets familie eller netværk. Den kan blive indgivet per e-mail, ved brug af en standardiseret formular på kommunens hjemmeside, et telefonopkald mv. Der er naturligvis også stor variation, når det gælder, hvad bekymringen i underretningen omhandler. Det kan for eksempel handle om, at barnet har et højt skolefravær, at



barnet er udeblevet fra en tandlægetid, at barnet er involveret i kriminalitet, en bekymring for omsorgssvigt, vanskelige opvækstbetingelser m.m. Endelig omhandler underretningerne også meget forskellige børn og unge, nogle med og nogle uden eksisterende sager i den kommunale forvaltning. Der er altså stor forskel i karakteren af de underretninger, kommunerne modtager og behandler, og samlet set betyder det, at kommuner på kort tid skal træffe nogle svære og samtidig vigtige beslutninger, som har betydning for, hvorvidt udsatte børn og unge får den hjælp, de har brug for.

Alle underretninger skal vurderes på lige fod, men de fleste, som arbejder med vurdering af underretninger ved, at det er svært at vurdere, hvilke børn der har brug for særlig støtte, og hvilke der ikke har. Enten fordi underretningen indeholder meget få konkrete informationer, og at det derfor er svært at læse ud fra underretningen, hvad bekymringen skyldes, eller fordi der i forvejen ligger en meget stor journal med en verserende sag (i hvilket tilfælde der kan være tale om en genvurdering), og at det derfor er svært at afgøre, hvorvidt der er grundlag for at justere det aktuelle bekymringsniveau og de aktuelle foranstaltninger.

Problemet med det stigende antal underretninger er ikke alene, at de lægger et stort pres på de kommunale forvaltninger, men også at det bliver sværere at adskille signaler fra støj. Så jo flere underretninger kommunen modtager om børn og unge med en lav risiko for at mistrives, des større er risikoen for, at den kommer til at forstyrre familier, som ikke har brug for det offentlige hjælp og indblanding. Omvendt stiger risikoen også for at overse børn og unge, som har brug for hjælp.

Det stigende antal underretninger har således gjort det endnu vigtigere at lede efter nye måder og metoder til at identificere, hvilke børn og unge, der er i risiko for mistrivsel, og hvilke der ikke er.

2 Beslutningsstøtten

2.1 Data

Udviklingen af Beslutningsstøtten er baseret på en række administrative registre i Danmarks Statistik, som er blevet stillet os til rådighed til dette forskningsprojekt. Det væsentligste register er underretningsregisteret, BUU (Børn og unge – Underretninger), som indeholder information om alle underretninger modtaget i de 98 danske kommuner i perioden fra og med april 2014 til og med december 2018. Det er kommunerne, der indsender informationen til Danmarks Statistik. Dette register indeholder bl.a. information om, hvem underretningen vedrører, datoen for modtagelsen af underretningen, samt hvem der står bag underretningen



og årsagen til bekymring. Dette betyder, at vi har et godt overblik over, hvilke børn, der har været involveret i en underretningssag i denne knap firårige periode.

I udviklingen af den endelige model betragtes kun de underretninger, som er modtaget i perioden fra og med april 2016 til og med december 2017. Ved kun at analysere disse underretninger, er det muligt at danne en toårig underretningshistorik for alle sagerne. Denne historik er potentielt vigtig i forhold til at forudsige risikoen for mistriivsel. Ydermere muliggør denne restriktion også, at vi kan kigge et år frem i tiden for alle sager, således at vi kan se, hvilke handlinger der blev – eller ikke blev - iværksat fra myndighedernes side det efterfølgende år som følge af underretningerne. I dette endelige datasæt analyseres således 173.044 underretninger vedrørende 90.644 forskellige børn.

Data hos Danmarks Statistik er pseudonymiserede, hvilket indebærer, at hver observation i BUU-registeret indeholder et unikt nummer, der gør det muligt at koble dette register med andre registre, der indeholder væsentlig information om det underrettede barn og dets familie. Det er vigtigt at påpege, at dette nummer ikke er barnets CPR nummer (som slet ikke kan tilgås af forskere under Danmarks Statistiks forskerordning).

For det barn, underretningen vedrører, inkluderer vi data om barnets alder, flytnehistorik, antal søskende og antallet af tidligere underretninger, forebyggende foranstaltninger og anbringelser inden for de seneste to år.

For forældrene til det barn, underretningen vedrører, inddrages data om deres alder ved barnets fødsel, civilstand, antal partnere de har børn med, samt en indikator for om de er afgået ved døden før underretningstidspunktet. Sluttelig indeholder modellen information om antallet af tidligere underretninger, forebyggende foranstaltninger og anbringelser inden for de seneste to år for det underrettede barns søskende. Data vedrørende underretninger, forebyggende foranstaltninger og anbringelser omfatter kun data for den kommune, barnet er bosat i på tidspunktet for den aktuelle underretning.

I denne forbindelse er det essentielt at bemærke, at de oplysninger som vi sammenkobler med BUU-registeret, *udelukkende* er information, som socialrådgiverne allerede har adgang til i deres systemer. Det betyder med andre ord, at den prædiktive risikomodell ikke inddrager viden, som socialrådgiverne ikke har til



rådighed. Modellen har således ikke flere informationer til rådighed, end socialrådgiverne har. Det modellen derimod kan tilbyde, er en standardisering af, hvorledes den tilgængelige information bliver anvendt til at vurdere, hvorvidt barnet eller den unge har behov for særlig støtte fra kommunen.¹

Ydermere er det vigtigt at bemærke, at vi ikke anvender information vedrørende barnets køn eller etniske oprindelse i modellen. Ved at udelade disse informationer kan vi bidrage til at sikre, at disse to faktorer aldrig direkte vil være årsagen til en given risikovurdering fra værktøjets side. Det er et væsentligt aspekt i udviklingen af datadrevne værktøjer at sørge for, at de ikke favoriserer bestemte befolkningsgrupper. Dette aspekt behandles nærmere i afsnit 4. Der korrigeres også for store aldersskævheder. Metoden hertil og konsekvenserne heraf beskrives mere indgående i afsnit 3.1.

2.2 Valg af outcome

Som nævnt er formålet med dette forskningsprojekt at udvikle et statistisk funderet værktøj, som kan vurdere risikoen for mistrivsel blandt børn og unge. Dette ville konceptuelt være relativt enkelt at udvikle, hvis mistrivsel var et direkte målbart karakteristikum. Dette er dog ikke tilfældet. Mistrivsel er en kompleks størrelse, som ikke umiddelbart kan reduceres til en enkelt observerbar faktor.² For at omgå denne problemstilling, anvendes en såkaldt *proxyvariabel* for mistrivsel. En proxyvariabel er en observerbar variabel, som kan approksimere en anden - uobserverbar - variabel i den forstand, at de to variable forventes at være højt korrelerede. I denne kontekst har vi søgt efter en observerbar variabel, hvormed det gælder, at denne har en høj association til mistrivsel.

Valget af mistrivselproxy, og dermed valget af outcome til den statistiske model, er faldet på en binær indikator for, hvorvidt det barn, en given underretning vedrører, er blevet anbragt uden for hjemmet inden for en periode på et år efter modtagelsen af underretningen. Motivationen for at anvende anbringelse som outcome er baseret på det rationale, at anbringelser kun sker i de tilfælde, hvor det vurderes, at et barns velbefindende er i høj fare. Dermed vil vi forvente, at en model, som forudsiger fremtidige anbringelser,

¹ I den eksterne validering af modellen indgår en række yderligere informationer. Disse ekstra data anvendes dog udelukkende til at vurdere modellens egenskaber og anvendes således ikke i Beslutningsstøtten.

² Danneskiold-Samsøe m.fl. (2020) diskuterer og analyserer muligheden for at anvende et allerede valideret redskab som fx SDQ (Strengths and Difficulties Questionnaire) som direkte mål for et barns trivsel eller mistrivsel. I den videnskabelige artikel eksperimenterer vi også med, hvorvidt vi ved hjælp af faktoranalyse kan identificere en latent faktor relateret til mistrivsel. Når den videnskabelige artikel er udgivet, vil vi indsætte en henvisning til den her.

også vil kunne bruges til at skelne mellem, hvilke børn der er i henholdsvis lav og høj risiko for mistrivsel. Denne hypotese bliver afprøvet i en ekstern validering af modellen, som kort præsenteres i afsnit 3.2.³

2.3 Den statistiske model

I forbindelse med udviklingen af Beslutningsstøtten har vi afprøvet forskellige prædiktive modeller fra klassen af statistiske modeller, der kan klassificeres som *maskinlæringsmodeller*. Fælles for alle de afprøvede modeltyper er, at de beregner sandsynligheden for, at det barn, en underretning vedrører, bliver anbragt i løbet af de næste 365 dage efter underretningen på baggrund af de tilgængelige data på tidspunktet for modtagelsen af underretningen. I udviklingsfasen har vi testet følgende fire modeltyper: (1) en lineær sandsynlighedsmodel, (2) en LASSO regulariseret logistisk regressionsmodel, (3) en random forest model og (4) en extreme gradient boosting (XGBoost) model.

Baseret på disse analyser konkluderes det, at XGBoost-modellen i den konkrete anvendelse giver den bedste prædiction (se afsnit 3.1). XGBoost-modellen er derfor valgt til at danne grundlag for algoritmen bag Beslutningsstøtten.

XGBoost er en metode baseret på *beslutningstræer*. Fremfor at bruge et enkelt beslutningstræ til at lave en modelprædiction, anvendes en række afhængige beslutningstræer, hvor det nuværende beslutningstræ forsøger at rette op på de prædiktionsfejl, som det blev lavet af det forrige beslutningstræ. Algoritmen starter typisk med en meget svag prædiktionsmodel, som i de efterfølgende iterationer af algoritmen bliver konverteret til en stærk prædiktionsmodel, ved sekventielt at opdatere modellen, så prædiktionsfejlene bliver minimeret. Dette er princippet bag *boosting*. For flere detaljer vedrørende teknikken henvises fx til Chen og Guestrin (2016). For at kunne implementere XGBoost-modellen, er det i praksis nødvendigt at specificere en lang række parametre, før træningen af algoritmen kan udføres. Måden, hvorpå vi har tacklet denne problemstilling, er ved at afprøve en lang række af forskellige parameterkonfigurationer på et træningsdatasæt, for derefter at vælge den konfiguration, der giver de bedste forudsigelser på et andet testdatasæt. Trænings- og testdatasæt er valgt ved lodtrækning.

XGBoost metodens gode egenskaber kan i denne sammenhæng forstås ved, at denne algoritme er specielt god til at modellere komplekse ikke-lineære relationer i data. Da en beslutning om en anbringelse uden for

³ Vi har også eksperimenteret med andre proxy-variable, som for eksempel modtagelse af flere alvorlige underretninger, implementering af forebyggende foranstaltninger, men den valgte proxyvariabel klarer sig bedre end disse mål – og kombinationer heraf – i den eksterne validering.

hjemmet typisk skyldes et sammenspil mellem mange forskellige faktorer, er det ikke overraskende, at XGBoost-modellens forudsigelser viser sig at være de mest præcise.

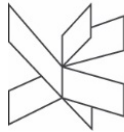
3 Modellens egenskaber

I dette afsnit analyserer vi XGBoost-modellens prædiktive egenskaber. I afsnit 3.1 betragtes modellens evne til at prædiktere outcome – anbringelse uden for hjemmet inden for et år – mens vi i afsnit 3.2 foretager en ekstern validering af modellen. I denne validering viser vi, at der også er en stærk sammenhæng mellem modellens forudsigelser og en række andre mål, der indikerer mistrivsel.

3.1 Forudsigelse af anbringelser

Som nævnt anvendes alle underretninger fra og med april 2016 til og med december 2017 vedrørende 0-17-årige børn og unge i udviklingen af modellen. I maskinlæringslitteraturen er det almindeligt praksis at opdele data i henholdsvis et træningsdatasæt og et testdatasæt. Vi anvender et tilfældigt 70-30 split, hvilket betyder, at 70% af de oprindelige underretninger tilfældigt udvælges og anvendes til at udvikle ('træne') algoritmen, mens de resterende 30% bruges til at evaluere modellens prædiktive egenskaber 'out-of-sample'. I træningsdatasættet indgår der 120.395 underretninger (svarende til 63.303 forskellige børn og unge), og i testdatasættet er de resterende 52.649 underretninger (svarende til 27.341 forskellige børn og unge). Det er værd at bemærke, at splittet er foretaget således, at børn med samme mor ikke kan optræde i begge datasæt. Dette gør, at der er mindre risiko for, at vi opnår kunstigt gode modelegenskaber (Chouldechova m.fl., 2018).

XGBoost-algoritmen giver en prædikteret sandsynlighed for, at det barn, en underretning vedrører, bliver anbragt i løbet af det næste år regnet fra underretningsdatoen. Til at evaluere modellens egenskaber anvendes et mål, som kaldes *Area Under the receiver operating characteristic Curve* (i det følgende blot AUC). AUC varierer mellem 50-100%, hvor højere værdier indikerer bedre modelegenskaber (den nedre grænse opstår, idet man ved ren tilfældighed vil ramme rigtigt i 50% af tilfældene). Mere formelt kan en AUC værdi fortolkes som sandsynligheden for – hvis vi tilfældigt udtrækker to børn, et som anbringes og et som ikke anbringes – at modellen tildeler en højere anbringelsessandsynlighed til det barn, som anbringes end til det barn, som ikke anbringes. Hvis AUC scoren er 100%, vil det betyde, at modellen er i stand til at



skelne perfekt mellem, hvilke underretninger, der fører til anbringelser, og hvilke der ikke gør. AUC værdier over 90% betragtes som 'excellent', værdier mellem 80-90% som gode, værdier mellem 70-80% som rimelige, og værdier under 70% som dårlige.

I Tabel 1 vises AUC-scoren for algoritmen bag Beslutningsstøtten. Modellen opnår en AUC på 83,95%, hvilket betyder, at modellen med en sandsynlighed på knap 84% vil vurdere et barn, der faktisk bliver anbragt til at være i højere risiko end et barn, der ikke bliver anbragt. I et lignende studie fra New Zealand opnåede man en AUC på 76% (Vaithianathan m.fl., 2013).

METODE	AUC (%)	95% KONFIDENSINTERVAL
Outcome: Anbringelse uden for hjemmet inden for et år		
XGBOOST	83,95	83,32 – 84,59

***Tabel 1:** I denne tabel angives AUC-værdien for XGBoost-modellen. Modellen er udviklet på baggrund af 120.395 underretninger (svarende til 63.303 forskellige børn) modtaget mellem april 2016 og december 2017. Modellernes prædiktive egenskaber er evalueret på baggrund af 52.649 underretninger (svarende til 27.341 forskellige børn) modtaget mellem april 2016 og december 2017.*

Når modellen anvendes i praksis af socialrådgivere på konkrete sager, er det hensigten, at Beslutningsstøtten skal tildele en risikoscore, der varierer mellem 1-10. Her angiver en risikoscore på 1 lavest mistrivsel, mens en score på 10 vil angive den højeste risiko for mistrivsel. Måden, hvorpå disse risikoscore dannes, er ved at transformere modellens prædikterede anbringelsessandsynligheder. For at omregne fra sandsynligheder til risikoscores har vi anvendt decil-afgrænsningerne i vores testdatasæt.

Da der historisk set i Danmark har været en stærk positiv sammenhæng mellem barnets alder og en beslutning om anbringelse, risikerer vi, at modellen kommer til at afspejle denne sammenhæng, selvom det vi er interesseret i at måle er mistrivsel, som forventes at være mindre aldersafhængig.

Vi har derfor valgt at gøre risikoscoren helt aldersneutral. Dette er gjort ved at anvende aldersspecifikke decil-afgrænsninger. Risikoscoren udregnes således ved at betragte decil-afgrænsningerne af anbringelsessandsynlighederne inden for hvert alderstrin. Eksempelvis vil alle underretninger blandt tiårige, hvis prædikterede sandsynlighed ligger under den første decil-grænse for tiårige, blive tildelt en risikoscore på 1, og på tilsvarende vis vil en underretning, der tilhører den tiende decil-gruppe blandt tiårige, få tildelt den højeste risikovurdering på 10. Ved at bruge aldersspecifikke decil-afgrænsninger sikres, at risikoscoren ikke bliver aldersafhængig. Der vil, med denne nye metode, nemlig være den samme andel, der får tildelt en



given risikoscore inden for hver aldersgruppe. Det er naturligvis vigtigt, når socialrådgiverne skal undervises i at anvende Beslutningsstøtten, at de informeres om, at risikoscoren er dannet, så den ikke påvirkes af barnets alder.

Som eksempel på, hvordan vi omregner fra anbringelsessandsynligheder til risikoscores, vises nedenfor omregningsmetoden fra prædikteret sandsynlighed til risikoscore for tiårige børn.⁴

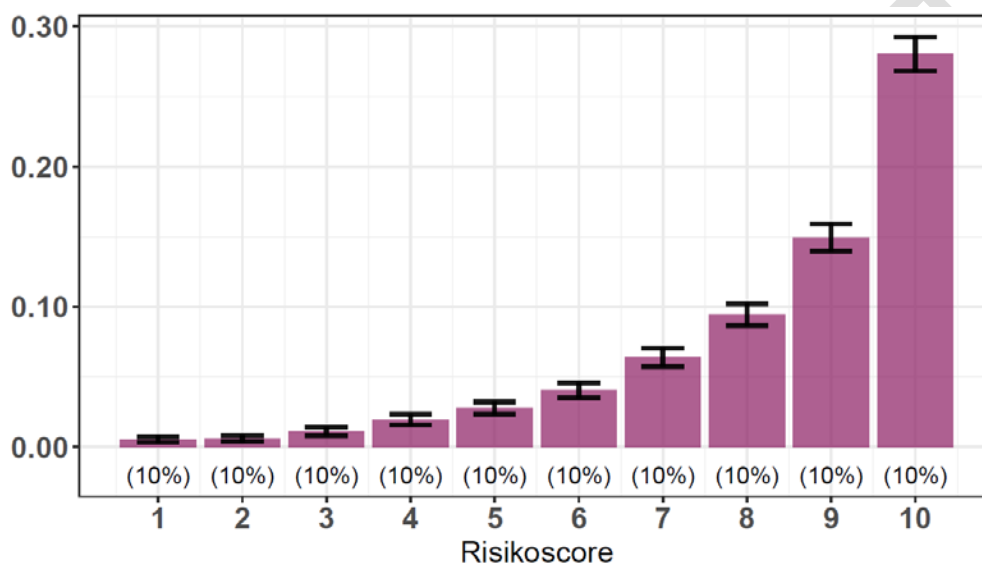
Score = 1, hvis den prædikterede anbringelsessandsynlighed tilhører intervallet [0; 0.00624]
Score = 2, hvis den prædikterede anbringelsessandsynlighed tilhører intervallet]0.00624; 0.00817]
Score = 3, hvis den prædikterede anbringelsessandsynlighed tilhører intervallet]0.00817; 0.01104]
Score = 4, hvis den prædikterede anbringelsessandsynlighed tilhører intervallet]0.01104; 0.01405]
Score = 5, hvis den prædikterede anbringelsessandsynlighed tilhører intervallet]0.01405; 0.01977]
Score = 6, hvis den prædikterede anbringelsessandsynlighed tilhører intervallet]0.01977; 0.02674]
Score = 7, hvis den prædikterede anbringelsessandsynlighed tilhører intervallet]0.02674; 0.03906]
Score = 8, hvis den prædikterede anbringelsessandsynlighed tilhører intervallet]0.03906; 0.05952]
Score = 9, hvis den prædikterede anbringelsessandsynlighed tilhører intervallet]0.05952; 0.10454]
Score = 10, hvis den prædikterede anbringelsessandsynlighed tilhører intervallet]0.10454;1]

Dette indebærer, at en underretning vedrørende et tiårigt barn, hvor den beregnede anbringelsesrisiko overstiger 10,45%, tildeles den højeste risikoscore på 10. Man kan undre sig over, at en anbringelsesrisiko på ca. 10% skal give anledning til så høj en risikovurdering, men det skyldes, at hensigten med Beslutningsstøtten ikke er at forudsige anbringelser, men at forudsige risikoen for, at barnet mistrives. I afsnit 3.2 viser vi, hvordan den således konstruerede risikoscore relaterer sig til andre mål for mistrivsel.

For de 52.649 underretninger i testdatasættet, har vi udregnet deres risikoscore baseret på algoritmen udviklet på træningsdatasættet. I Figur 1 illustreres sammenhængen mellem risikoscores og de faktiske anbringelsesrater for de børn og unge, der indgår i testdatasættet. Som det fremgår af figuren, er der en klar monoton sammenhæng mellem modellens prædiktion og de faktiske hændelser. Det vil sige, at når den

⁴ De øvrige cut-off værdier der er benyttet til omregning fra prædikterede sandsynligheder til risikoscore for andre aldersgrupper, kan rekvireres ved henvendelse til førsteforfatteren.

observerede anbringelsesrate stiger, så ser vi også en stigning i risikoscoren. I de tilfælde, hvor modellen har vurderet højest risiko, er der næsten 30%, der faktisk blev anbragt inden for det efterfølgende år. I den modsatte ende af risikoskalaen ser vi, at der næsten ikke fandt nogle anbringelser sted i de tilfælde, hvor modellen har tildelt en lav risikoscore.⁵

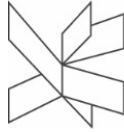


Figur 1: I denne figur vises sammenhængen mellem de modelberegnete risikoscores og de faktiske anbringelsesrater for de børn og unge, som indgår i testdatasættet.

3.2 Ekstern validering af modellen

I det foregående afsnit har vi vist, at den udviklede model har gode prædiktive egenskaber. Hensigten med Beslutningsstøtten er imidlertid, som allerede nævnt, ikke at forudsige anbringelser, men at forudsige det enkelte barn/unges relative risiko for at mistrives. Det er derfor vigtigt at undersøge, om modellens forudsigelser også har en sammenhæng med andre – eksterne - mål for mistrivsel. Det er i den forbindelse vigtigt, at 1) de anvendte mål tidsmæssigt ligger senere end underretningen, og 2) de anvendte mål ikke indgår i udviklingen af algoritmen.

⁵ Bemærk at denne figur ville højere anbringelsesrater i 10. decil og lavere anbringelsesrater i de lavere deciler, hvis den var baseret direkte på de ikke-alderskorrigerede prædikterede sandsynligheder. Denne kosmetiske forbedring ville imidlertid afspejle aldersafhængigheden i anbringelsesraterne, som vi med den anvendte transformation antager er urelaterede til mistrivsel.

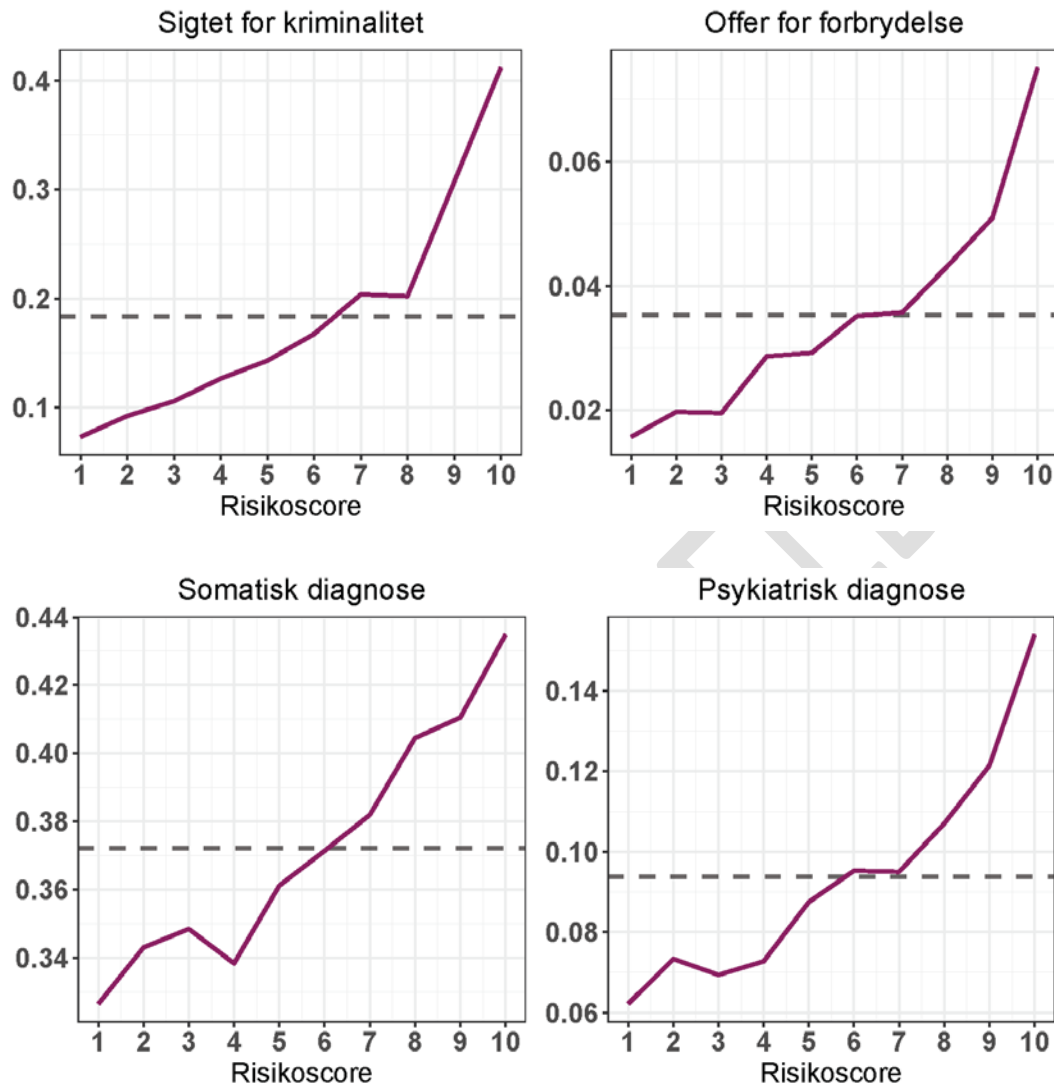


Vi har undersøgt denne meget væsentlige problemstilling ved at se på sammenhængen mellem risikoscores for børnene i testdatasættet, og en række trivselsrelaterede udfald, der er målt i løbet af 2018. Ved at tage udgangspunkt i året 2018, er vi sikre på, at der ikke er overlap mellem den information, som algoritmen kender til og de faktiske hændelser, da modellen som bekendt er udviklet på baggrund af data fra og med april 2016 til og med december 2017. I alt har vi undersøgt sammenhængen mellem risikoscoren og 20 andre observerbare variable, der kan indikere forskellige aspekter af mistrivsel. Dette er grundigt belyst i den videnskabelige artikel, som vil blive offentliggjort, når analyserne er helt afsluttede, og den har været fagfællebedømt. I denne korte fremstilling har vi udvalgt fire af disse variable: (i) anholdelser, (ii) offer for forbrydelse, (iii) somatiske diagnoser, og (iv) psykiatriske diagnoser.⁶ Disse er indikative for de resterende resultater.

I figur 2 er resultaterne af denne eksterne validering illustreret. Figuren viser, at selvom modellen er bygget ved at forudsige fremtidige anbringelser, så er der en meget tydelig og stærk positiv sammenhæng mellem de modelberegne risikoscores og de eksterne mistrivselsvariable. Eksempelvis blev 40% af de 14-17 årige, som blev tildelt en risikoscore på 10, anholdt i løbet af 2018, mens dette tal var under 10% for de med en score på 1. I tilfældet, hvor vi betragter hvorvidt et barn/ung har været offer for en forbrydelse, ses der også en firdobling af denne andel fra den lave ende af risikokalaen (knap 2%) til den høje ende (knap 8%). For nye somatiske diagnoser stiger andelen fra knap 33% for børn med risikoscoren 1 til godt 43% for børn med risikoscoren 10 (kigger vi på specifikke diagnoser ses en stigning fra 13% til 20% i frakturer). For psykiatriske diagnoser er stigningen endnu stærkere, idet prævalensen mere end fordobles fra risikoscore 1 (ca. 6%) til 10 (ca. 15%). Det er vigtigt igen at understrege, at disse variable ikke har haft nogen indflydelse på den beregnede risikoscore. Vi afprøver først efterfølgende, hvorledes samspillet er mellem modellens risikovurdering og disse variable.

Det er endvidere værd at bemærke fra figuren, at mange af sammenhængene mellem risikoscores og trivselsmål er konvekse i den forstand, at de stiger stejlere ved de højere risikoscores. Det indebærer, at forøgelsen af mistrivsel er større, når risikoscoren stiger fra 9 til 10, end når den stiger fra fx 1 til 2 eller fra 5 til 6.

⁶ Analyseresultaterne for de øvrige variable kan rekvireres ved henvendelse til førsteforfatteren.



Figur 2: I denne figur illustreres sammenhængen mellem modellens risikoscores, og faktiske hændelser målt i løbet af 2018. De stiplede linjer angiver de gennemsnitlige værdi af udfaldsmålne blandt personer i testdatasættet. I figuren med sigtelser i øverste venstre hjørne betragtes kun de 14-17 årige i testdatasættet.

Den gode eksterne validitet af modellen er betryggende i den forstand, at selvom modellen er bygget ved at forudsige anbringelser, så giver den eksterne validering evidens for, at modellens risikovurderinger kan tolkes ind i en bredere kontekst af mistrivsel.

4 Diskussion

Modellen medtager, som allerede nævnt, ikke information om køn eller etnicitet, for ikke at komme til at videreføre eventuelle biases i historiske beslutninger i underretningssager. Der kan imidlertid stadig være variable inkluderet i modellem, som er korreleret med køn og/eller etnicitet, og som derved fører til, at modellen giver forskellige prædiktioner for disse grupper. Nogle gange kan disse forskelle være velbegrundede, hvis for eksempel seksuelle overgreb ses hyppigere mod piger end mod drenge. Det er dog stadig en god ide at analysere sådanne forskelle, da eventuelle forskelle også kan skyldes historisk skævhed i beslutningerne, hvilket kan tolkes som diskrimination/unfairness. Sådanne forskelle bør ikke afspejles og dermed videreføres i de prædikterede sandsynlighed og risikoscores.

Baseret på en række statistiske mål for bias/fairness opstillet af Chouldechova (2017) viser vi, at modellen ikke fører til væsentligt forskellige prædiktioner, når de opdeles på køn og etnicitet.⁷ Dette udsagn skal forstås i den forstand, at der for en given risikoscore ikke anbringes væsentligt flere piger end drenge (og vice versa). På samme vis finder vi, at der heller ikke er substantielle forskelle i anbringelsesraterne mellem børn af ikke-vestlig- og vestlig (herunder dansk) baggrund.

Til gengæld viser modellen, at modellen prædikterer meget skævt, når de børn, underretningerne vedrører, opdeles efter forældrenes socioøkonomiske status. Her ses en klar tendens til flere anbringelser for børn i familier med lavere socioøkonomisk status, men dette ses ved alle risikoscores og afspejler et historisk mønster i de foregående beslutninger, som modellen er beregnet på grundlag af. Hvorvidt dette er et reelt problem, eller om det afspejler, at højere socioøkonomisk status kan være en proxy for andre beskyttende faktorer, er vi i færd med at analysere.

5 Konklusion

Det konkluderes forsigtigt, at den opstillede statistiske model har gode prædiktive egenskaber, og at den også har stor ekstern validitet. Vi vurderer derfor, at den udgør et rimeligt præcist mål og samtidig er et godt mål for faktisk mistrivsel.

⁷ Konkret anvendes fairness begreberne 'calibration' og 'predictive parity'.



Endelig vurderes, at modellens egenskaber med hensyn til videreførelse af bias med hensyn til køn og etnicitet er tilfredsstillende, og at en store skævhed med hensyn til alder kan elimineres ved at foretage transformationen fra prædikterede sandsynligheder til risikoscores på en måde, så modellen gøres helt aldersneutral. Således vil den model, der afprøves fremadrettet, være fri for opdagede utilsigtede skævheder.⁸ Bemærk, at modellen under alle omstændigheder bidrager til at bringe disse skævheder, som afspejler historisk praksis i de afgørelser, som socialrådgiverne har foretaget i perioden fra og med april 2016 til og med december 2017, frem i lyset.

5.1 Analyser under udarbejdelse

Under arbejdet med modellen har vi modtaget en lang række kommentarer og feedback fra fagfæller og andre interesserede, bl.a. i forbindelse med indsendelse af artiklen til et videnskabeligt tidsskrift med henblik på publicering af den videnskabelige artikel. Disse kommentarer har ført til yderligere arbejde med at analysere modellens egenskaber i en række dimensioner.

Som allerede nævnt er vi i færd med at analysere eventuelle biases i modellen og undersøge forskellige muligheder for at eliminere disse enten i modelfasen eller *ex post* ved for eksempel at foretage decil-inddelingen inden for de grupperinger, hvor skævhederne opstår. I første omgang har vi gjort dette inden for aldersgrupper, men det er sandsynligt, at denne decil-inddeling senere kommer til at inddrage flere faktorer end blot alder.

Vi arbejder også på at udvikle et mere generelt mål for mistrivsel baseret på en faktoranalyse af alle de 20 variable, som indgår i den eksterne validering af modellen. Hensigten med dette mål er primært (i første omgang) anvendelse i den eksterne valideringsanalyse.

I næste fase af forskningsprojektet afprøves Beslutningsstøtten i endnu en pilotafprøvning, hvor den nye statistiske model samt en ny brugerflade skal afprøves, inden Beslutningsstøtten i løbet af 2022 forhåbentlig vil være tilstrækkeligt valideret til afprøvning i et større lodtrækningsforsøg.

⁸ Der kan naturligvis stadig være skævheder vi ikke er opmærksomme på, og som modellen derfor ikke tager højde for, men i det omfang vi bliver opmærksomme herpå, vil vi analysere disse problemstillinger og justere modellerne efter behov.

6 Referencer

Arad-Davidzon, B., Benbenishty, R., 2008. The role of workers' attitudes and parent and child wishes in child protection workers' assessments and recommendation regarding removal and reunification. *Children and Youth Services Review*, 30(1), 107-121.

Arad-Davidzon, B., Benbenishty, R., 2010. Contribution of child protection workers' attitudes to their risk assessments and intervention recommendations: A study in Israel. *Health & Social Care in the Community*, 18(1), 1-9.

Bartelink, C., van Yperen T. A., Ingrid, J., 2015. Deciding on child maltreatment: A literature review on methods that improve decision-making. *Child Abuse & Neglect*: 49, 142-153.

Benbenishty, R., Davidson-Arad, B., López, M., Devaney, J., Spratt, T., Koopmans, C., Knorth, E. J., Witteman, C. L. M., Del Valle, J. F., Hayes, D., 2015. Decision making in child protection: An international comparative study on maltreatment substantiation, risk assessment and interventions recommendations, and the role of professionals' child welfare attitudes. *Child Abuse & Neglect*: 49, 63-75.

Chen, T., Guestrin, C., 2016. Xgboost: A scalable tree boosting system. *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*, 785-794.

Chouldechova, A., 2017. Fair prediction with disparate impact: A study of bias in recidivism prediction instruments. *Big Data*, 5, 153-163.

Chouldechova, A., Benavides-Prado, D., Fialko, O., Vaithianathan, R., 2018. A case study of algorithm-assisted decision making in child maltreatment hotline screening decisions. In: *Conference on Fairness, Accountability and Transparency*. 134–148.

Christoffersen, M. N., C. Armour, M. Lasgaard, T. E. Andersen, A. Elklit. (2013). The prevalence of four types of childhood maltreatment in Denmark. *Clinical practice and epidemiology in mental health: CP & EMH*, 9, 149.

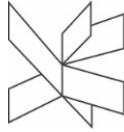
Coulthard, B., Mallett, J., Taylor, B.J., 2020. Better decisions for children with 'big data': can algorithms promote fairness, transparency and parental engagement? *Societies*, 10, 97. [Link til artikel.](#)

Danneskiold-Samsøe, S., Baviskar, S., Bergström, M., 2020. Kan målinger af anbragte børns trivsel understøtte samfundets ansvar? *Udenfor Nummer*, 40, 24-33.

- Drake, B., Jonson-Reid, M., Ocampo, M. G., Morrison, M., Dvalishvili, D., 2020. A Practical Framework for Considering the Use of Predictive Risk Modeling in Child Welfare. *The ANNALS of the American Academy of Political and Social Science*, 692(1), 162-181. doi: [Link til artikel](#).
- Ejrnæs, M., Jakobsen, I. S., Lyngholm, L. T. M., Michelsen, R. R., Moesby-Jensen, C. K., Monrad, M., 2020. Professionelle praktikers risikovurderinger – og de faktiske risici. *Udenfor Nummer*, 40, 34-47.
- Enosh, G., Bayer-Topilsky, T., 2015. Reasoning and bias: Heuristics in safety assessment and placement decisions for children at risk. *The British Journal of Social Work*, 45(6), 1771-1787.
- Fluke, J. D., López, M. L., Benbenishty, R., Knorth, E. J., Baumann, J. D., 2021. *Decision-Making and Judgment in Child Welfare and Protection*. Oxford University Press.
- Hillis, S. J. Mercy, Amobi, A., Kress, H., 2016. Global prevalence of past-year violence against children: a systematic review and minimum estimates. *Pediatrics*, 137(3).
- O’Neil, C., 2016. *Weapons of Math Destruction – how big data increases inequality and threatens democracy*. Penguin Random House UK.
- Rossi, P. H., Schuerman, J., Budde, S. (1999). Understanding decisions about child maltreatment. *Evaluation Review*, 23(6), 579-598.
- Rutter, M., 2012a. Resilience: Causal Pathways and Social Ecology. P.33- 42. I: Ungar, M (ed): *The Social Ecology of Resilience: A Handbook of Theory and Practice*. Springer Science + Business Media LLC, 2012.
- Rutter, M., 2012b. Resilience as a dynamic concept. *Developmental and Psychopathology*, 24, 335-344.
- Servicebogen. ”Bekendtgørelse af lov om social service”. [Retsinformation](#) (januar 2021).
- Schuerman, J., Rossi, P. H., Budde, S., 1999. Decisions on placement and family preservation: Agreement and targeting. *Evaluation Review*, 23(6), 599-618.
- Søbjerg, L. M., Taylor, B. J., Przeperski, J., Horvat, S., Nouman, H., Harvey, D., 2020. Using risk-factor statistics in decision making: prospects and challenges. *European Journal of Social Work*. [Link til artikel](#).
- Taylor, B.D., 2021. Teaching and Learning Decision-Making in Child Welfare and Protection Social Work. In Fluke et al. (eds.).
- Villumsen, A. M. A., 2015. Hvorfor det ikke er så lige til med udsathed hos børn og unge. I. D. Graversen (red.), *Pædagogik: introduktion til pædagogens grundfaglighed* (1 udg., Bind 1). Systime.



TRYGFONDENS
BØRNEFORSKNINGSCENTER
INSTITUT FOR ØKONOMI
AARHUS UNIVERSITET



VIA University
College

TrygFonden

Websites:

[Projektside hos TrygFondens Børneforskningscenter](#)

[Projektside på UC Viden](#)

DELVIST FORÆLDET