

Et scoping review omhandlende begrænsninger ved udviklingen af prognostiske kliniske beslutningsstøttesystemer og en validering af det prognostiske beslutningsstøttesystem ERVIN



Abstrakt

Titel: Et scoping review omhandlende begrænsninger ved udviklingen af prognostiske kliniske beslutningsstøttesystemer og en validering af det prognostiske beslutningsstøttesystem ERVIN

Uddannelse: Klinisk Videnskab og Teknologi 4. semester, speciale projekt

Vejleder: Kristian Kjær-Staal Pedersen

Bivejleder: Jan Brink Valentin

Projektgruppe: 505

Forfattere: Mads Thøgersen, 20210538 og Ditmer Kiis, 20211859

Introduktion og formål: Artrose er en globalt udbredt lidelse med store fysiske, psykiske, sociale og økonomiske konsekvenser. Artrose behandles enten konservativt eller kirurgisk, men der foreligger ikke klare retningslinjer for, hvornår patienten skal tilbydes den ene behandling frem for den anden. Kliniske prognostiske beslutningsstøtte systemer (CDSS) er statistiske værktøjer, der igennem patientinddragelse kan hjælpe klinikeren og patienten til at træffe en beslutning om behandling. Et af disse systemer er ERVIN, der er udviklet til at prædikterer blandt andet reoperation, korttidskomplikationer for både total hoftealloplastik (THA) og total knæalloplastik (TKA) patienter og luksationer for THA patienter, det er dog uvist hvor pålidelig ERVIN er til at prædiktere de førnævnte ricisi udfald. Der findes flere begrænsninger ved udviklingen af prognostiske CDSS'er, det er dog ubekendt hvilke der opgives og i hvilket omfang de opgives. **Metode:** Metoden i dette projekt var todelt. Første del bestod af et systematisk scoping review i databaserne Pumbed, Embase og IEEE med formål at afdække begrænsninger ved udviklingen af prognostiske CDSS'er. Anden del af metoden bestod af en temporal validering af CDSS'en ERVIN. Data til validering af reoperation og korttidskomplikationer for både THA- og TKA-patienter samt luksationer for THA-patienter blev indhentet fra Aalborg Universitets Hospital Farsø via BI og Analyse i Region Nordjylland. Der blev udført kalibrering-i-det-store (A), kalibreringshældning (B), diskriminering (AUC) og en beslutningskurveanalyse. **Resultater:** I scoping review blev 99 studier inkluderet hvori der var opgjort 355 begrænsninger fordelt på bias, datakvalitet, generaliserbarhed, manglende ekstern validering, manglende udfald, overtilpasning, populationsdatashift, problemer med prædiktorer, rodekasse og sample size. I valideringen af ERVIN er THA reopration (A: -1,569. B: 0,096. AUC: 0,42 [95% konf. interval 0 1]), THA korttidskomplikationer (A: 0,950. B: 0,215. AUC: 0,57 [95% konf. interval 0,33 0,82]), THA luksationer (-), TKA reopration (A: -0,779. B: 0. AUC: 0,46 [95% konf. interval 0 1]), TKA korttidskomplikationer (A: 1,810. B: -0,117. AUC: 0,48 [95% konf. interval 0,34 0,63]). **Konklusion:** Der fremkommer stor variation i hvilke, i hvilket omfang og hvordan begrænsninger ved udvikling af prognostiske CDSS'er opgøres. Der findes dog væsentlige mangler i udviklernes forståelse og håndtering af disse. Resultaterne for valideringen af ERVIN indikerer, at ERVIN ikke pålideligt kan prædiktere reoperation, korttidskomplikationer for både THA- og TKA-patienter og luksationer for THA-patienter, resultaterne er dog påvirket ringe sample size.

Antal sider: 92

Bilag: 6 stk

Referencesystem: Vancouver

Abstract

Title: A Scoping Review on limitations of the development of prognostic Clinical Decision Support Systems and validation of the prognostic clinical decision support system ERVIN

Education: Clinical Science and Technology, 4th Semester, a master's thesis.

Supervisor: Kristian Kjær-Staal Pedersen

Co-supervisor: Jan Brink Valentin

Project Group: 505

Authers: Mads Thøgersen, 20210538 and Ditmer Kiis, 20211859

Introduction and purpose: Osteoarthritis is a globally prevalent disorder with significant physical, psychological, social, and economic consequences. Osteoarthritis is treated either conservatively or surgically, but clear guidelines for determining which treatment to offer the patient are lacking. Clinical decision support systems (CDSS) are statistical tools that, through patient involvement, can assist clinicians and patients in making treatment decisions. One of these systems is ERVIN, which is developed to predict, among other things, reoperation, short-term complications for both total hip arthroplasty (THA) and total knee arthroplasty (TKA) patients, and dislocations for THA patients. However, it is uncertain how reliable ERVIN is in predicting the outcomes. There are several limitations to the development of prognostic CDSSs, but it is unknown which limitations are reported and to what extent they are reported. **Method:** The method in this project was divided into two parts. The first part consisted of a systematic scoping review in the databases PubMed, Embase, and IEEE, with the aim of identifying limitations in the development of prognostic decision support systems. The second part of the method involved a temporal validation of the CDSS ERVIN. Data for validating reoperation, short-term complications for both THA and TKA patients, and dislocations for THA patients were obtained from Aalborg University Hospital Farsø through BI and Analysis in Region North Jutland. Calibration-in-the-large (A), calibration slope (B), discrimination (AUC), and decision curve analysis were performed. **Results:** The scoping review included 99 studies in which 355 methodological limitations were identified, distributed across bias, data quality, generalizability, lack of external validation, missing outcomes, overfitting, population drift, predictor issues, mixed problems, and sample size. The result for the validation of ERVIN were THA reoperation (A: -1.569, B: 0.096, AUC: 0.42 [95% confidence interval 0-1]), THA short-term complications (A: 0.950, B: 0.215, AUC: 0.57 [95% confidence interval 0.33-0.82]), THA dislocations (-), TKA reoperation (A: -0.779, B: 0, AUC: 0.46 [95% confidence interval 0-1]), and TKA short-term complications (A: 1.810, B: -0.117, AUC: 0.48 [95% confidence interval 0.34-0.63]). **Conclusion:** There is significant variation in how, to what extent, and how the limitations of the development are reported. However, there are substantial deficiencies in the developers' understanding and handling of these limitations. The results of the ERVIN validation indicate that ERVIN cannot reliably predict reoperation, short-term complications for both THA and TKA patients, and dislocations for THA patients, although the results are influenced by a small sample size.

No. of pages: 92

Appendix: 6 pcs

Reference style: Vancouver

Forord

Dette specialprojekt er udarbejdet af projektgruppe 505 som en del af 4. semester på Klinisk Videnskab og Teknologi på Aalborg Universitet. Gruppen består af to specialestuderende med baggrund som fysioterapeuter. Specialet er udarbejdet fra 1. februar til 1. juni 2023.

Formålet med dette specialeprojekt er at identificere videnshuller ift. begrænsninger ved udvikling af prognostiske kliniske beslutningsstøttesystemer og validering af det kliniske beslutningssystem ERVIN, der anvendes på Ortopædkirurgisk Afdeling Aalborg Universitetshospital, Farsø. Projektet henvender sig til personer som har interesse i udvikling, især validering, af prognostiske kliniske beslutningsstøttesystemer.

En stor tak skal sendes til vejledere for specialet, senior statistiker og vidtfavnende Jan Brink Valentin og Ph.d., Dr.Med. Kristian Kjær-Staal Petersen, for sparring og konstruktiv vejledning i udarbejdelsen af dette projekt. Ligeledes en stor tak til Ph.d.-studerende Nanna Kastrup for at bidrage med viden, undervisning og dokumentation omhandlende ERVIN.

Indhold

1.0 Begrebsafklaring.....	7
2.0 Problemanalyse.....	8
2.1 Artrose.....	8
2.1.1 Artroses prævalens, incidens og sundhedsøkonomiske konsekvenser	8
2.1.2 Diagnosticering af artrose	9
2.1.3 Behandling af artrose	9
2.1.4 Fordele og ulemper ved konservativ behandling	10
2.1.5 Fordele og ulemper ved kirurgisk behandling.....	10
2.1.6 Patientinddragelse	11
2.2 Kliniske beslutningsstøttesystemer.....	11
2.2.1 Diagnostiske og prognostiske beslutningsstøttesystemer	12
2.2.2 Udvikling af et prognostisk beslutningsstøttesystemer	12
2.2.3 Opdatering af prognostiske kliniske beslutningsstøttesystemer	13
2.2.4 Udfordringer ved udviklingen af kliniske beslutningsstøttesystemer	14
2.2.5 ERVIN - et prognostisk beslutningsstøttesystem	15
3.0 Problemafgrænsning	17
4.0 Problemformulering	18
5.0 Metode	18
5.1 Scoping review.....	18
5.1.1 Metodisk ramme for scoping review	18
5.1.2 Systematisk litteratursøgning.....	19
5.1.3 In- og eksklusionskriterier for den systematiske litteratursøgning	20
5.1.4 Systematisk udvælgelse af litteratur	21
5.1.5 Udtrækning af data	22
5.2 Temporal validering af ERVIN.....	23
5.2.1 Datakilde	23
5.2.2 Databehandling	23
5.2.3 Kalibrering-i-det-store (A)	23
5.2.4 Kalibreringshældningen (B)	24
5.2.5 Diskriminering (C).....	24
5.2.6 Beslutningskurveanalyse (D)	25

6.0 Resultater.....	27
<i>6.1 Resultaterne for scoping review:</i>	<i>27</i>
6.1.1. Inkluderede studier	28
6.1.2 Selvopgjorte begrænsninger ved udviklingen i inkluderede studier	28
<i>6.2 Temporal validering af ERVIN.....</i>	<i>32</i>
6.2.1 Demografi på valideringsdata	33
6.2.2 Validering af THA-prædiktioner	34
6.2.3 Validering af TKA-prædiktioner.....	37
<i>6.3 Opsummering af resultater</i>	<i>40</i>
7.0 Diskussion	41
<i>7.1 Resultatdiskussion</i>	<i>41</i>
7.1.1 Resultatdiskussion af scoping review	41
7.1.2 Resultatdiskussion af valideringen af ERVIN	42
<i>7.2 Metodediskussion.....</i>	<i>44</i>
7.2.1 Metodediskussion af scoping review	44
7.2.2 Metodediskussion af valideringen af ERVIN.....	45
8.0 Konklusion	46
9.0 Perspektivering	47
10.0 Referenceliste	48
11.0 Bilag	64
<i>Bilag 1: Patient Reported Outcomes (PRO)</i>	<i>64</i>
<i>Bilag 2: Valideringstyper.....</i>	<i>65</i>
<i>Bilag 3: Risici udfald.....</i>	<i>66</i>
<i>Bilag 4: Søgestreng.....</i>	<i>67</i>
<i>Bilag 5: Inkluderede studier</i>	<i>68</i>
<i>Bilag 6: Begrænsninger ved udviklingen i de enkelte studier</i>	<i>81</i>

1.0 Begrebsafklaring

Artrose: Artrose, også kendt som slidgigt, er en degenerativ ledsgdom, der sker nedbrydning af brusk og knoglevæv i et eller flere led, hvilket resulterer i smerte, stivhed og nedsat funktionsevne(1,2).

Total hoftealloplastik: En total hoftealloplastik (THA) er en kirurgisk operation, hvor man erstatter et beskadiget hoftelede med en kunstig hofteprotese(3,4).

Total knæalloplastik: En total knæalloplastik (TKA) er en kirurgisk operation, hvor man erstatter et beskadiget knæled med en kunstig knæprotese(5,6).

Klinisk beslutningsstøttesystem: Clinical Decision Support System (CDSS) er en type software, der hjælper klinikere med at træffe kliniske beslutninger baseret på sundhedsdata(7,8).

Effekt og Risiko Ved INdgreb (ERVIN): Et klinisk beslutningsstøtte system til at hjælpe ortopædkirurger med at træffe en oplyst beslutning om operation eller konservativ behandling for hofte- eller knæartrose(9–11).

Diskriminering: Diskriminering refererer til modellens evne til at differentiere mellem forskellige udfald og eller grupper af data(8,12).

Kalibrering: Omhandler modellens evne til at forudsige overensstemmende prædikterede udfald og faktisk udfald(12).

Intern validering: Ved intern validering af en prognostisk model evalueres modellens ydeevne ved hjælp af data, der blev brugt under modellens træning og tilpasning(8).

Ekstern validering: Ved ekstern validering af en prognostisk model evalueres modellens ydeevne ved hjælp af separate og uafhængige datasæt, der ikke blev brugt under modellens træning(8).

Temporal validering: Ved temporal validering af en prognostisk model evalueres modellens ydeevne ved at teste den på data, der stammer fra en anden tidsperiode end den der blev brugt under modellens træning(8).

Populationsdatashift: Ændringer i det CDSS evne til at forudsige udfald på grund af ændringer i data i populationen(13).

Interventionsdatashift: Ændringer i det CDSS evne til at forudsige udfald på grund af interventionen bliver den primære prædiktor for udfald(13).

2.0 Problemanalyse

Følgende afsnit er en analyse de problemstillinger, der ligger til grund for problemafgrænsningen og problemformuleringen i dette projekt.

2.1 Artrose

Patologien ved artrose er kendtegnet ved en gradvis degeneration af bruskvævet og osteofytdannelse i ledet samt inflammation i synovialmembranen, hvilket resulterer i ændringer i ledets form og funktion samt ledsmærter(1,2,14,15). Over tid kan artrose forårsage symptomer som hævelse, stivhed og begrænsning af bevægelighed(1,16), dog vil 50 % af personer med strukturelle ændringer svarende til artrose være asymptotiske(15). For patienter med smerter grundet artrose kan sygdommen resulterer i komplekse smerteproblematikker, herunder refererede smerter(1,17). Risikoen for artrose øges særligt af adipositas og alder, men også andre faktorer kan øge risikoen herunder medicinering, skader, genetiske faktorer, køn, fejlernæring, tidlige operationer i ledet samt systemiske sygdomme(14,18). Artrose har betydelige konsekvenser for den enkelte patients livskvalitet, hvor patienterne oplever begrænsninger i deres fysiske dagligdags aktiviteter, arbejdsliv og sociale interaktioner, dårlig søvn, udtrætning og tab af uafhængighed(18–24). Dette medfører at artrosepatienter har højere udgifter til sundhedsydelsler end raske(24). I et studie af Vitaloni, Bemden, Sciortino et al. (2020) oplevede 27,49 % af patienterne at deres følelsesmæssige, psykologiske og mentale sundhed påvirkes af artrose. I samme studie angav 24,6 % af patienterne, at de ikke forstår deres behandlingsmuligheder og de dertilhørende risici(20). Hofte og knæ er blandt de hyppigst afficerede led(18,22).

2.1.1 Artroses prævalens, incidens og sundhedsøkonomiske konsekvenser

Artrose er en globalt udbredt gigtsygdom og prævalensen steg på verdensplan med 113,25 % fra 247,51 millioner i 1990 til 527,81 millioner i 2019 og incidensen var 6348,25 pr. 100.000 i 2019(25). I Danmark angav 19% af befolkningen i 2014 symptomer på artrose, i 2021 var dette tal steget til 22,5%, hvilket medfører at artrose er den næstmest forekommende langvarige sygdom i Danmark(26). Ifølge Sundhedsstyrelsens seneste rapport om Sundhedsbyrden i Danmark fra 2022 kostede artrose i 2017 4,6 milliarder i sundhedsomkostninger og 12,4 milliarder i funktionstab, hvilket gør artrose til den femte mest omkostningstunge sygdom i Danmark(27).

2.1.2 Diagnosticering af artrose

Hofte- og knæartrose kan diagnosticeres på grundlag af patient historik, en fysisk undersøgelse samt en radiologisk undersøgelse(28). Det kliniske billede vil typisk være smerte og stivhed i det ramte led. Stivhed er typisk værst fra morgenen, eller efter at patienten har været stillesiddende i længere tid, og forbedres i løbet af 30 minutter. Smerterne er i sygdomsbegyndelsen aktivitetsafhængige, men bliver mere uforudsigelige over tid(28,29). Den fysiske undersøgelse består typisk af inspektion og palpation af ledet, bevægelighedsundersøgelse og funktionsundersøgelse(28,29). Studier har vist at der begrænset sammenhæng mellem radiologisk påvist artrose og smerter hos patienter med artrose i hofte eller knæ(1,15,16). På billede 1 ses et rask (A) og et artroseramt (B) højre knæ på røntgen anteriort.



Billede 1: Billede A foreviser et rask knæ og billede B foreviser et artrose ramt knæ med medialt sammenfald, osteofytter, ujævnheder og sklerosering(30).

2.1.3 Behandling af artrose

Ifølge guidelines fra Osteoarthritis Research Society International (OARSI), skal behandlingen af artrose målrettes mod, at reducere ledsmerte, stivhed og tab funktionsniveau, fastholde og forbedre ledbevægelighed og sundhedsrelateret livskvalitet, begrænse progression af ledskade samt at uddanne patienterne i forståelse og håndtering af artrose(31,32). Behandlingen af artrose består af henholdsvis konservativ behandling og kirurgisk behandling(6,32–35). Patienterne starter ofte med at blive tilbuddt konservativ behandling, men ved vedvarende smerter og funktionstab tilbydes patienterne ofte operation(28). Den konservative behandling bør som minimum bestå af en eller flere af følgende; øvelsesterapi herunder styrkestræning, kredsløbstræning og balancetræning samt vægttab og patientuddannelse(34,36), og kan kombi-

neres med farmakologisk behandling(32,34,37). Den kirurgiske behandling består af primært alloplastikker, men kan også være atroskopier eller aksekorrigerende osteomier(28,33,38,39). Ved total hoftealloplastik (THA) udskiftes hele ledet i form af caput femoris, collum femoris samt acetabulum og opdeles i cementerede og ucementerede operationer(3,4). Ved cementerede THA fikseres alloplastiken med ”cement” imens den ved ucementerede skrues fast(40). Ved total knæalloplastik (TKA) udskiftes både tibia-pleauet og kondylerne udskiftes(6). I 2021 blev henholdsvis 10262 og 10549 opereret med enten en hofte- eller knæalloplastik i Danmark(41,42).

Behandlingsvalget influeres af mange faktorer såsom forhøjet BMI, patientens alder, smerte- og funktionsniveau samt patientens og klinikerens præferencer(6,43). Både konservativ- og kirurgiskbehandling er effektive behandlingsformer, findes der ikke klarer nationale eller internationale retningslinjer for, hvornår patienten bør tilbydes operation frem for konservativ behandling(10,39,44).

2.1.4 Fordeler og ulemper ved konservativ behandling

Det er påvist at konservativ behandling til patienter med artrose, især kombination af træning, information, farmakologisk behandling, patientrådgivning og vægtab, reducerer smerter og funktionstab(28,33,37,45–47). De positive effekter af konservativ behandling fører nødvendigvis ikke til øget fysisk aktivitet hos patienterne(48). Ligeledes kan mangel på efterlevelse hos patienterne være en barriere ift. at opnå fuld effekt af konservativ behandling(49). Et studie af Selten, Vreizkolk, Greenen et. al (2016) fastslår, at konservativ behandling er underanvendt, og at forklaringerne på dette kan være frygt for medicineringsbivirkninger, bevægelse eller at den konservative behandling ikke følger kliniske retningslinjer samt at patientens præferencer og værdier ikke indtænkes i behandlingen(50). Et studie af Crawford, Miller og Block (2013) argumenterer for, at konservativ behandling er en mangelfuld strategi til behandling af knæartrose på grund af ineffektiv smertelindring, potentielle bivirkninger ved NSAID og manglende evne til at forsinke sygdomsprogression(51). Personer med fremskreden artrose eller manglende effekt af konservativ behandling kan være passende kandidater til THA eller TKA(28,52,53).

2.1.5 Fordeler og ulemper ved kirurgisk behandling

THA og TKA er succesfulde kirurgiske behandlinger, som øger livskvalitet for patienter med fremadskreden artrose, ved at behandle smerte og forbedre funktion(6,43,52,54–56). Der findes dog også flere potenti-

elle komplikationer ved THA og TKA, det kan f.eks. være risiko for infektioner, dyb venøs trombose, smerter, luksationer, nerveskader bevægelsesindskrænkninger og komplikationer med protesen(53,57–59). Andelen af reoperationer i THA-patienters led er på 2,05% imens andelen for TKA-patienter er på 1,18%(60). Ydermere understreger et studie risikoen for at nogle patienter får mere end én reoperation(61). Komplikationer indenfor 30 dage er for THA 3,67 % og 3,49 % for TKA(60) og 2 % af THA patienter oplever hofteluksation indenfor det første år(35). Studier påviser at op til 30% af patienterne oplever utilfredsstillende resultater som følge af THA eller TKA(10,55,56,62,63), hvilket kan skyldes et misforhold mellem forventninger og resultatet(63). Ydermere indikerer et studie Alokozai, Berstein, Samuel et al. (2021) at op imod en tredjedel af alle TKA operationer var unødvendige og at årsagen til dette kan være manglende patientinddragelse(44).

2.1.6 Patientinddragelse

Patientinddragelse er en helhedsorienteret tilgang til behandling, hvor det sundhedsfaglige personale opmuntrer patienten til at tage ansvar for deres egen behandling og tage aktiv del i beslutningerne om deres helbred(64,65). En stor del af den medicinske verden bevæger sig i disse år imod større patientinddragelse(65–67). Flere studier beskriver at patientinddragelse i forhold til beslutningstagning omhandlende patienternes egen behandling forbedrer patienttilfredshed, behandlingskvaliteten, patientefterlevelse og sundhedsudfald som f.eks. vægtab(44,67,68). Ligeledes hjælper patientinddragelse det faglige personale med at vejlede patienter meningsfuldt(65). Ved fælles beslutningstagning, hvor patient og kliniker arbejder som ligestillede, kan der arbejdes evidensbaseret samtidig med at patientens ønsker og værdier inddrages i beslutningstagningen(69,70). Øget patientinddragelse kan realiseres ved at anvende "Patient Reported Outcomes" (PRO)(71,72), se bilag 1: Patient Reported Outcomes. Ifølge et studie af Stacey, Taljaard, Dervin et al. (2016), kan anvendelse af beslutningsstøtteværktøjer til patientinddragelse hos patienter med hofte- eller knæartrose medvirke til, at patienterne bedre kan tage stilling til om de ønsker konservativ eller kirurgisk behandling(73). Samme studie viste en forbedring i patienternes forventninger til smertelindring ved gang efter kirurgiske behandling for THA og TKA patienter(73).

2.2 Kliniske beslutningsstøttesystemer

Et klinisk beslutningsstøttesystem (CDSS) er en statistiske prædiktionsmodel, der giver sundhedspersonale og patienter computerbaseret støtte til at træffe oplyste beslutninger i forhold til patientbehandling

og forebyggelse. Formålet er at forbedre og målrette den kliniske viden samt give patientspecifik information, hvilket forbedre patientinddragelsen(74). Dette gøres ved at prædiktere kliniske udfald eller identificere patienter med risiko for at udvikle sygdom, baseret på multiple data fra patienten(75–77). CDSS'er kan muligvis bruges til at forbedre kvaliteten og effektiviteten af sundhedsydeler og fastsatte udfald for modellerne. Prædiktionsmodeller kan ikke anvendes til at forklare årsagen til at sygdom opstår(74), men kan, gennem en multivariabel tilgang, give et billede af patientens helbredstilstand og deres risiko for komplikationer. Dette kan give klinikere og patienter i samråd mulighed for at vurdere, om mere invasive eller omkostningsfulde undersøgelser er nødvendige for at få en mere præcis diagnose eller behandling, eller om billigere og mindre invasive alternativer kan være lige så effektive(76). Modellerne kan anvendes i forskellige omgivelser til forskellige formål, herunder både diagnostiske og prognostiske formål(75–77).

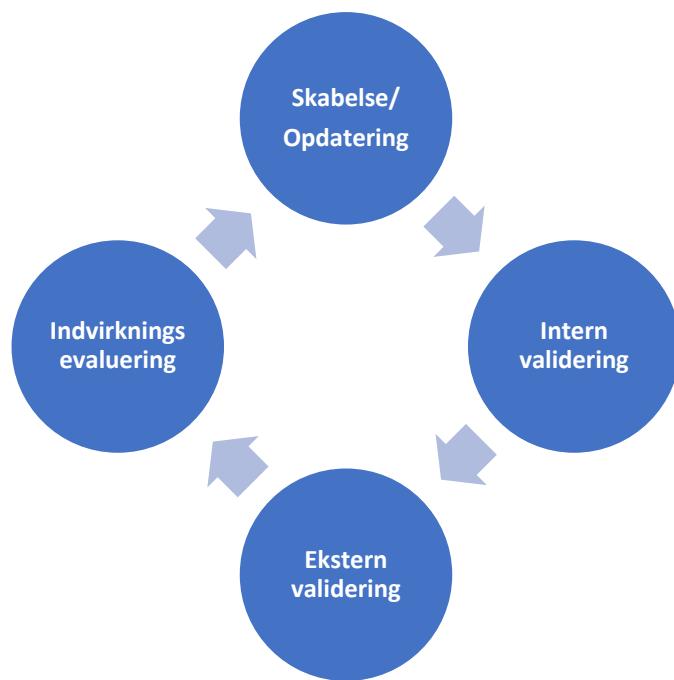
2.2.1 Diagnostiske og prognostiske beslutningsstøttesystemer

Diagnostiske CDSS'er er designet til at give sundhedspersonalet vejledning i at træffe den korrekte diagnose. Disse systemer kombinerer typisk patientinformationer, symptomer mv. til at generere en liste over mulige diagnoser samt deres sandsynlighed. Klinikere kan bruge disse informationer til udføre yderligere relevante kliniske test(74,78). Prognostiske CDSS'er kombinerer typisk patientspecifikke informationer, PRO-data, biomarkører mv. til at prædiktere et givent udfald ved et bestemt scenarie. Dette kan anvendes til vejlede læger og patienter i, hvad der bør gøres i et givent scenarie(7,8). Forskellen mellem de to typer CDSS'er ligger primært i deres formål og anvendelse(79). Dette projekt vil udelukkende omhandle prognostiske CDSS'er.

2.2.2 Udvikling af et prognostisk beslutningsstøttesystemer

Et prognostisk CDSS gennemgår i sin udvikling flere faser(8,80). Den første fase omhandler skabelsen af modellen, men forudsætter samtidig, at behovet for et CDSS er identificeret i det pågældende emnefelt. En fundamental del af den tidlige udvikling er at indsamle en sufficient mængde udviklingsdata, håndterer manglende data, udvælge relevante prædiktorer og deres vægtning samt at beslutte hvilke statistiske metoder der anvendes til prædiktionen(8,77). I den anden fase evalueres modellens ydeevne igennem en intern validering, se Bilag 2: Valideringstyper, hvor kalibreringen og diskrimineringen undersøges i en afledt prøve, men fra data modellen er udviklet på, hvilket evaluerer reproducerbarheden og forhindrer overtilpasning af nuværende data(8,81). Kalibrering henviser til, hvor præcist en model forudsiger sandsynlighederne for et given udfald. En velkalibreret model forudsiger begivenheder der afspejler den faktiske

frekvens af den givne begivenhed i den pågældende population. Diskriminering refererer til modellens evne til at differentiere mellem forskellige udfald og eller grupper af data(8). Både kalibrering og diskriminering er vigtig fordi en model der prædikterer den samme beregnede risiko for alle individer svarende til forekomsten, vil vise perfekt kalibrering, men er ubrugelig, da modellen ikke skelner mellem positive og negative udfald. Samtidigt vil en model med perfekt diskriminering, men ukorrekt kalibrering ikke kunne give præcise forudsigelser af patienternes faktiske risiko(8). Den tredje fase, som bør gennemføres inden implementeringen i klinisk praksis, er ekstern validering. Ved ekstern validering testes modellens generalisierbarhed på en anden cohorte af patienter end den er udviklet på(8,77,82), se Bilag 2: Valideringstyper. I den sidste fase implementeres CDSS'et i klinisk praksis, hvor indvirkningen undersøges og om CDSS'et faktisk forbedre beslutningstagningsprocessen, de kliniske udfald eller hvad modellen var bygget til at forbedre(76). Efter evalueringen bør modellen løbende opdateres eller deimplementeres(12,13). Udviklingen af en prognostiske model er illustreret i figur 1.



Figur 1: Udviklingen af et prognostisk CDSS. Skabelse/opdatering; identificering af behovet for et CDSS for det givne område, identificering af prædiktorer som vægtes, indsamling af testdata og ydeevnen testes i form af intern validering. Validering; Ydeevnen testes via ekstern validering. Indvirkningsevaluering; Indvirkningen undersøges og om CDSS forbedre beslutningstagningsprocessen, de kliniske udfald eller hvad modellen var bygget til at forbedre(12,76).

2.2.3 Opdatering af prognostiske kliniske beslutningsstøttesystemer

Mange forskere og programmører anerkender, at der udvikles for mange nye modeller(12). Såfremt der er udviklet en nogenlunde veldiskreminered og -kalibreret model indenfor emnefeltet, bør fokus være at opdatere og validere denne. Flere ting kan ændre et prognostisk CDSS's ydeevne over tid og i sidste ende

medføre, at CDSS'et bliver redundant, herunder ændringer i patientdemografien, diagnosticeringen og behandlingsprocedure. Over tid vil dette føre til fejlkalibrering, hvorfor det er derfor vigtigt at opdatere CDSS'et(12,83). Opdatering af modeller opdeles typisk i ren opdatering, herunder rekalibrering og revision, samt modeludvidelse, hvor der eksempelvis inkorporeres nye prædiktorer eller udfald i modellen(12). Efter opdatering bør en model revalideres og reevalueres. Anvendelse af CDSS'er kan derfor ikke anses som en enkeltstående hændelse, men en konstant iterativ proces(12,83).

2.2.4 Udfordringer ved udviklingen af kliniske beslutningsstøttesystemer

Kliniske beslutningsstøttesystemer i form af prognostiskemodeller har fået en større betydning og rolle i sundhedssystemet i løbet af de seneste år(79,84). Ligeledes er der blevet udgivet en betydelig mængde metodologisk litteratur og publikationer om hvordan prædiktionsmodeller udvikles, men størstedelen af de skabte prædiktionsmodeller lider af metodiske mangler(8). Metodiske mangler beskrives i litteraturen eksempelvis som manglende ekstern validering, for lille sample size til udvikling af modellen, overtilpasning til træningsdata og problemer med prædiktorer(12,80,85). Et studie af Lenert, Matheny og Walsh (2019) beskriver ydermere to øvrige problemer ved udviklingen af CDSS'er(13). Det første problem, populationsdatashift, skyldtes at CDSS'er over tid vil blive dårligere til at forudsige deres respektive udfald på grund af ændringer i data, som beskrevet i 2.2.3. Det andet problem, interventionsdatashift, skyldtes at CDDS'en medfører ændringer i arbejdsgangene, fordi brugerne handler på modellens forudsigelser, hvilket der ikke tages højde for i modeludviklingen og -opdateringen. Brugernes handlinger på CDDS'en prædiktion medfører, at fordelingen af udfaldene ændres fordi interventionen på baggrund af prædiktionen bliver den væsentligste prædiktor, hvilket medfører at sammenhængen mellem øvrige prædiktorer og udfald mindskes(13), se figur 2.



Figur 2: Illustration af interventionsdatashift. En velkalibreret CDSS forudsiger et udfald, som en kliniker handler ud fra. Dette resulterer i en ændring af det forudsagte udfald, hvilket medfører, at CDSS fejlkalibrerer.

2.2.5 ERVIN - et prognostisk beslutningsstøttesystem

Der findes flere forskellige beslutningsstøttesystemer indenfor THA og TKA, som forsøger at give en prognose på udfald som f.eks. patienttilfredshed, bevægelighed, postoperativ smerte, genindlæggelse, revision, PRO-data og perioperative komplikationer(86). ERVIN står for "Effekt og Risiko Ved INgreb" og er et dansk klinisk beslutningsstøttesystems, udviklet i Region Nordjylland med patientdata fra Farsø sygehus, til at hjælpe klinikere og patienter med at træffe en veloplyst beslutning om konservativ eller kirurgisk behandling. ERVIN er udviklet i samarbejde mellem KMD og Aalborg Universitetshospital(11,87). ERVIN indeholder flere prædiktorer, herunder PRO-data såsom Oxford Hofte-Score (OHS), Oxford Knæ-Score (OKS) og EQ-5D-3L samt patientdata som BMI, indlæggelser mv.(9,10). Spørgeskemaerne udfyldes af patienterne i en E-kiosk inden forundersøgelsen hos ortopædkirurgen og kombineres med data fra LABKA (elektronisk laboratorie system) og PAS (Patient administrativt system)(10). Udfaldene er opdelt i to kategorier 1) Funktion og livskvalitet, defineret som gevinst udfald og 2) Reoperation og korttidskomplikation for hofte og knæ samt luksation for hofte, der er defineret som risici udfald(10,88), se tabel 1. Se bilag 3: Risici udfald.

Gevinst udfald	Risici udfald
Funktion	Reoperation
Livskvalitet	Korttidskomplikationer Luksation

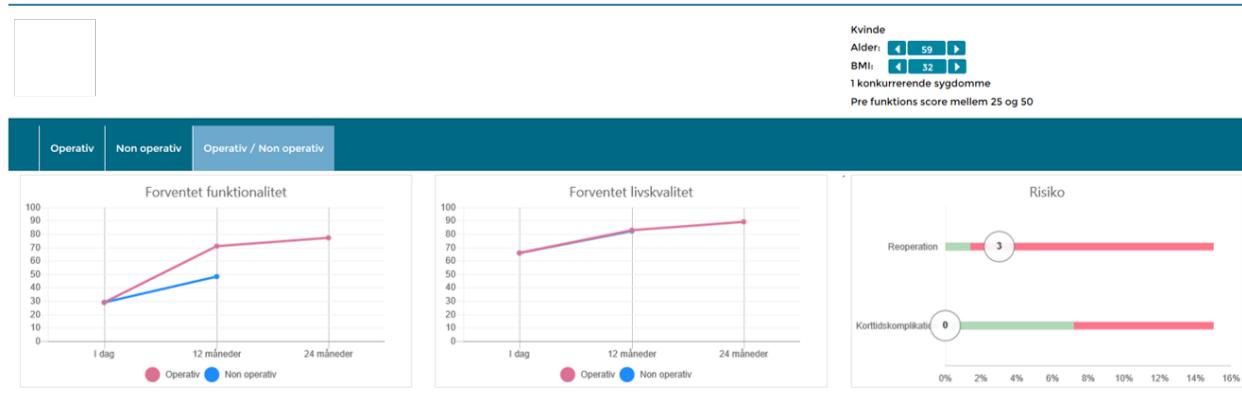
Tabel 1: opgørelse over udfald(10,88)

ERVIN sammenligner reallid-score på konsultationsdagen med et prædikteret gevinst udfald 12- eller 24 måneder efter kirurgisk eller konservativ behandling, ydermere prædikterer ERVIN et risici udfald for reoperation og korttidskomplikationer for både total hoftealloplastik og total knæalloplastik patienter samt luksationer for total hoftealloplastik patienter(11). På denne måde forsøges patientinddragelsen understøttet, således at patienter, der tilbydes enten operation eller konservativ behandling, er bekendt med de kliniske og livskvalitetsmæssige forventede udfald(10). Grafikken som forevises knæpatienten og kirurgen fremgår af Figur 3.

Nyt knæ - højre side

Patientens ønsker til effekten af indgribet

Baggrundsguppe



Hvad kan jeg forvente

Antal sammenlignelige patienter: 142

Konkurrerende sygdomme

Figur 3: Udfald fra en knæpatient fra ERVIN(89) som kirurgen og patienten forevises og kommunikerer ud fra. I venstre side af billedet ses forventet funktionalitet efter 12 og 24 måneder for henholdsvis TKA (rød) og 12 måneder for konservativ behandling (blå). I midten ses forventet livskvalitet efter 12 og 24 måneder for henholdsvis TKA (rød) og 12 måneder for konservativ behandling (blå). Til højre ses risici udfaldende for henholdsvis reoperation og kortidskomplikationer.

Det er dog uvist om ERVIN rent faktisk opfylder dette da den udelukkende er internt valideret. Et andet problem er at det aktuelt er uvist om ERVIN kan prædiktere de opstillede risici udfald(10,11).

3.0 Problemafgrænsning

Artrose er en degenerativ sygdom(1,2,14,15), som kan føre til komplekse smerteproblemetikker(1,17), og har alvorlige konsekvenser for patientens funktionsniveau og livskvalitet(1,16,18–24). Hofte og knæ er blandt de hyppigst ramte led(22), og behandlingen skal rettes mod at reducere smerte og forbedre funktion samt livskvalitet. Konservativ og kirurgisk behandling er begge effektive behandlingsmetoder med hver deres fordele og ulemper, men der er ingen klare nationale eller internationale retningslinjer for, hvornår operation bør overvejes frem for konservativ behandling(10,39,44). Kliniske undersøgelser og patienthistorien kan sammen med patientinddragelse understøtte beslutningstagningen, skabe meningsfulde forløb, forbedre efterlevelsen af instrukser og øge patienttilfredsheden(44,64,65,67). Prognostiske CDSS'er anvendes til at forbedre og målrette den kliniske viden samt give patientspecifik information, hvilket forbedre patientbeslutningsstøtten, kvaliteten og effektiviteten af sundhedsydelerne og patientresultaterne(74). Det er vigtigt at prognostiske CDSS'er har god ydeevne, og at der er udført en ekstern validering før CDSS'erne implementeres i klinisk praksis(8,77,82). Det er anerkendt, at der udvikles for mange nye modeller(12), og størstedelen af de skabte prædiktionsmodeller lider af metodiske mangler(8). Metodiske mangler beskrives i litteraturen eksempelvis som manglende ekstern validering, lille sample size til udvikling af modellen, overtilpasning til træningsdata, problemer med prædiktorer og datashift(76,80,81,82). Hvorvidt og i hvilket omfang disse begrænsninger eller andre begrænsninger beskrives af udviklerne af prognostiske CDSS'er er ikke beskrevet, ligeledes er det ikke undersøgt, hvilke begrænsninger udviklerne af CDSS'er beskriver. Et eksempel på et prognostisk CDSS, der anvendes i Region Nordjylland til at prædiktere udfald hos patienter med hofte- eller knæartrose ved konservativ og kirurgisk behandling er ERVIN(89). Det er dog uvist om ERVIN kan prædiktere risici udfaldene reoperation og korttidskomplikationer for både THA- og TKA-patienter samt luksationer for THA-patienter, da Ervin udelukkende er internt valideret(10,11).

4.0 Problemformulering

"Hvad beskriver litteraturen som begrænsninger ved udvikling af prognostiske kliniske beslutningsstøttesystemer, og hvor pålidelig er det kliniske beslutningsstøttesystem ERVIN til at prædiktere reoperation og korttidskomplikationer for både total hoftealloplastik og total knæalloplastik patienter samt luksationer for total hoftealloplastik patienter?"

5.0 Metode

Følgende afsnit omhandler de metodiske principper for dette projekt. Metoden er opdelt i henholdsvis et systematisk scoping review og et valideringsstudie af ERVIN.

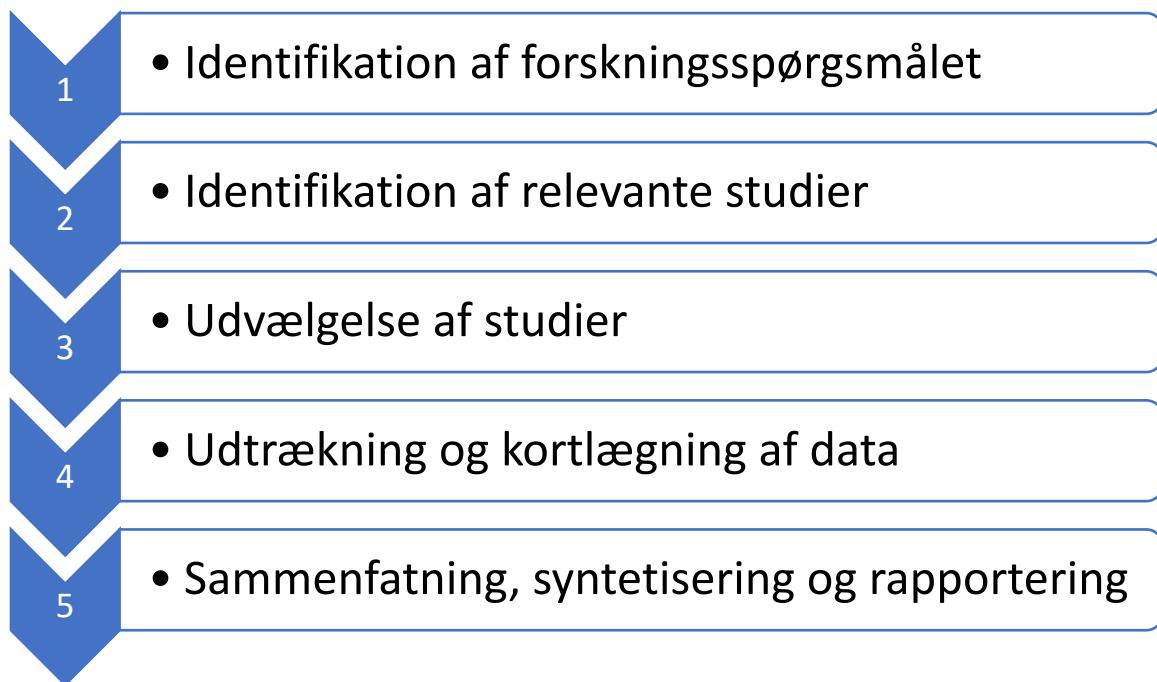
5.1 Scoping review

Scoping reviews er publikationer, der kortlægger dybden og bredden af litteraturen inden for et givent emnefelt. Gennem denne systematiske kortlægning beskriver forfatterne de vigtigste begreber, hovedkoncepter, teorier, emner, kilder og videnshuller indenfor emnefeltet. Scoping reviews er ofte drevet af brede, udforskende forskningsspørgsmål(90–93). I dette projekt vil hovedkonceptet omfatte analyse af selvopgjorte begrænsninger ved udviklingen af prognostiske kliniske beslutningsstøttesystemer.

5.1.1 Metodisk ramme for scoping review

For at undersøge den første del af problemformuleringen blev et systematisk scoping review gennemført med Arksey og O'Malley metodiske ramme, som indeholder fem trin. Det første trin var at identificerer forskningsspørgsmålet(94), som i dette projekt blev fremanalyseret gennem problemanalysen og formuleret som første del af problemformuleringen. Det andet trin var identificeringen af relevante studier(94), som blev behandlet under den systematiske søgning. Det tredje trin var udvælges af studier(94), der blev foretaget på baggrund af udvalgte in- og eksklusionskriterier. Fjerde trin var udtrækning og kortlægning af data fra de inkluderede studier(94), hvilket indebar kortlægning af centrale informationspunkter i forhold til forskningsspørgsmålet, som blev opstillet i figurer i resultatafsnittet og i tabeller i bilag. I dette projekt blev der anvendt en narrativ tilgang, der syntetiserede og konkretiserede de opstillede begrænsninger. Det femte og sidste trin indebar en systematisk og struktureret sammenfatning af de inkluderede

studier og de syntetiserede resultater, som blev rapporteret i form af metatekst, der bidrog til at besvare forskningsspørgsmålet(94). I dette projekt blev dette behandlet under resultatafsnittet. Den metodiske ramme fremgår af figur 4.



Figur 4: Metodisk ramme for scoping review inspireret af Arksey og O’Malley(94).

5.1.2 Systematisk litteratursøgning

For at undersøge emneområdet og besvare første del af problemformuleringen blev der udført systematiske søgninger i databaserne Pubmed, Embase og IEEE(95–97). Den systematiske søgning blev gennemført med emneord fra områdefeltet, der blev identificeret via ustrukturerede litteratursøgninger. Søgningen blev udformet som bloksøgning med to blokke for at øge søgningens sensitivitet(98). Den første blok omhandlede klinisk beslutningsstøtte, og anden blok omhandlede prædiktionsmodeller. Internt i blokkene blev søgeordene kombineret med den boolske operator OR, mens blokkene imellem blev kombineret med den boolske operator AND(98). Der blev anvendt både fritekstord og tesarus termer i begge blokke, hvilket havde til formål øge søgningssensitivitet. Der blev i søgningen anvendt trunkering for at udvide endelser og frasering for at ordrækker blev søgt i den angivne rækkefølge(98,99). Bloksøgningen i Pubmed fremgår af tabel 2. Bloksøgningerne i Embase og IEEE fremgår af Bilag 4: Søgestrenge.

Blok 1	Blok 2
"Clinical decision support system*"	"Prediction model*"
"Decision support systems, clinical"	"Prognostic model*"
CDSS	"Prediction rule*"
"Clinical Decision Support"	
"Medical decision support system*"	
"Decision-making tool for clinicians"	
"Clinical decision-making software"	
"Decision support technology"	
"Decision support technologies"	
"Clinical decision support software"	

Tabel 2: In- og eksklusioner for den systematiske litteratursøgning i Pubmed(100).

5.1.3 In- og eksklusionskriterier for den systematiske litteratursøgning

Inden behandling af litteraturen blev der opstillet in- og eksklusionskriterier, der var udvalgt med baggrund i problemanalysen, som havde til formål at sikre, at den inkluderede litteratur var relevant for besvarelse af problemformuleringen(101). Inklusionskriteriet prognostiske modeller blev defineret som systemer, der anvender patientspecifikke informationer til at prædiktere et givent udfald ved et bestemt scenarie(8). Klinisk beslutningsstøtte blev valgt for at sikre, at de prognostiske modeller blev anvendt som støtte til at træffe beslutninger i forhold til patientbehandling og forebyggelse(76). Inklusionen af begrænsninger ved udviklingen af prognostiske CDSS'er blev foretaget med henblik på at opnå en indsigt i de begrænsninger, der bliver rapporteret i forbindelse med udviklingen af prognostiske CDSS'er samt undersøge omfanget af disse rapporterede begrænsninger. Begrænsninger omfatter i dette projekt begrænsninger under skabelse, intern validering, ekstern validering, indvirkningsevaluering og opdatering af prognostiske modeller, se afsnit 2.2.2. Der blev kun inkluderet publicerede videnskabelige artikler for at sikre, at de inkluderede studier havde en vis kvalitet, da der ikke blev foretaget kvalitetsvurdering af de enkelte studier i dette projekt. Adgang til fuldtekst og sprogkravet blev opstillet for at sikre, at gruppens medlemmer kunne læse og forstå de inkluderede studier.

Protokoller blev ekskluderet, fordi denne type studier kunne indeholde teoretiske begrænsninger. Pilot CDSS'er blev ekskluderet, da det blev vurderet, at denne type studier kunne indeholde forventede begrænsninger, som ikke nødvendigvis er relaterede til det færdigudviklede CDSS. Diagnostiske CDSS'er

blev ekskluderet for ikke at inddrage CDSS'er, der var både prognostiske og diagnostiske. Review og metaanalyser blev ekskluderet af flere årsager. Først og fremmest ville inklusion af disse risikerer, at begrænsninger fra det samme studie kunne fremgå flere gange. Ydermere ville begrænsningerne rapporteret i reviews eller metaanalyser ikke være selvrapparterede, men en fortolkning af nye forfattere. Alle in- og eksklusionskriterier fremgår af tabel 3.

Inklusionskriterier	Eksklusionskriterier
Skal omhandle prognostiske modeller	Protokol CDSS
Skal inkludere klinisk beslutningsstøtte	Pilot CDSS
Opgør en eller flere begrænsninger ved udviklingen et prognostisk CDSS	Diagnostisk CDSS
Publiceret videnskabelige artikler	Reviews og metaanalyser
Fuldtekst tilgængelig	
Dansk, svensk, norsk eller engelsk	

Tabel 3: In- og eksklusionskriterier for den systematiske litteratursøgning.

5.1.4 Systematisk udvælgelse af litteratur

Efter den systematiske litteratursøgning i Pubmed, Embase og IEEE blev søgeresultaterne samlet i Zotero for at fjerne duplikationer(95–97,100). Herefter blev alle fremsøgte resultater overført til Rayyan, der anvendes til at behandle systematisk litteratur(102), som blev anvendt til sorterer artiklerne ud fra in- og eksklusionskriterier. Gruppens medlemmer vurderede på baggrund af titel og abstrakt, hvorvidt studiet skulle ”inkluderes”, ”måske inkluderes” eller ”ekskluderes”. I Rayyan blev ord relevante for inklusionskriterierne markeret med grønt, og ord relevante for eksklusionskriterierne markeret med rødt, hvilket bidrog til hurtigere og mere effektiv gennemgang af studierne. I denne del af den systematiske udvælgelse var gruppens medlemmer blindede for hinanden, hvilket sikrede at gruppens medlemmer ikke påvirkede hinanden i udvælgelsen af studier. Ved uoverensstemmelser og ”måske inkluderes” blev studierne gennemgået og diskuteret af hele gruppen ud fra opsatte in- og eksklusionskriterier, hvorefter studiet blev

"inkluderet" eller "ekskluderet". Efter gennemgang af titel og abstrakt blev de inkluderede studier vurderet for fuldtekst tilgængelighed, ikke tilgængelige studier blev ekskluderet. Efterfølgende blev de tilgængelige studiers relevans vurderet ved at fuldtekst læse diskussion og begrænsningsafsnit, hvilket blev udført med samme procedure som ved gennemgang af titel og abstrakt. Processen for den systematiske udvælgelse kan ses i figur 5.



Figur 5: Den systematiske udvælgelse af studier.

5.1.5 Udtrækning af data

Udvælgelsesprocessen af studier blev databehandlet i Rayyan(102). Udtrækningen af resultater foregik ved, at begge gruppemedlemmer separat nedskrev, hvilke begrænsninger den respektive litteratur opgjorde. Resultaterne for den manuelle nedskrivning blev sammenlignet, og studier uden selvopgjorte begrænsninger ved udviklingen blev ekskluderet. Efter manuel nedskrivning af begrænsninger i Rayyan blev inkluderet litteratur overflyttet til kodningsprogrammet NVIVO(103). Anvendelsen af NVIVO indebar kodning i hovedkoder og underkoder for specifikke identificerede temae ift. begrænsninger ved udvikling. Anvendelse af NVIVO sikrede en systematisk udtrækning af opgjorte begrænsninger ved udvikling i den inkluderede litteratur. Kodningen blev foretaget løbende med udgangspunkt i en datastyret tilgang(104). Efter dette trin blev alle koder og data revideret for at reducere antallet af koder og sikre, at kun relevant data blev medtaget. Koder med tætte relationer og lidt data blev samlet eller fjernet, og data som var blevet fejlkodet, blev rettet eller slettet. Processen for udtrækning af data fremgår af figur 6.



Figur 6: Den systematisk proces ved kodning

5.2 Temporal validering af ERVIN

For at besvare den anden del af problemformuleringen blev der udført en temporal validering, se bilag 2: Valideringstyper, hvor observerede og prædikterede data blev anvendt for henholdsvis reoperation og korttidskomplikationer for både THA- og TKA-patienter samt luksationer for THA-patienter.

5.2.1 Datakilde

Ved den temporelle validering blev der anvendt nyt data, der ikke var anvendt i udviklingen af ERVIN, som blev indsamlet fra januar 2020 til november 2022 på Ortopædkirurgisk Afdeling Aalborg Universitetshospital, Farsø. Afdelingen udfører blandt andet ortopædkirurgi på patienter med hofte- og knæartrose og modtager 3000 nye henvisninger til THA- eller TKA-konsultationer om året, hvoraf 1000 patienter får foretaget en operation. Patienterne henvises fra egen læge til konsultation med kirurgen om hvorvidt der skal udføres THA eller TKA(10). Data er indhentet fra BI og Analyse i Region Nordjylland og inkluderer data som BMI, alder, operationsdato, rygning, Oxford- hofte og knæscore, risikoestimater og køn på 125 THA- og TKA-patienter(105). Præsentation af demografisk data præsenteres i tabelform i resultatafsnittet.

5.2.2 Databehandling

Steyerberg og Vergouwes (2014) ABCD blev anvendt som metodisk ramme for udførelsen af den temporelle validering(8). De fire elementer (ABCD) omhandler kalibrering-i-det-store (A), kalibreringshældning (B), diskriminering (C) og beslutningskurveanalyse (D)(8). Resultaterne for kalibrering og diskriminering blev samlet og fremstillet i henholdsvis kalibreringskurver og receiver operating characteristic (ROC) for henholdsvis reoperation, korttidskomplikationer for både THA og TKA og luksationer for THA, se bilag 3: Risici udfald. Beslutningskurveanalysen blev præsenteret grafisk med en figur med respektive tærskels sandsynligheder. Til databehandling blev statistikprogrammet STATA 18, udviklet af StataCorp LLC., og kodepakkerne 'pmcalplot', 'roctab' og 'dca' anvendt(106).

5.2.3 Kalibrering-i-det-store (A)

Kalibrering henviser til nøjagtigheden af de forudsagte sandsynligheder, der produceres af modellen. En velkalibreret model vil producere forudsagte sandsynligheder, der nøje matcher de observerede proportioner af positive tilfælde i testdata(8). Kalibring-i-det-store, eller middelkalibreringen, er et gennemsnit

af, hvor tæt de prædikterede udfald er på de faktiske udfald. Kalibrering-i-det-store præsenteres i et kalibreringsplot, og perfekt kalibrering-i-det-store angives med 0(8,82).

5.2.4 Kalibreringshældningen (B)

Kalibreringshældningen henviser til spredning og om de forudsagte sandsynligheder varierer for meget (hældning mindre end 1) eller ikke varierer nok (hældning større end 1)(8,107)

En regressionshældning på 1 indikerer, at modellen er perfekt kalibreret, hvilket betyder, at de observerede og forudsagte værdier af responsvariablen har samme spredning. En regressionshældning på mindre end 1 indikerer, at modellen undervurderer forudsigelserne, mens en regressionshældning på mere end 1 indikerer, at modellen overvurderer forudsigelserne forudsat ens kalibrering-i-det-store/Y-akseskæring. En model der overvurderer kan føre til, at flere patienter anses for at have højere risiko for et givet udfald end de faktisk har, mens en model som undervurderer kan føre til, at patienter med højere risiko for et givent udfald ikke identificeres og ikke får den nødvendige behandling(8,82). Kalibreringshældningen præsenteres i et kalibreringsplot.

5.2.5 Diskriminering (C)

Diskriminering henviser til modellens evne til at skelne mellem to grupper, f.eks. patienter i høj risiko og lav risiko. Til at beskrive diskrimineringen ift. reoperation, korttidskomplikationer efter hofte- og knæalloplastik samt luksation efter hoftealloplastik blev der udarbejdet C-statistik, visualiseret som en "reciver operating characterisctic" (ROC)-kurve(8,108). C-statistikken beregner en prædiktionsmodels evne til at skelne mellem sande og falske positive resultater ved at beregne ROC-kurven(108). Dermed kan C-statistik fortolkes som sandsynligheden for, at et tilfældigt udvalgt subjekt, som har oplevet det undersøgte udfald, har større forudsagt sandsynlighed for at udfaldet forekommer, end et tilfældigt udvalgt subjekt, som ikke oplevede det undersøgte udfald(108). Arealet under ROC-kurven,"Area Under the Curve" (AUC), er en samlet betegnelse for hele modellens diskrimineringsevne. AUC er angivet med et tal mellem 0 og 1, hvor 1 repræsenterer en perfekt diskriminering, der altid skelner mellem sandt positive og sandt negative. 0,5 repræsenterer en model der tilfældigt inddeler sandt positive og sandt negative(8,81).

ROC-kurven en visualisering af sensitiviteten (sandt-positiv rate) overfor 1-specificiteten (falsk-positiv rate).

Sensitiviteten beregnes for en given tærskelværdi med følgende formel:

$$\text{Sensitivitet} = \frac{\text{Sandt positive}}{\text{Sandt positive} + \text{Falsk negtaive}}$$

Sensitivitet refererer til modellens evne til at opdage sandt positive tilfælde(8). Sensitivitet følger Y-aksen.

Specificitet beregnes for en given tærskelværdi med følgende formel:

$$(1 - \text{specificitet}) = \frac{\text{Falsk positive}}{\text{Falsk positive} + \text{Sandt negtaive}}$$

Specificitet refererer til modellens evne til at afvise sandt negative tilfælde(8). 1-specificitet følger X-aksen.

En model med god diskriminering vil producere høje forudsagte sandsynligheder for positive tilfælde og lave forudsagte sandsynligheder for negative tilfælde(8).

5.2.6 Beslutningskurveanalyse (D)

En beslutningskurve er en visuel fremstilling af en models "nettofordel" ift. en klinisk fastsat "tærskelsandsynlighed". Nettofordel er andelen af sandt positive, der identificeres ved en given tærskelværdi for et udfald. En nettofordel ved en given tærskelsværdi angiver, hvor stor gevinst eller skadevirkning der opnås ved en given behandlingsstrategi(109). En nettofordel på en given værdi angiver, hvor stor gevinst eller skadevirkning der opnås ved en given behandlingsstrategi. I dette projekt blev der anvendt beslutningskurve analyse til at vurdere den kliniske anvendelighed ved reoperation, korttidskomplikationer for både THA og TKA patienter og luksationer for THA patienter. Med beslutningskurve analysen blev der beregnet kliniske nettofordele for hvert af de undersøgte udfald, som blev sammenlignet med at behandle alle eller ingen patienter(8,109). Nettofordelene beregnes via flere tærskelsandsynligheder, som er minimumsandsynligheden for sygdom, som kræver yderligere behandling, med formlen(8):

$$\text{Nettofordel} = (\text{Sandt positiv} - \text{Tærskelssandsynlighed} * \text{Falsk positive})/N$$

Grænseværdierne for tærskelsandsynlighederne er fastsat ud fra notatet "Grænseværdier, cutpoints og transformering - hofte-knæ prædiktionsmodel". Tærskelsandsynlighederne er fastsat ud fra nationale registre og tilgængelige studier, ydermere er de drøftet af en gruppe regionale klinikere med ekspertviden indenfor emnet(88).

Grænseværdier for tærskelsandsynligheder	
Hofte reoperation	15 %
Hofte luksation	8 %
Hofte korttidskomplikation	15 %
Knæ reoperation	5 %
Knæ korttidskomplikation	15 %

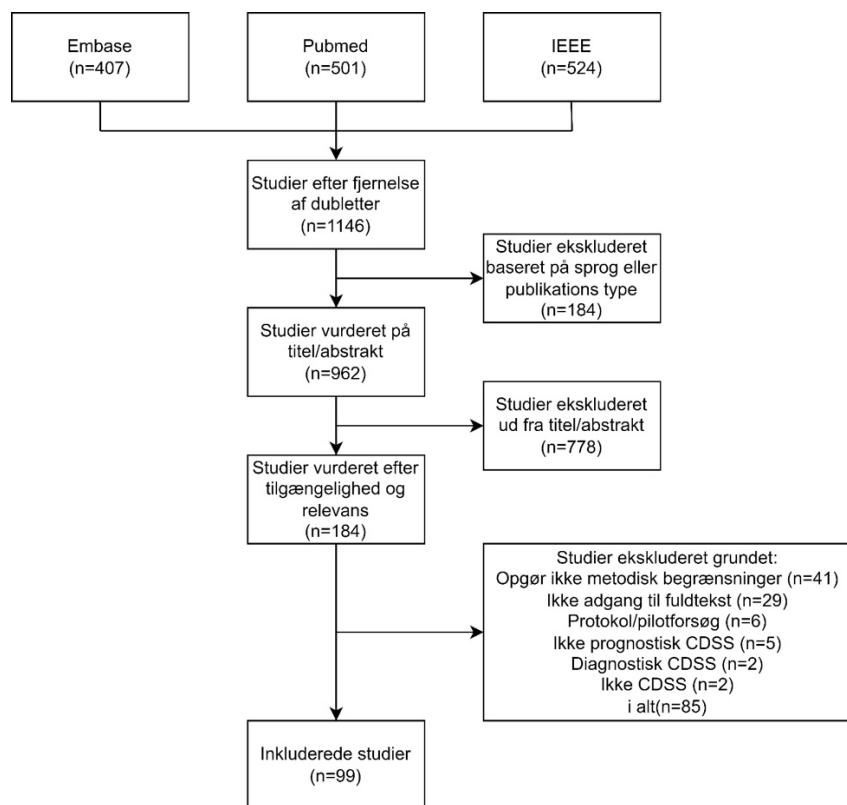
Tabel 4: Grænseværdier for tærskelsandsynlighederne ved reoperation, korttidskomplikationer for både hofte og knæ og luksationer for hofte.

6.0 Resultater

I følgende afsnit præsenteres resultaterne af dette speciale. Først præsenteres resultaterne for det gennemførte scoping review omhandlende de metodologiske begrænsninger ved prognostiske CDSS'er. Efterfølge ned præsenteres valideringen af ERVIN.

6.1 Resultaterne for scoping review:

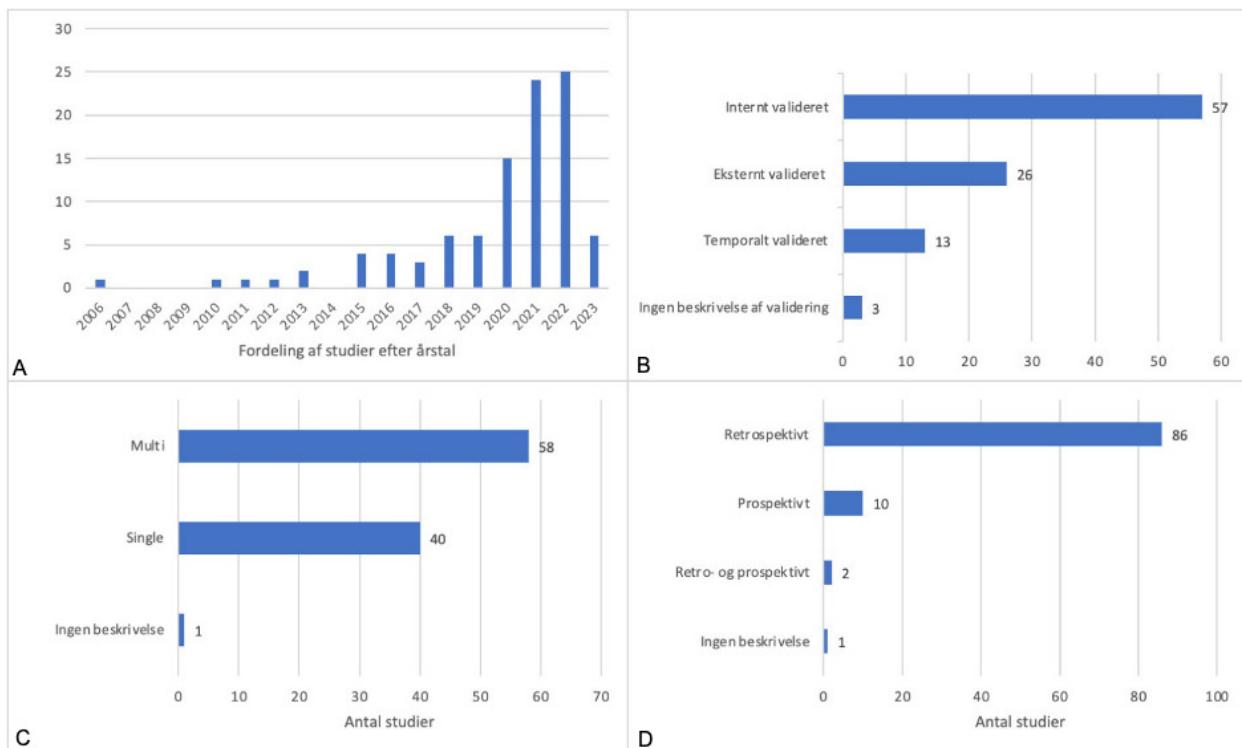
Der blev fremsøgt 1432 artikler, hvoraf 501 kom fra Pubmed, 407 kom fra Embase og 524 kom fra IEEE(95–97). Efter fjernelse af 286 dubletter blev 1146 studier vurderet med hensyn til publikationstype og sprog. Ved vurdering på titel og abstraktniveau blev 778 artikler ekskluderet, hvilket resulterede i, at 184 artikler blev vurderet på relevans og fuldtekst-tilgængelighed. I alt er 99 studier inkluderet i dette scoping review, se bilag 5: Inkluderede studier. En oversigt over hele processen kan ses i figur 7, som viser et flowchart over resultaterne af det gennemførte scoping review.



Figur 7: Flowchart over resultaterne for det gennemførte scoping review.

6.1.1. Inkluderede studier

I alt er 99 studier inkluderet i dette scoping review, studierne daterer fra 2006 til 2023, se figur 8 (A), og sample size varierer fra 72 til 721.339, se bilag 5: Inkluderede studier. 57 af studierne er internt valideret, 26 er eksternt valideret, 13 er temporalt valideret og tre beskriver ikke validering, se figur 8 (B). I 58 af studierne er data indsamlet multiple steder (multi), 40 af studierne er data indsamlet på en enkelt lokalitet (single) og et studie beskriver ikke lokalitet for indsamling af data, se figur 8 (C). I studierne er data indsamlet retrospektivt i 86 tilfælde, prospektivt i 10 tilfælde, både retro- og prospektivt i to tilfælde og et studie angiver ikke dataindsamlingsmetode, se figur 8 (D). Alle inkluderede studier fremgår af bilag 5: Inkluderede studier.

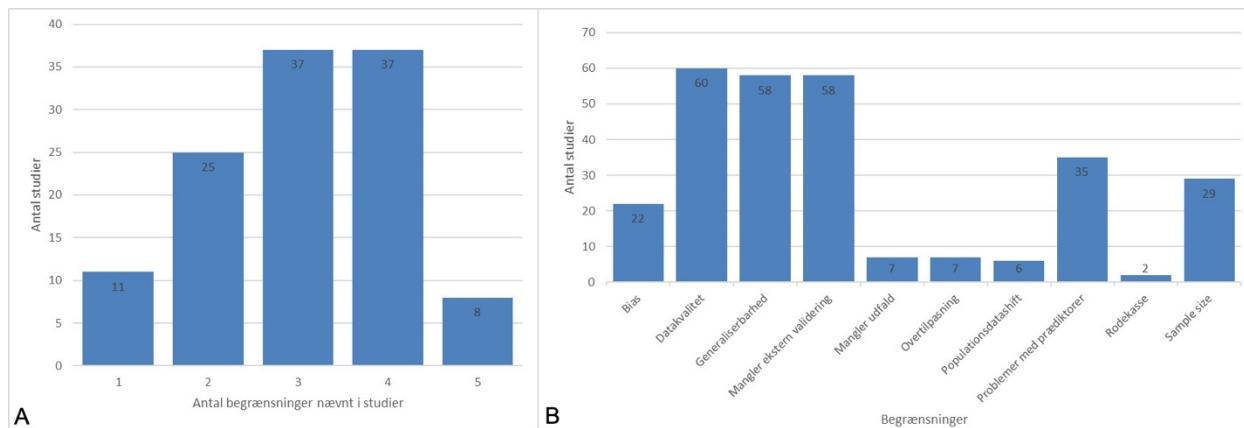


Figur 8: Oversigt over udgivelsesår for inkluderede studie (A), inkluderede studiers validering (B), inkluderede studiers data-opsamling lokalitet (C) og inkluderede studiers dataopsamlingsmetode (D). Figurer er konstrueret med data fra bilag 5: Inkluderede studier.

6.1.2 Selvopgjorte begrænsninger ved udviklingen i inkluderede studier

11 studier opgjorde en begrænsning ved udviklingen, 25 opgjorde to begrænsninger ved udviklingen, 37 studier opgjorde tre begrænsninger ved udviklingen, 37 studier opgjorde fire begrænsninger ved udviklingen og otte studier opgjorde fem begrænsninger ved udviklingen, se figur 9 (A). De enkelte studiers begrænsninger ved udviklingen er opgjort i bilag 6: Begrænsninger ved udviklingen i de enkelte studier. I

kodningen af studierne er der identificeret 10 hovedkoder med i alt 13 underkoder. Fordelingen af begrænsninger ved udviklingen er opgjort i figur 9 (B). Koderne og underkoderne er præsenteret under hvert af de følgende afsnit.



Figur 9: Antallet af begrænsninger ved udviklingen i alle 99 inkluderede studier (A) samt fordelingen af begrænsninger ved udviklingen (B). Figurer er konstrueret med data fra bilag 6: Begrænsninger ved udviklingen i de enkelte studier.

Bias

Hovedkoden 'bias' indeholder uspecificerede oplyste bias samt underkoder som 'tidsvindue-bias', 'udvægelseselsesbias', 'confounders' og 'svarbias'. 22 af de 99 inkludererede studier rapporterer bias. Udvægelsesbias rapporteres i 13 studier, confounders rapporteres i 5 studier og andre former for bias rapporteres i mindre omfang. I 19 studier er data indsamlet retrospektivt og i tre tilfælde er data indsamlet prospektivt. I seks af studierne er data indsamlet på en enkelt lokalitet og i 16 tilfælde er data opsamlet på multiple lokaliteter. I 11 af studierne er modellerne internt valideret, i tre tilfælde er modellen temporalt valideret og i otte tilfælde er modellen eksternt valideret. Selvrapportererde bias indikerer ikke et tydeligt mønster ift., hvor og hvordan data er indsamlet samt hvordan det er valideret.

Datakvalitet

Hovedkoden 'datakvalitet' indeholder problemer med datakvaliteten, herunder misklassificering, ubalance data og upræcist data, samt underkoderne 'manglende data' og 'retrospektiv data'. Manglende data inkluderer både ufuldendte dataset og insufficient datamængde til at inkludere relevant prædiktor. I 60 studier rapporteres begrænsninger relateret til datakvaliteten, i 22 studier rapporteres begrænsninger relateret til datakvaliteten, i 29 studier rapporteres manglende data og i 30 studier er retrospektivt indsamling af data rapporteret som en begrænsning. I 56 studier er data indsamlet retrospektivt, og i to studier er data indsamlet prospektivt. I 22 studier er data indsamlet på en lokalitet og i 36 studier er data

indsamlet multiple steder. 35 studier er internt valideret, seks studier er temporalt valideret, 16 er eksternt valideret og et studie beskriver ikke valideringen. Selvrapporterede problemer med datakvaliteten indikerer en overvægt af studier, hvor data var indsamlet retrospektivt, men der indikeres ikke et tydeligt mønster med hensyn til, hvor data er indsamlet samt hvordan det er valideret.

Generaliserbarhed

Hovedkoden 'generaliserbarhed' indeholder problemer med generaliserbarhed af modellen, samt underkoden 'geografisk generaliserbarhed', som omfatter at modellen kun er generaliserbar indenfor et afgrænset geografisk område, og 'population specifik generaliserbarhed', som omfatter at modellen kun er generaliserbar på en specifik population. 58 studier rapporteres begrænsninger relateret til generaliserbarheden, i 38 studier rapporteres begrænsninger relateret til geografisk generaliserbarhed og i 20 studier rapporteres begrænsninger relateret til populations specifik generaliserbarhed. I 51 studier er data indsamlet retrospektivt, i seks studier er data indsamlet prospektivt og i et studie er data indsamlet både retro- og prospektivt. I 29 studier er data indsamlet på en enkelt lokalitet, i 28 studier er data opsamlet multiple steder og et studie beskriver ikke datalokalitet. 38 studier er internt valideret, otte er temporal valideret, 10 er eksternt valideret, og to studier beskriver ikke valideringen. Der rapporteres en overvægt af problemer med generaliserbarheden ved studier, hvor valideringen er foretaget internt, men der indikeres ikke et tydeligt mønster ift., hvor og hvordan data er indsamlet.

Valideringsbegrænsninger

Hovedkoden 'valideringsbegrænsninger' indeholder problemer relateret til manglende ekstern validering og underkoden 'begrenset ekstern validering', som omfatter begrænsninger relateret til den temporelle eller eksterne validering. 58 studier rapporterer begrænsninger relateret til validering, i 45 studier rapporteres manglende ekstern validering og i 13 studier rapporteres begrenset ekstern validering. I 50 studier er data indsamlet retrospektivt, i seks studier er data indsamlet prospektivt, i et studie er data indsamlet både retro- og prospektivt og i et studie er der ingen beskrivelse af datalokalitet. I 24 studier er data indsamlet på en lokalitet og i 34 studier er data indsamlet multiple steder. 35 studier er internt valideret, ni studier er temporalt valideret, 12 er eksternt valideret og to studier beskriver ikke valideringen. Selvrapporterede valideringsbegrænsninger indikerer ikke et tydeligt mønster ift., hvor og hvordan data er indsamlet samt hvordan det er valideret.

Manglende udfald

Hovedkoden 'manglende udfald' indeholder begrænsninger relateret til manglende udfald. Syv studier rapporterer manglende udfald, og i alle studier er data indsamlet retrospektivt. I tre studier er data indsamlet på en lokalitet, og i fire studier er data indsamlet på multiple lokaliteter. To studier er internt valideret, tre er temporalt valideret og to er eksternt valideret. Selvrapporterede manglende udfald indikerer ikke et tydeligt mønster ift., hvor og hvordan data er indsamlet samt hvordan det er valideret.

Overtilpasning

Hovedkoden 'overtilpasning' indeholder problemer relateret til overtilpasning ved udvikling af modellen. Syv studier rapporterer begrænsninger med overtilpasning. I fem studier er data indsamlet retrospektivt, i et studie er data indsamlet prospektivt, og i et studie er data indsamlet både retro- og prospektivt. I et studie er data indsamlet på en lokalitet, og i seks studier er data indsamlet multiple steder. Et studie er internt valideret, to er temporalt valideret, og fire er eksternt valideret. Der rapporteres en overvægt af problemer med rapportering af overtilpasning ved studier, hvor valideringen er foretaget temporalt og eksternt samt hvor data er indsamlet på multiple lokaliteter, men der indikeres ikke et tydeligt mønster ift., hvordan data er indsamlet.

Populationsdatashift

Hovedkoden 'populationsdatashift' indeholder problemer relateret til ændringer i modellen over tid grundet ændringer i populationens sammensætning. Seks studier rapporterer begrænsninger med populationsdatashift. I fem studier er data indsamlet retrospektivt, og i et studie er data indsamlet prospektivt. I to studier er data indsamlet på en lokalitet, og i fire studier er data indsamlet multiple steder. To studie er internt valideret, to er temporalt valideret, og to er eksternt valideret. I studier med selvrapporterede problemer med populationsdatashift indikeres ikke et tydeligt mønster ift., hvor og hvordan data er indsamlet samt hvordan det er valideret

Problemer med prædiktorer

Hovedkoden 'problemer med prædiktorer' indeholder problemer relateret til de prædiktorer modellen anvender til at forudsige et udfald og underkoderne 'for mange prædiktorer' og 'vægtning af prædiktorer'. 'For mange prædiktorer' henviser til at studiet har inkluderet potentielt ikke vigtige eller forkerte prædiktorer i modellen og 'vægtning af prædiktorer' henviser til at forholdet mellem prædiktorerne potentielt er vægtet forkert. 35 studier rapporterer problemer med prædiktorer. 32 af studierne rapporterer manglende eller potentielt manglende prædiktorer såsom, at studierne kunne have ønsket at indsamle andre

prædiktorer. To studier rapporterer for mange eller muligvis for mange prædiktorer, og i et studie rapporterer begrænsninger ved vægtningen af anvendte prædiktorer. I 29 studier er data indsamlet retrospektiv og i seks studier er data indsamlet prospektivt. I 15 studier er data indsamlet på en lokalitet, og i 20 tilfælde er data indsamlet på multiple lokaliteter. 16 studier er internt valideret, syv studier er temporalt valideret, 11 er eksternt valideret, og et studie beskriver ikke valideringen. I studier med selvrapporтерede problemer med prædiktorer indikeres ikke et tydeligt mønster ift., hvor og hvordan data er indsamlet samt hvordan det er valideret.

Sample size

Hovedkoden 'sample size' indeholder problemer med relateret til sample size og er underkodet 'sample size skabelsesdata' og 'sample size valideringsdata'. 29 studier rapporterer begrænsninger relateret til sample size. 26 studier rapporterer begrænsninger til underkoden sample size på skabelsessample, og fem studier rapporterer begrænsninger til underkoden sample size valideringsdata. I 23 studier er data indsamlet retrospektivt, i fem studier er data indsamlet prospektivt, i et studie er data indsamlet både retro- og prospektivt. I 13 studier er data indsamlet på en lokalitet, og i 16 tilfælde er data indsamlet multiple steder. 17 studier er internt valideret, fire studier er temporalt valideret, syv er eksternt valideret, og et studie beskriver ikke valideringen. I studier med selvrapporтерede problemer med sample size indikeres ikke et tydeligt mønster ift., hvor og hvordan data er indsamlet samt hvordan det er valideret. Sample size for studier, der rapporterede begrænsninger relateret til sample size, havde en gennemsnits sample size på 6099 (± 16.353) sammenlignet med en gennemsnits sample size på 34981 (± 100.155) ved studier, der ikke rapporterede begrænsninger relateret til sample size.

6.2 Temporal validering af ERVIN

I følgende afsnit præsenteres resultaterne for kalibrering-i-det-store, kalibreringshældning, diskriminering og beslutningskurveanalyse for reoperation og korttidskomplikationer for THA og TKA, se bilag 3: Risici udfald. ERVIN beregnede et gennemsnitligt risikoestimat på 0,025 for hoftepatienters tendens til at opleve hofteluksationer efter THA, svarende til, at 2,52% af patienterne risikerer at få en hofteluksation. Ingen af patienterne oplevede en hofteluksation i valideringsdatasættet, hvilket gør det umuligt at foretage analyser.

6.2.1 Demografi på valideringsdata

Valideringen af ERVIN blev foretaget på data fra 125 patienter, der alle havde fået foretaget enten en THA- eller TKA-operation. Mediantiden mellem baseline og follow-up var på 410 dage for THA og 435 dage for TKA, 5 THA- og 20 TKA-patienter havde ikke follow-up data. Den deskriptiv beskrivelse af valideringsdata ses i tabel 5, herunder demografi, rygning, gennemsnitsværdier på BMI og baseline for Oxford hofte- og knæscore.

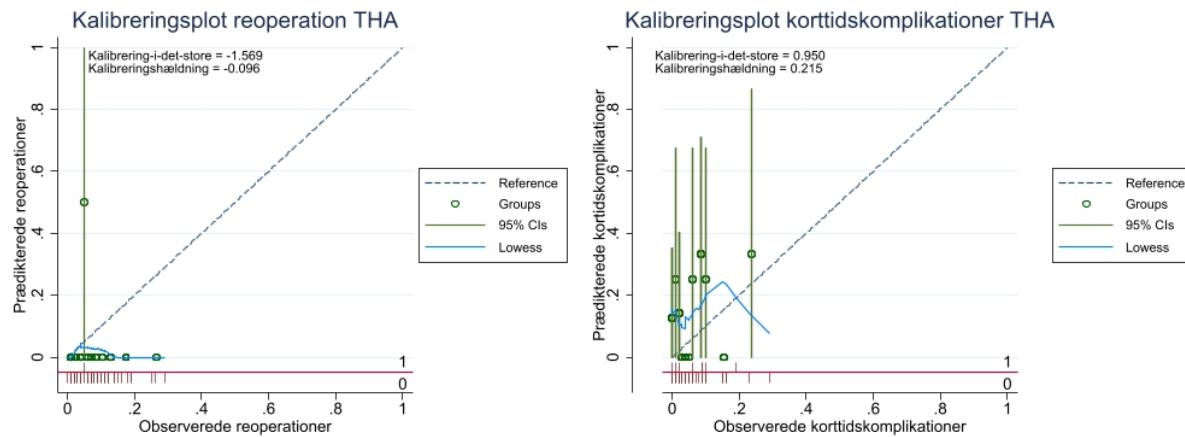
	<i>Frekvens</i>	<i>Procent</i>
Hoftepatienter		
Mænd	26	48,1 %
Kvinder	28	51,9 %
Total	54	100 %
Rygning (ja/nej)	7/47	13 %/87 %
<i>Gennemsnit</i>		
BMI	27,10 ($\pm 4,6$)	
Oxford hoftescore baseline	21,22 ($\pm 7,1$)	
Knæpatienter		
Mænd	27	38 %
Kvinder	44	62 %
Total	71	100 %
Rygning (ja/nej)	8/63	11,3 %/ 88,7 %
<i>Gennemsnit</i>		
BMI	30,39 ($\pm 5,4$)	
Oxford knæscore baseline	20,89 ($\pm 6,7$)	

Tabel 5: Deskriptiv beskrivelse af valideringsdata herunder demografi, rygning, gennemsnitsværdier på BMI og Oxford baseline for hofte og knæ-patienter.

6.2.2 Validering af THA-prædiktioner

Kalibrering

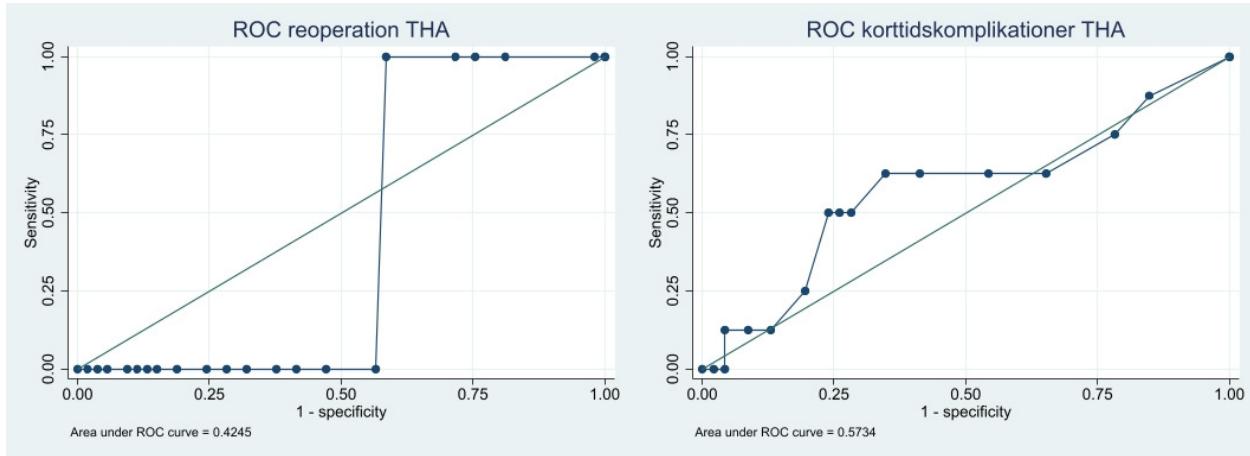
Kalibreringen for henholdsvis reoperation og korttidskomplikationer efter THA ses i figur 10. For reoperation er kalibrering i det store (A) -1,57 og kalibreringshældningen (B) er -0,1. For korttidskomplikationer er kalibrering i det store (A) 0,95 og kalibreringshældningen (B) er 0,22.



Figur 10: Kalibreringsplot reoperationer og korttidskomplikationer THA; Observerede risici udfald er repræsenteret på x-aksen, mens de prædikterede risici udfald er på y-aksen, den stippled linje repræsentere perfekt kalibrering med kalibrering i det store på 0 og en hældning på 1, den blå linje repræsenterer modellens kalibrering. De grønne cirkler viser deciler af patienternes gennemsnits estimerede risiko med 95% konfiden-sinterval. Den røde linje beskriver fordelingen af faktiske udfald, hvor 0 er ikke oplevet og 1 er oplevet.

Diskriminering

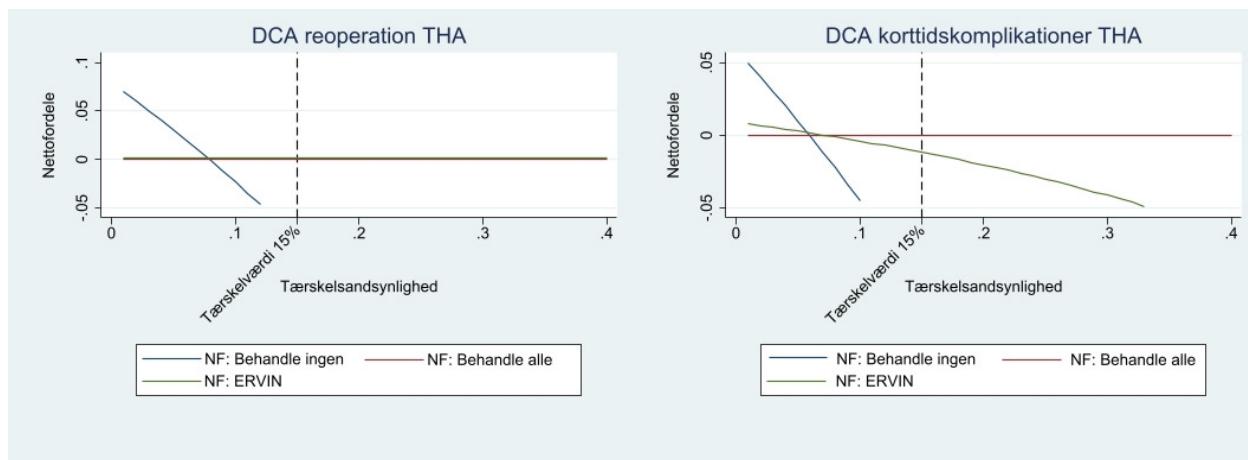
Diskrimineringen for reoperation og korttidskomplikationer efter THA er beregnet og illustreret i ROC kurver i figur 11. For reoperation forevises ROC-Kurven og AUC er 0,42 [95% konf. interval 0,1] og for korttidskomplikationer forevises ROC-Kurven og AUC er 0,57 [95% konf. interval 0,33, 0,82].



Figur 11: ROC reoperation og korttidskomplikationer THA; 1-specificitet (falsk-positiv rate) er repræsenteret på x-aksen og sensitiviteten (sandt-positiv rate) er repræsenteret på y-aksen, ROC-kurven er indtegnet med den blå linje. Den grønne linje repræsenterer tilfældig inddeling af patienterne. AUC beskriver modellens evne til at diskrimineringsevne.

Beslutningskurve analyse

Beslutningskurver for reoperation og korttidskomplikationer for hofte, med tilhørende tærskelværdier, er beskrevet i figur 12. Tærskelværdierne for både reoperation og korttidskomplikationer er på 15 %(88). Ved reoperation viser beslutningskurve analysen, at der ikke opnås en nettofordel ved at anvende ERVIN fremfor at operere alle ved den fastsatte tærskelværdi. Ved korttidskomplikationer viser beslutningskurve analysen, at den største nettofordel opnås ved tærskelværdien, hvis alle bliver opereret. Nettofordelen ved anvendelsen af modellen er negativ, da flere patienter får korttidskomplikationer ved tærskelværdien end hvis alle blev opereret.

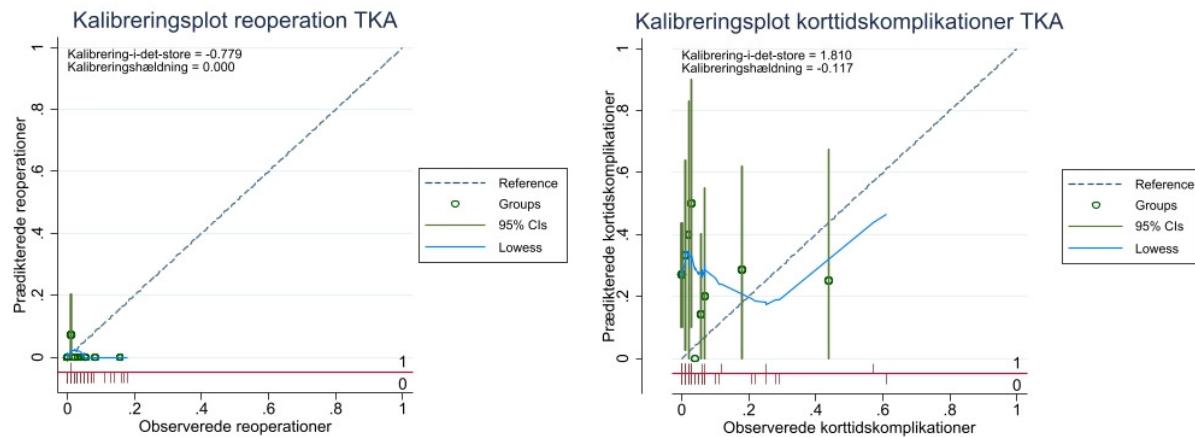


Figur 12: DCA reoperation og korttidskomplikationer THA; beslutningskurven beskriver tærskelsandsynligheden (x-aksen) over for Nettofordele (NF) (y-aksen). Blå linje beskriver NF, hvis ingen behandles, rød linje beskriver NF, hvis alle behandles og den grønne linje beskriver NF, hvis ERVIN anvendes, stiplede linje markerer tærskelværdien for reoperation og korttidskomplikationer efter THA på 15%.

6.2.3 Validering af TKA-prædiktionser

Kalibrering

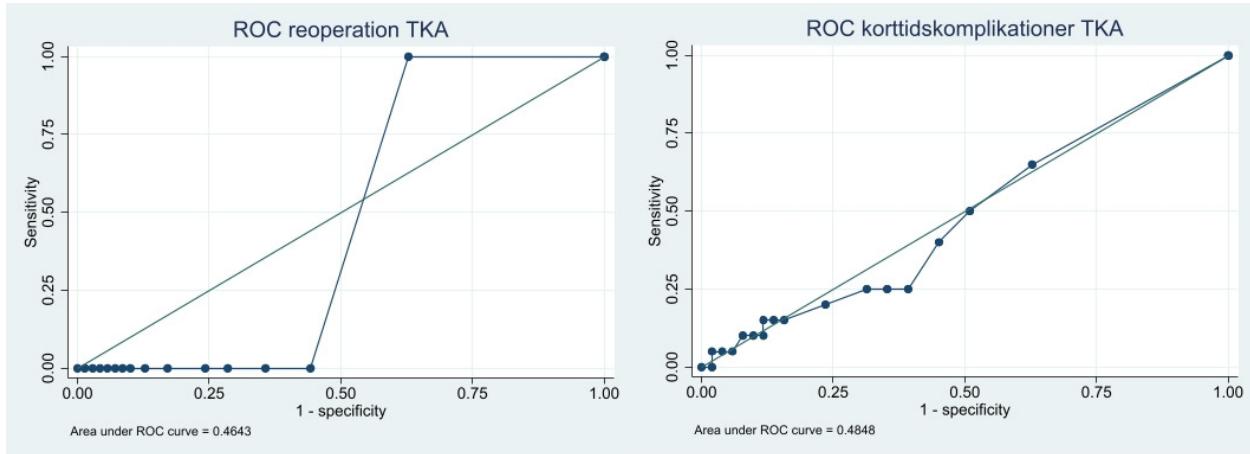
Kalibreringen for henholdsvis reoperation og korttidskomplikationer efter TKA ses i figur 13. For reoperation er kalibrering i det store (A) -0,78 og kalibreringshældningen (B) er 0. For korttidskomplikationer er kalibrering i det store (A) 1,81 og kalibreringshældningen (B) -0,117



Figur 13: Kalibreringsplot reoperationer og korttidskomplikationer TKA; Observerede udfald er repræsenteret på x-aksen, mens de prædikterede risici udfald er på y-aksen, den stippled linje repræsentere perfekt kalibrering med kalibrering i det store på 0 og en hældning på 1, den blå linje repræsenterer modellens kalibrering. De grønne cirkler viser deciler af patienternes gennemsnits estimerede risiko med 95% konfidenstinterval. Den røde linje beskriver fordelingen af faktiske udfald, hvor 0 er ikke oplevet og 1 er oplevet.

Diskriminering

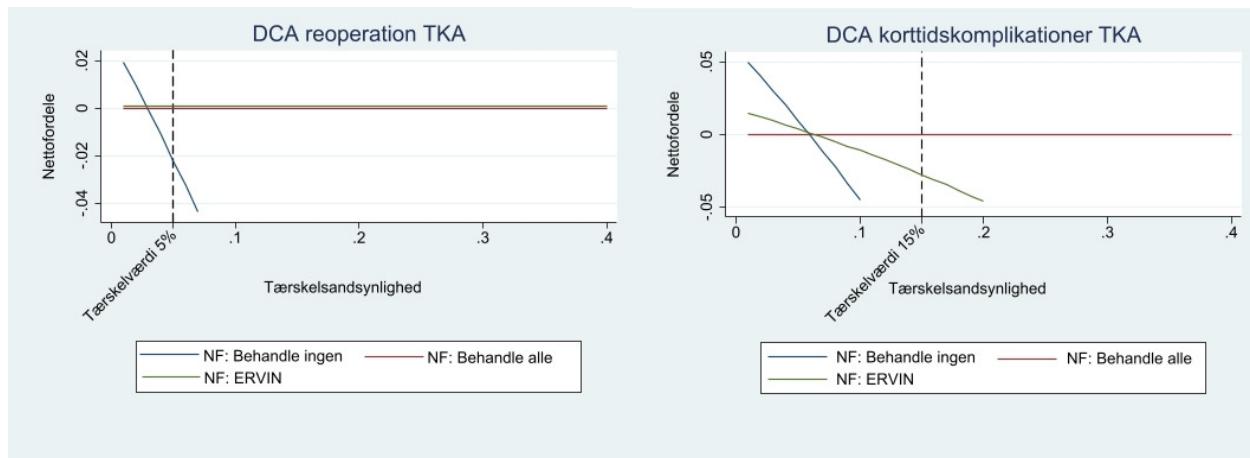
Diskrimineringen for reoperation og korttidskomplikationer efter TKA er beregnet og illustreret i ROC kurver i figur 14. For reoperation forevises ROC-Kurven og AUC er 0,46 [95% konf. interval 0 1]. For korttidskomplikationer forevises ROC-Kurven og AUC er 0,48 [95% konf. interval 0,34 0,63].



Figur 14: ROC reoperation og korttidskomplikationer TKA; 1-specificitet (falsk-positiv rate) er repræsenteret på x-aksen og sensitiviteten (sand-positiv rate) er repræsenteret på y-aksen, ROC-kurven er indtegnet med den blå linje. Den grønne linje repræsenterer tilfældig inddeling af patienterne. AUC beskriver modellens diskrimineringsevne.

Beslutningskurve analyse

Beslutningskurver for reoperation og korttidskomplikationer for knæ, med tilhørende tærskelværdier, er beskrevet i figur 15. Tærskelværdierne for reoperation er på 5% og korttidskomplikationer på 15%(88). Ved reoperation viser beslutningskurve analysen, at der ikke opnås en nettofordel ved at anvende ERVIN fremfor at operere alle ved den fastsatte tærskelværdi. Ved korttidskomplikationer viser beslutningskurve analysen, at den største nettofordel ved tærskelværdien opnås ved at operere alle. Nettofordelen ved anvendelsen af modellen er negativ da flere patienter får korttidskomplikationer ved tærskelværdien end hvis alle blev opereret.



Figur 15: DCA reoperation og korttidskomplikationer TKA; beslutningskurven beskriver tærskelsandsynligheden (x-aksen) over for Nettofordel (NF) (y-aksen). Blå linje beskriver NF, hvis ingen behandles, rød linje beskriver NF, hvis alle behandles og den grønne linje beskriver NF, hvis ERVIN anvendes, stiplede linje markerer tærskelværdien for reoperation efter TKA på 5% og 15% ved korttidskomplikationer.

6.3 Opsummering af resultater

I scoping review blev 99 studier inkluderet hvori der var opgjort 355 begrænsninger ved udviklingen fordelt på bias, datakvalitet, generaliserbarhed, manglende ekstern validering, manglende udfald, overtilpasning, populationsdatashift, problemer med prædiktorer, rodekasse og sample size.

Ved den temporale validering viste resultaterne ved korttidskomplikationer for THA, at ERVIN overvurderede risikoen og den skelnede marginalt dårligere end tilfældigt mellem oplevede og ikke oplevede ricisi udfald, beslutningskurveanalyse viste at det var bedre at opererer alle end at anvende ERVIN. Ved reoperation for THA undervurderede ERVIN risikoen og den skelnede marginalt dårligere end tilfældigt mellem oplevede og ikke oplevede ricisi udfald, beslutningskurveanalyse viste ingen nettofordel ved at anvende ERVIN. Ved korttidskomplikationer for TKA overvurderede ERVIN risikoen og den skelnede marginalt dårligere end tilfældigt mellem oplevede og ikke oplevede ricisi udfald, beslutningskurveanalyse viste at det var bedre at opererer alle end at anvende ERVIN. Ved reoperation for TKA undervurderede ERVIN risikoen og den skelnede marginalt dårligere end tilfældigt mellem oplevede og ikke oplevede ricisi udfald, beslutningskurveanalyse viste ingen nettofordel ved at anvende ERVIN.

7.0 Diskussion

Følgende afsnit omhandler diskussionen af dette projekt. Første del er en diskussion af resultaterne fra henholdsvis scoping review omhandlede begrænsninger ved udviklings af prognostiske CDSS'er og validering af det prognostiske CDSS ERVIN. Anden del er en diskussion af metoderne for henholdsvis scoping review og validering af ERVIN.

7.1 Resultatdiskussion

7.1.1 Resultatdiskussion af scoping review

I 22 ud af 57 internt validerede studier blev det ikke angivet som en begrænsning ved udvikling, at studiet manglede ekstern validering. Der kan diskuteres om dette tal er for højt, da prædiktionsmodeller generelt præsterer dårligere ved ekstern validering end ved intern validering(8). Ydermere er ekstern validering i nogen form nødvendig for at vurdere generaliserbarhed, hvorfor alle studier der udelukkende er intern validerede bør rapportere manglende ekstern validering som en begrænsning ved udviklingen(8,110). Resultaterne i dette scoping review indikerer dog at flere anvende ekstern validering end tidligere da 39 ud af 99 studier (39,39%) er eksternt eller temporalt valideret i modsætning til 5% af studierne i studiet af Ramspek, Jager, Dekker et al. (2020)(110). Af de 26 eksternt validerede og 13 temporalt validerede rapporterede ni eksternt validerede studier(111–119) og tre temporalt validerede studier(120–122) at deres eksterne validering er begrænset, mens de resterende ikke angiver nogle begrænsninger ved deres eksterne validering. Dette kan tyde på at der er en forskellighed i forståelsen af ekstern validering og hvornår et prognostisk CDSS er valideret succesfuldt. Ifølge Altman, Vergouwe, Royston et al. (2009) er temporal validering ikke en sand ekstern validering, men type validering mellem intern og sand ekstern validering(85), omvendt beskriver studiet af Steyerberg og Vergouwe (2014), at temporal validering også er en ekstern validering, men at den eksterne validering er stærkere, hvis den udføres med data fra en anden lokation og tid(8). Der kan dog godt argumenteres for at en temporal validering i nogle tilfælde kan være mere meningsfuld end en fuldstændig ekstern validering. Hvis en algoritme skal bruges på én lokalitet på én population, er det ikke hensigtsmæssigt at teste modellen en anden lokalitet på en anden population. Ifølge Van Calster, Steyerberg, Wynants et al. (2023) så kan prædiktions modeller aldrig kaldes succesfuldt valideret fordi der konstant sker ændringer i populationens sammensætningen, prædiktorer, udfald og målemetoder(123). Eksempelvis nævner ingen af de tre Covid-19 studier(111,119,124) populationsdatashift som et metodisk problem, på trods af denne population er kendtegnet ved hurtige ændringer i både sygdomskarakteristika og befolkningsdemografi(125).

Problemet vedrørende CDSS'ers selvødelæggende natur med interventionsbias opgørs bemærkelsesværdigt ikke i et eneste studie(13). Eksempelvis vil en effektiv forudsigelse af delir medfører et klinisk respons der forhindre delir i at opstå, hermed bliver forudsigelsen om delir en falsk positiv. At denne begrænsning ved udviklingen i af CDDS'er ikke blev nævnt i studier før 2019, kan forklares med, at de pågældende studier blev offentliggjort før udgivelsestidspunktet for før omtalte studie. Men 70 ud 99 inkluderede studier er udgivet i 2020 og fremefter, kan det indikere, at dette metodologiske problem ikke har fået tilstrækkelig opmærksomhed fra udviklere af CDSS'er.

De fleste studier opgør enten to, tre eller fire begrænsninger ved udviklingen af deres respektive CDSS. Et større antal selvopgjorte begrænsninger ved udviklingen er dog langt fra ensbetydende med at det er en ringere model. At forfatterne er opmærksomme på begrænsninger ved udviklingen, kan ses som en ressource da der så kan effektueres på at begrænse disse. Et eksempel er at kun syv studier opgør overtilpasning som en begrænsning ved udviklingen(111,114,126–130), og af disse studier er seks ekstern valideret(111,114,127–130). Dette indikerer at overtilpasning som begrænsning potentielt er underrapporteret blandt internt validerede studier og udviklere ikke med sikkerhed afgøre, om en model er overtilpasset, før der er foretaget en ekstern validering(110).

Tilsvarende typer af begrænsninger blev gjort opgjort på forskelligvis i de respektive studier. Et eksempel på dette er ”manglende generaliserbar”, hvilket kan fortolkes på flere måder. Med manglende generaliserbarhed menes der, at modellen kan bruges til patienter i en anden lokation, men det samme kan skrives med manglende ekstern validering(110,131). Et andet eksempel kan ses i opdelingen i hovedkoderne datakvalitet, overtilpasning, problemer med udfald, problemer med prædiktorer og sample size, da de alle på varierende vis beskriver datakvalitet. Som konsekvens heraf kunne en alternativ tilgang i dette projekt have resulteret i et mindre antal, men betydeligt mere omfattende og uigennemsuelige hovedkoder. Måden hvorpå studierne opgør begrænsningerne er derfor afgørende for hvordan de opgøres i dette projekt, på trods af den bagvedliggende årsag kan være identisk.

7.1.2 Resultatdiskussion af valideringen af ERVIN

Et problem i dette projekt var den lille mængde data, som var til rådighed til at udføre valideringen, bestående af 125 patienter indsamlet i en periode på to år og ti måneder. Som beskrevet i afsnit 5.2.1 så bliver der foretaget omkring 3000 totale alloplastik operationer om året(10), hvilket kunne betyde at ERVIN kun er blevet anvendt til ca. 1,5 % af patienterne i tidsperioden, hvor data blev indsamlet. Data i dette projekt blev indsamlet samtidig med at der blev gennemført et sideløbende studie af Kastrup, Bjerregaard,

Laursen et al. (2023), hvor ERVIN blev anvendt(10). Dette kan have betydet at ERVIN blev anvendt hyppigere end, hvis førnævnte studie ikke var blevet gennemført. Ydermere kan den begrænsede anvendelse indikerer at der foreligger barriere ved implementering og anvendelsen af ERVIN, barriererne kan være både menneskelige, organisatoriske og tekniske(132).

Et problem der med stor sandsynlighed har påvirket resultaterne i dette projekt, er valideringsdatas sammenstilling. Udviklingsdatasættet på ERVIN inkluderede både data på patienter, der blev konservativt behandlet og patienter der blev kirurgisk behandlet(10). I dette projekt er hele valideringssættet på kirurgisk behandlede patienter, hvorfor valideringspopulationen er en anden end udviklingspopulationen. Et andet problem der foreligger med valideringsdatasættet, er den manglende fuldkomne opfølgningstid tid for henholdsvis reoperation ved THA og TKA og luksationer ved THA, da data indsamles fra 31 dage til to år efter operation. Patienter opereret i 2022 kan stadigt risikere reoperation ved THA og TKA og luksationer ved THA indenfor to år, uden det er registreret i valideringsdata, hvorfor resultaterne påvirkes heraf. Problemet er dog ikke gældende for korttidskomplikationer for hverken THA og TKA da de registreres indenfor 30 dage.

Et af dette projekts formål var at undersøge, hvorvidt ERVIN kan prædiktere korttidskomplikationer og reoperation for hofte- og knæpatienter efter THA og TKA og luksation for hoftepatienter efter THA, ved at udføre kalibrering og diskriminering samt beslutningskurve analyser. Ud fra dette projekts resultater kan ERVIN ikke pålideligt prædiktere nogle af de ovenstående risici udfald hos hverken THA- eller TKA-patienter, ligeledes viser beslutningskurve analyserne, at der ikke opnås en nettofordel ved at anvende ERVIN. Resultaterne skal dog tolkes med varsomhed grundet den lille sample size på valideringsdatasættet, hvilket med stor sandsynlighed påvirker resultaterne i sådan en grad, at det ikke er muligt at konkludere om ERVIN prædiktere de undersøgte risici udfald pålideligt. Resultaterne i dette projekt indikerer ligeledes en overestimering af korttidskomplikationer ved THA. Eftersom denne estimering har fungeret som beslutningsstøtte kan overestimeringen af risikoen for korttidskomplikationer potentielt have medført at færre patienter er blevet opereret end hensigtsmæssigt. Den kliniske betydning af CDSS'er kan derfor ikke undervurderes da beslutninger truffet på et fejlbart grundlag kan have store konsekvenser for patienten.

Eftersom ERVIN er udviklet til beslutningsstøtte, og ikke beslutningstagen, kan de fastsatte tærskelværdier diskuteres(9). Beslutningsstøtteværktøjer vil ofte have et interval af tærskelværdier i stedet for en enkelt værdi, hvor modellen er anvendelig(109). Et sådan tærskelværdis-spænd ville kunne anvendes i patientinddragelsen ift. til at træffe den rigtige beslutning for individet.

Om beslutningskurveanalyse er meningsfuld at gennemføre når AUC i alle tilfælde er kategoriseret som unacceptable, kan diskuteres(133). Det blev dog metodisk valg at præsenterer beslutningskurve analysen da en AUC ikke er direkte overførbar til nettofordel, hvorfor præsentation af nettofordelen blev vurderet relevant(109). Ydermere understregede fremstillingen af DCA’erne loyalitet overfor den valgte metode til valideringen(8).

7.2 Metodediskussion

7.2.1 Metodediskussion af scoping review

En ulempe ved at udføre scoping reviews er, at på grund af problemformuleringens brede karakter kan resultaterne også være brede og kræver yderligere trin fra forfatterne for at syntetisere og drage brugbare konklusioner fra dem. Metoden gør det dog muligt at lave et systematisk overblik over en emnesfære indenfor en begrænset tidsramme, hvorfor metoden alligevel blev vurderet mest egnet til at besvare problemformuleringen(93). For at sikre en grundig og transparent udarbejdelse af projektets scoping review blev Arksey og O’Malley metodiske ramme anvendt(94). På trods af den systematiske tilgang findes der visse svagheder ved det udarbejdede scoping review. Et problem kan være at det kun var studier, som beskrev begrænsninger ved kliniske beslutningsstøtte, som blev inkluderet, hvilket kan have resulteret i at potentielt relevante studier blev ekskluderet fordi de ikke tydeligt beskrev i titel/abstrakt, om deres prognostiske model blev anvendt til klinisk beslutningsstøtte. Ifølge Martin, Jenkins, Bull et al. (2020) så kan udvælgelse af studier på baggrund af titel/abstrakt være problematisk og ofte føre til et stort antal af falsk-positive, som skal fuld screenes(134). Dette var også et problem i vores scoping review, hvor det ofte var nødvendigt at fortolke, hvorvidt en model kunne anses som et CDSS eller ej, hvilket kan have ført til udvælgelses bias. For at forsøge at reducere dette blev alle studier gennemgået af begge projektmedlemmer. For at gøre udvælgelsesprocessen yderligere transparent og forbedre den metodiske kvalitet, kunne vi med fordel have opgjort de tilfælde, hvor projektmedlemmerne var uenige ift. in- eller eksklusion af et givent studie.

En anden mulig svaghed ved dette scoping review er, at de inkluderede studiers evidens ikke blev kvalitetsvurderet, hvilket betyder at alle studier vægtes lige på trods af forskellig evidensniveau/risiko for bias og kvalitet. Dette kan være problematisk når resultaterne skal analyseres, da det vil være projektmedlemmernes narrativ, som redegør for de inkluderede studier, hvilket et systematisk review bedre håndterer, men samtidig er mere ressourcekrævende(94). Derfor kan resultaterne fra scoping reviews blandt andet ikke anvendes til at drage konklusioner, men identificere videnshuller(92). Formålet med dette projekt er

netop at identificere begrænsninger ved udviklingen af prognostiske CDSS'er, og ikke drage konklusioner eller anbefalinger om, hvorvidt CDSS'er er anvendelige eller ikke anvendelige. I dette perspektiv vurderer vi derfor at et scoping review er et passende design til at besvare problemformuleringen.

7.2.2 Metodediskussion af valideringen af ERVIN

Et problem ved valideringen af ERViN var mængden af events i det inkluderede data. En tommelfingerregel siger at man bør et datasæt med minimum 100-200 hændelse og 100-200 ikke hændelse for at lave en gyldig validering(135–137), Riley, Debray, Colling et al. (2021) har dog påvist at denne regel ikke altid giver præcise estimer af en models forudsigende præstationsmål(135,136). En sample size udregning blev dog ikke fundet nødvendig da datamængden ikke kunne udvides. Problemet var tydeligt ved validering af prædikterede hofteluksationer da ingen patienter i det inkluderede data havde oplevet en hofteluksation postoperativt, hvorfor validering var umulig. Hvis det ikke er muligt at opnå en tilstrækkelig sample size, kan der desuden argumenteres for kun at evaluere kalibreringen ud fra kalibrering-i-det-store og hældningen(82,135). I dette projekt har vi på trods af den begrænsede sample size valgt at evaluere kalibreringen ud fra både en kalibreringskurve, kalibrering-i-det-store og hældningen da en kurve med en kalibrering-i-det-store tæt på nul og en hældning tæt på et, ikke nødvendigvis betyder, at kurven er tæt på den perfekte kalibreringskurve, hvis kurven er meget fleksibel(82).

8.0 Konklusion

Ved udarbejdelsen af dette specialeprojekt blev der udarbejdet følgende problemformulering:

"Hvad beskriver litteraturen som begrænsninger ved udvikling af prognostiske kliniske beslutningsstøttesystemer, og hvor pålidelig er det kliniske beslutningsstøttesystem ERVIN til at prædiktere reoperation og korttidskomplikationer for både total hofteallograft og total knæallograft patienter samt luksationer for total hofteallograft patienter?"

Den første problemstilling fokuserede på litteraturens selvopgjorde begrænsninger ved udvikling prognostiske beslutningsstøttesystemer. I alt blev der gennemgået 99 studier, der samlet identificerede 355 begrænsninger ved udviklingen. Disse begrænsninger omhandlede områder som bias, datakvalitet, generalisierbarhed, manglende ekstern validering, manglende udfald, overtilpasning, populationsdatashift, problemer med prædiktorer, rodekasse og sample size. Det skal bemærkes, at der forekom variation i hvilke, i hvilket omfang samt hvordan begrænsningerne ved udviklingen blev opgjort i studierne.

Den anden problemstilling omhandlede pålideligheden af er det kliniske beslutningsstøttesystem ERVIN til at prædiktere reoperation, korttidskomplikationer for både THA- og TKA-patienter og luksationer for THA-patienter. Resultaterne fra dette projekt indikerer, at ERVIN ikke pålideligt kan prædiktere reoperation, korttidskomplikationer for både THA- og TKA-patienter og luksationer for THA-patienter. Det er dog vigtigt at bemærke, at resultaterne baserer sig på et lille sample på valideringsdata, hvilket potentielt kan have påvirket resultatet.

Med udgangspunkt i projektets to problemstillinger og de dertilhørende resultater, kan det konkluderes, at prognostiske beslutningsstøttesystemer har en række begrænsninger, som skal håndteres og adresseres for at opnå påidelige resultater.

9.0 Perspektivering

Dette projekt har givet et indblik i tendenserne i litteraturen med hensyn til selvopgjorte begrænsninger ved udviklingen af prognostiske CDSS'er. For at sikre fremtidig anvendelighed er det nødvendigt at øge opmærksomheden omkring udviklingsbegrænsningerne ved prognostiske CDSS'er. Yderligere forskning er derfor nødvendigt for at identificere og håndtere begrænsninger ved udviklingen, og værktøjer som 'Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis' (TRIPOD) bør anvendes til at rapportere udviklingen af prognostiske CDSS'er(138). Der er flere af begrænsningerne, som er identificeret i dette projektets scoping review, der kan have influeret ERVIN'S valideringsresultater. Den første mulige forklaring på resultaterne fra ERVIN-valideringen er manglende eksterne validering, hvilket kan have resulteret i overtilpasning, der potentielt kan forklare ERVIN's manglende evne til at prædiktere reoperation, korttidskomplikationer for både THA- og TKA-patienter og luksationer for THA-patienter i ny data. En anden forklaring på resultaterne er, at ERVIN blev udviklet med for lille sample size. Det kræver en betydelig højere sample size at forudsige reoperation, korttidskomplikationer for både THA og TKA og luksationer for THA på grund af hyppighedsratioer for de givne risici udfald. En tredje forklaring kan være at ERVIN'S prædiktioner risikeres at være påvirket af selektionsbias, da modellen kun er brugt på ganske få patienter på Farsø Sygehus. Hvis ERVIN var blevet anvendt på alle THA- og TKA-patienter i perioden kunne det potentielt have påvirket resultaterne. En fjerde forklaring på ERVIN's resultater er interventionsdatashift eftersom dette problem ikke er taget højde for ved udviklingen, var dette problem blevet håndteret havde interventionen ikke påvirket kalibreringen af modellen. Problemerne ved udviklingen bør tages i betragning ved udvikling og opdatering af fremtidige CDSS'er, herunder ERVIN, for at sikre pålidelige og effektive CDSS'er, der kan bidrage til forbedret kliniske beslutningsstøtte, prognose og behandling.

10.0 Referenceliste

1. Kristian Kjær Petersen. Mechanistic Pain Profiling of Patients with Osteoarthritis : Current Knowledge and Future Directions [Internet]. Aalborg Universitetsforlag; 2021. Tilgængelig hos: https://vbn.aau.dk/ws/portalfiles/portal/450213713/1061952_thesis_mechanistic_pain_profiling_of_patients_with_osteoarthritis_kkp.pdf
2. Aboulenain S, Saber AY. Primary Osteoarthritis. I: StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing; 2022 [henvist 31. januar 2023]. Tilgængelig hos: <http://www.ncbi.nlm.nih.gov/books/NBK557808/>
3. Wirries N, Budde S. [Partial hip resurfacing-a critical analysis]. Orthopade. februar 2021;50(2):119–23.
4. Siopack JS, Jergesen HE. Total hip arthroplasty. West J Med. marts 1995;162(3):243–9.
5. Beard DJ, Davies LJ, Cook JA, MacLennan G, Price A, Kent S, m.fl. Total versus partial knee replacement in patients with medial compartment knee osteoarthritis: the TOPKAT RCT. Health Technol Assess. april 2020;24(20):1–98.
6. Price AJ, Alvand A, Troelsen A, Katz JN, Hooper G, Gray A, m.fl. Knee replacement. The Lancet. 3. november 2018;392(10158):1672–82.
7. Dekker A, Dumontier M, Kubben P, redaktører. Fundamentals of Clinical Data Science. 1st ed. 2019. Cham: Springer International Publishing : Imprint: Springer; 2019. 1 s.
8. Steyerberg EW, Vergouwe Y. Towards better clinical prediction models: seven steps for development and an ABCD for validation. Eur Heart J. 1. august 2014;35(29):1925–31.
9. Ny it-loesning og data skal hjælpe læger med at udvælge egnede patienter til operation [Internet]. [henvist 1. februar 2023]. Tilgængelig hos: <https://www.kmd.dk/indsigter/ny-it-loesning-og-data-skal-hjaelpe-laeger>
10. Kastrup N, Bjerregaard HH, Laursen M, Valentin JB, Johnsen SP, Jensen CE. An AI-based patient-specific clinical decision support system for OA patients choosing surgery or not: study protocol for a single-centre, parallel-group, non-inferiority randomised controlled trial. Trials. 12. januar 2023;24(1):24.
11. Hedensted Bjerregaard H, Falk Hvidberg M. Værdibaseret styring hofte-knæ: Principper for prædiktionsmodeller, baggrund, noter og teknik.
12. Binuya MAE, Engelhardt EG, Schats W, Schmidt MK, Steyerberg EW. Methodological guidance for the evaluation and updating of clinical prediction models: a systematic review. BMC Medical Research Methodology. 12. december 2022;22(1):316.
13. Lenert MC, Matheny ME, Walsh CG. Prognostic models will be victims of their own success, unless.... J Am Med Inform Assoc. 3. september 2019;26(12):1645–50.
14. Chen D, Shen J, Zhao W, Wang T, Han L, Hamilton JL, m.fl. Osteoarthritis: toward a comprehensive understanding of pathological mechanism. Bone Res. 17. januar 2017;5:16044.
15. Hunter DJ, Guermazi A, Roemer F, Zhang Y, Neogi T. Structural correlates of pain in joints with osteoarthritis. Osteoarthritis Cartilage. september 2013;21(9):1170–8.
16. Dieppe PA, Lohmander LS. Pathogenesis and management of pain in osteoarthritis. The Lancet.

12. marts 2005;365(9463):965–73.
17. Referred pain. J Appl Oral Sci. december 2009;17(6):i.
18. Zhang Y, Jordan JM. Epidemiology of Osteoarthritis. Clin Geriatr Med. august 2010;26(3):355–69.
19. Kawano MM, Araújo ILA, Castro MC, Matos MA. Assessment of quality of life in patients with knee osteoarthritis. Acta ortop bras. december 2015;23(6):307–10.
20. Vitaloni M, Botto-van Bemden A, Sciortino R, Carné X, Quintero M, Santos-Moreno P, m.fl. A patients' view of OA: the Global Osteoarthritis Patient Perception Survey (GOAPPS), a pilot study. BMC Musculoskelet Disord. december 2020;21(1):727.
21. Carr AJ. Beyond disability: measuring the social and personal consequences of osteoarthritis. Osteoarthritis and Cartilage. marts 1999;7(2):230–8.
22. Fautrel B, Hilliquin P, Rozenberg S, Allaert FA, Coste P, Leclerc A, m.fl. Impact of osteoarthritis: results of a nationwide survey of 10,000 patients consulting for OA. Joint Bone Spine. maj 2005;72(3):235–40.
23. Long H, Liu Q, Yin H, Wang K, Diao N, Zhang Y, m.fl. Prevalence Trends of Site-Specific Osteoarthritis From 1990 to 2019: Findings From the Global Burden of Disease Study 2019. Arthritis Rheumatol. juli 2022;74(7):1172–83.
24. Hawker GA. Osteoarthritis is a serious disease. Clin Exp Rheumatol. 2019;37 Suppl 120(5):3–6.
25. Long H, Liu Q, Yin H, Wang K, Diao N, Zhang Y, m.fl. Prevalence Trends of Site-Specific Osteoarthritis From 1990 to 2019: Findings From the Global Burden of Disease Study 2019. Arthritis Rheumatol. juli 2022;74(7):1172–83.
26. DANSKERNES SUNDHED, Den Nationale Sundhedsprofil 2021 [Internet]. Danmark: Sundhedsstyrelsen; 2021 s. 196. Report No.: 1.0. Tilgængelig hos: https://www.sst.dk/-/media/Udgivelser/2022/Sundhedsprofil/Sundhedsprofilen.ashx?sc_lang=da&hash=5C9A9A81483F6C987D5651976B72ECB2
27. 2022 Sygdomsbyrden i Danmark [Internet]. København: Sundhedsstyrelsen; 2022 s. 460. Tilgængelig hos: <https://www.sst.dk/-/media/Udgivelser/2023/Sygdomsbyrden-2023/Sygdomme-Sygdomsbyrden-2023.ashx>
28. Katz JN, Arant KR, Loeser RF. Diagnosis and treatment of hip and knee osteoarthritis: A review. JAMA. 9. februar 2021;325(6):568–78.
29. Bijlsma JW, Berenbaum F, Lafeber FP. Osteoarthritis: an update with relevance for clinical practice. The Lancet. 18. juni 2011;377(9783):2115–26.
30. Altman RD, Gold GE. Atlas of individual radiographic features in osteoarthritis, revised. Osteoarthritis and Cartilage. 1. januar 2007;15:A1–56.
31. Zhang W, Moskowitz RW, Nuki G, Abramson S, Altman RD, Arden N, m.fl. OARSI recommendations for the management of hip and knee osteoarthritis, Part I: Critical appraisal of existing treatment guidelines and systematic review of current research evidence. Osteoarthritis and Cartilage. 1. september 2007;15(9):981–1000.
32. Zhang W, Moskowitz RW, Nuki G, Abramson S, Altman RD, Arden N, m.fl. OARSI recommen-

dations for the management of hip and knee osteoarthritis, Part II: OARSI evidence-based, expert consensus guidelines. *Osteoarthritis Cartilage*. februar 2008;16(2):137–62.

33. Lim WB, Al-Dadah O. Conservative treatment of knee osteoarthritis: A review of the literature. *World J Orthop*. 18. marts 2022;13(3):212–29.
34. Bannuru RR, Osani MC, Vaysbrot EE, Arden NK, Bennell K, Bierma-Zeinstra SMA, m.fl. OARSI guidelines for the non-surgical management of knee, hip, and polyarticular osteoarthritis. *Osteoarthritis and Cartilage*. 1. november 2019;27(11):1578–89.
35. Dargel J, Oppermann J, Brüggemann GP, Eysel P. Dislocation Following Total Hip Replacement. *Dtsch Arztbl Int*. december 2014;111(51–52):884–90.
36. Latham N, Liu C ju. Strength training in older adults: The benefits for osteoarthritis. *Clin Geriatr Med*. august 2010;26(3):445–59.
37. Sinatti P, Sánchez Romero EA, Martínez-Pozas O, Villafaña JH. Effects of Patient Education on Pain and Function and Its Impact on Conservative Treatment in Elderly Patients with Pain Related to Hip and Knee Osteoarthritis: A Systematic Review. *Int J Environ Res Public Health*. 19. maj 2022;19(10):6194.
38. Artrose, knæ - Lægehåndbogen på sundhed.dk [Internet]. [henvist 31. januar 2023]. Tilgængelig hos: <https://www.sundhed.dk/sundhedsfaglig/laegehaandbogen/ortopaedi/tilstande-og-sygdomme/knae/artrose-knae/>
39. NATIONAL KLINISK RETNINGSLINJE FOR HOFTEARTROSE – IKKE-KIRURGISK BEHANDLING OG GENOPTRÆNING EFTER TOTAL HOFTEALLOPLASTIK [Internet]. Danmark: Sundhedsstyrelsen; 2016 s. 60. Tilgængelig hos: <https://www.sst.dk/da/udgivelser/2016/~/media/64C2665956B24E7CB624E7FF2EA7D892.ashx>
40. Abdulkarim A, Ellanti P, Motterlini N, Fahey T, O’Byrne JM. Cemented versus uncemented fixation in total hip replacement: a systematic review and meta-analysis of randomized controlled trials. *Orthop Rev (Pavia)*. 15. marts 2013;5(1):e8.
41. Dansk Hoftealloplastik Register - National årsrapport for 2021 [Internet]. Danmark: Dansk Hoftealloplastik Register; 2021 s. 204. Tilgængelig hos: http://danskhoftealloplastikregister.dk/wp-content/uploads/2022/07/DHR-aarsrapport-2021_Udgivet-2022_offentligjort-version.pdf
42. Dansk Knæalloplastikregister Årsrapport 2021 [Internet]. Danmark: Regionernes Kliniske Kvalitetsudviklingsprogram ; 2021 s. 204. Tilgængelig hos: https://www.sundhed.dk/content/cms/99/4699_dkr-aarsrapport-2021-budgivet2022offentligjortversion.pdf
43. Ferguson RJ, Palmer AJ, Taylor A, Porter ML, Malchau H, Glyn-Jones S. Hip replacement. *The Lancet*. 3. november 2018;392(10158):1662–71.
44. Alokozai A, Bernstein DN, Samuel LT, Kamath AF. Patient Engagement Approaches in Total Joint Arthroplasty: A Review of Two Decades. *J Patient Exp*. 19. august 2021;8:23743735211036524.
45. Fransen M, McConnell S, Hernandez-Molina G, Reichenbach S. Exercise for osteoarthritis of the hip. *Cochrane Database Syst Rev*. 22. april 2014;(4):CD007912.
46. Fransen M, McConnell S, Harmer AR, Van der Esch M, Simic M, Bennell KL. Exercise for osteoarthritis of the knee. *Cochrane Database Syst Rev*. 9. januar 2015;1:CD004376.
47. Skou ST, Simonsen ME, Odgaard A, Roos EM. Predictors of long-term effect from education and

- exercise in patients with knee and hip pain. *Dan Med J.* juli 2014;61(7):A4867.
48. Basisbehandling af artrose [Internet]. [henvist 12. februar 2023]. Tilgængelig hos: <https://www.glaid.dk/pdf/BasisBehArtrose.pdf>
49. Campbell R, Evans M, Tucker M, Quilty B, Dieppe P, Donovan JL. Why don't patients do their exercises? Understanding non-compliance with physiotherapy in patients with osteoarthritis of the knee. *J Epidemiol Community Health.* februar 2001;55(2):132–8.
50. Selten E, Vreizkold J, Greenen R, Roelien VDM, Nijhof M. Reasons for Treatment Choices in Knee and Hip Osteoarthritis: A Qualitative Study. [henvist 1. februar 2023]; Tilgængelig hos: <https://onlinelibrary.wiley.com/doi/10.1002/acr.22841>
51. Crawford DC, Miller LE, Block JE. Conservative management of symptomatic knee osteoarthritis: a flawed strategy? *Orthop Rev (Pavia).* 22. februar 2013;5(1):e2.
52. Skou ST, Roos EM, Laursen MB, Rathleff MS, Arendt-Nielsen L, Simonsen O, m.fl. A Randomized, Controlled Trial of Total Knee Replacement. *N Engl J Med.* 22. oktober 2015;373(17):1597–606.
53. Total Knee Replacement. *Ont Health Technol Assess Ser.* 1. juni 2005;5(9):1–51.
54. Marsh M, Newman S. Trends and developments in hip and knee arthroplasty technology. *J Rehabil Assist Technol Eng.* 8. februar 2021;8:2055668320952043.
55. Kahlenberg CA, Nwachukwu BU, McLawhorn AS, Cross MB, Cornell CN, Padgett DE. Patient Satisfaction After Total Knee Replacement: A Systematic Review. *HSS J.* juli 2018;14(2):192–201.
56. Mancuso CA, Salvati EA, Johanson NA, Peterson MGE, Charlson ME. Patients' expectations and satisfaction with total hip arthroplasty. *The Journal of Arthroplasty.* 1. juni 1997;12(4):387–96.
57. Healy WL, Della Valle CJ, Iorio R, Berend KR, Cushner FD, Dalury DF, m.fl. Complications of Total Knee Arthroplasty: Standardized List and Definitions of The Knee Society. *Clin Orthop Relat Res.* januar 2013;471(1):215–20.
58. Healy WL, Iorio R, Clair AJ, Pellegrini VD, Della Valle CJ, Berend KR. Complications of Total Hip Arthroplasty: Standardized List, Definitions, and Stratification Developed by The Hip Society. *Clin Orthop Relat Res.* februar 2016;474(2):357–64.
59. Hsu H, Siwiec RM. Knee Arthroplasty [Internet]. StatPearls [Internet]. StatPearls Publishing; 2022 [henvist 8. marts 2023]. Tilgængelig hos: <https://www.ncbi.nlm.nih.gov/books/NBK507914/>
60. George J, Chughtai M, Khlopas A, Klika AK, Barsoum WK, Higuera CA, m.fl. Readmission, Re-operation, and Complications: Total Hip vs Total Knee Arthroplasty. *The Journal of Arthroplasty.* 1. marts 2018;33(3):655–60.
61. Glassou EN, Pedersen AB, Hansen TB. Risk of re-admission, reoperation, and mortality within 90 days of total hip and knee arthroplasty in fast-track departments in Denmark from 2005 to 2011. *Acta Orthopaedica.* 1. september 2014;85(5):493–500.
62. Cochrane JA, Flynn T, Wills A, Walker FR, Nilsson M, Johnson SJ. Clinical Decision Support Tools for Predicting Outcomes in Patients Undergoing Total Knee Arthroplasty: A Systematic Review. *J Arthroplasty.* maj 2021;36(5):1832–1845.e1.
63. Hafkamp FJ, Gosens T, de Vries J, den Oudsten BL. Do dissatisfied patients have unrealistic expectations? A systematic review and best-evidence synthesis in knee and hip arthroplasty patients. *EFORT Open Rev.* 2. april 2020;5(4):226–40.

64. Waljee J, McGlinn EP, Sears ED, Chung KC. Patient Expectations and Patient-Reported Outcomes in Surgery: A Systematic Review. *Surgery*. maj 2014;155(5):799–808.
65. Sacristán JA, Aguarón A, Avendaño-Solá C, Garrido P, Carrión J, Gutiérrez A, m.fl. Patient involvement in clinical research: why, when, and how. *Patient Prefer Adherence*. 27. april 2016;10:631–40.
66. Bombard Y, Baker GR, Orlando E, Fancott C, Bhatia P, Casalino S, m.fl. Engaging patients to improve quality of care: a systematic review. *Implementation Science*. 26. juli 2018;13(1):98.
67. Birkeland S, Bismark M, Barry MJ, Möller S. Is greater patient involvement associated with higher satisfaction? Experimental evidence from a vignette survey. *BMJ Qual Saf*. 1. februar 2022;31(2):86–93.
68. La S, Je L. Where is the evidence? A systematic review of shared decision making and patient outcomes. *Medical decision making : an international journal of the Society for Medical Decision Making* [Internet]. januar 2015 [henvist 27. maj 2023];35(1). Tilgængelig hos: <https://pubmed.ncbi.nlm.nih.gov/25351843/>
69. L PB, M C, D S, S P. Protocol for systematic review: patient decision aids for total hip and knee arthroplasty decision-making. *Systematic reviews* [Internet]. 1. april 2021 [henvist 27. maj 2023];10(1). Tilgængelig hos: <https://pubmed.ncbi.nlm.nih.gov/33397488/>
70. L PB, M C, S P, D S. Effectiveness of patient decision aids for total hip and knee arthroplasty decision-making: a systematic review. *Osteoarthritis and cartilage* [Internet]. oktober 2021 [henvist 27. maj 2023];29(10). Tilgængelig hos: <https://pubmed.ncbi.nlm.nih.gov/34302958/>
71. Weldring T, Smith SMS. Patient-Reported Outcomes (PROs) and Patient-Reported Outcome Measures (PROMs). *Health Serv Insights*. 4. august 2013;6:61–8.
72. Joeris A, Zhu TY, Lambert S, Wood A, Jayakumar P. Real-world patient data: Can they support decision making and patient engagement? *Injury* [Internet]. 16. december 2021 [henvist 10. marts 2023]; Tilgængelig hos: <https://www.sciencedirect.com/science/article/pii/S0020138321010020>
73. Stacey D, Taljaard M, Dervin G, Tugwell P, O'Connor AM, Pomey MP, m.fl. Impact of patient decision aids on appropriate and timely access to hip or knee arthroplasty for osteoarthritis: a randomized controlled trial. *Osteoarthritis and Cartilage*. 1. januar 2016;24(1):99–107.
74. Sutton RT, Pincock D, Baumgart DC, Sadowski DC, Fedorak RN, Kroeker KI. An overview of clinical decision support systems: benefits, risks, and strategies for success. *npj Digit Med*. 6. februar 2020;3(1):1–10.
75. Vogenberg FR. Predictive and Prognostic Models: Implications for Healthcare Decision-Making in a Modern Recession. *Am Health Drug Benefits*. 2009;2(6):218–22.
76. Moons KGM, Royston P, Vergouwe Y, Grobbee DE, Altman DG. Prognosis and prognostic research: what, why, and how? *BMJ*. 23. februar 2009;338(feb23 1):b375–b375.
77. Han K, Song K, Choi BW. How to Develop, Validate, and Compare Clinical Prediction Models Involving Radiological Parameters: Study Design and Statistical Methods. *Korean J Radiol*. 14. april 2016;17(3):339–50.
78. Berner ES. Diagnostic Decision Support Systems: How to Determine the Gold Standard? *J Am Med Inform Assoc*. 2003;10(6):608–10.

79. Hendriksen JMT, Geersing GJ, Moons KGM, de Groot J a. H. Diagnostic and prognostic prediction models. *Journal of Thrombosis and Haemostasis*. 2013;11(s1):129–41.
80. Royston P, Moons KGM, Altman DG, Vergouwe Y. Prognosis and prognostic research: Developing a prognostic model. *BMJ*. 31. marts 2009;338:b604.
81. Alba AC, Agoritsas T, Walsh M, Hanna S, Iorio A, Devereaux PJ, m.fl. Discrimination and Calibration of Clinical Prediction Models: Users' Guides to the Medical Literature. *JAMA*. 10. oktober 2017;318(14):1377–84.
82. Van Calster B, McLernon DJ, van Smeden M, Wynants L, Steyerberg EW, Bossuyt P, m.fl. Calibration: the Achilles heel of predictive analytics. *BMC Medicine*. 16. december 2019;17(1):230.
83. Moons KGM, Altman DG, Vergouwe Y, Royston P. Prognosis and prognostic research: application and impact of prognostic models in clinical practice. *BMJ*. 4. juni 2009;338:b606.
84. van Smeden M, Reitsma JB, Riley RD, Collins GS, Moons KG. Clinical prediction models: diagnosis versus prognosis. *Journal of Clinical Epidemiology*. 1. april 2021;132:142–5.
85. Altman DG, Vergouwe Y, Royston P, Moons KGM. Prognosis and prognostic research: validating a prognostic model. *BMJ*. 28. maj 2009;338:b605.
86. Chang WJ, Naylor J, Natarajan P, Liu V, Adie S. Evaluating methodological quality of prognostic prediction models on patient reported outcome measurements after total hip replacement and total knee replacement surgery: a systematic review protocol. *Systematic Reviews*. 10. august 2022;11(1):165.
87. RCT Measuring the Effect of the ERVIN Software - Full Text View - ClinicalTrials.gov [Internet]. [henvist 15. februar 2023]. Tilgængelig hos: <https://clinicaltrials.gov/ct2/show/NCT04332055>
88. Grænseværdier, cutpoints og transformering - hofte-knæ prædiktionsmodel. (Klassificeret ikke tilgængeligt internt dokument).
89. ERVIN.2. 2023.
90. Munn Z, Pollock D, Khalil H, Alexander L, McInerney P, Godfrey CM, m.fl. What are scoping reviews? Providing a formal definition of scoping reviews as a type of evidence synthesis. *JBI Evidence Synthesis*. april 2022;20(4):950.
91. Maggio LA, Larsen K, Thomas A, Costello JA, Artino AR. Scoping reviews in medical education: A scoping review. *Med Educ*. juni 2021;55(6):689–700.
92. Khalil H, Peters MD, Tricco AC, Pollock D, Alexander L, McInerney P, m.fl. Conducting high quality scoping reviews-challenges and solutions. *J Clin Epidemiol*. februar 2021;130:156–60.
93. Hanneke R, Asada Y, Lieberman L, Neubauer LC. The Scoping Review Method: Mapping the Literature in “Structural Change” Public Health InterventionsChange” Public Health Interventions [Internet]. 2017. Tilgængelig hos: <https://digitalcommons.montclair.edu/cgi/viewcontent.cgi?article=1083&context=public-health-facpubs>
94. Arksey H, O’Malley L. Scoping studies: towards a methodological framework. *International Journal of Social Research Methodology*. 1. februar 2005;8(1):19–32.
95. PubMed [Internet]. PubMed. [henvist 3. marts 2023]. Tilgængelig hos: <https://pubmed.ncbi.nlm.nih.gov/>
96. Welcome - Embase [Internet]. [henvist 3. marts 2023]. Tilgængelig hos: <https://www.embase.com/landing?status=grey>

97. IEEE Xplore [Internet]. [henvist 3. marts 2023]. Tilgængelig hos: <https://ieeexplore.ieee.org/Xplore/home.jsp>
98. Bramer WM, de Jonge GB, Rethlefsen ML, Mast F, Kleijnen J. A systematic approach to searching: an efficient and complete method to develop literature searches. *J Med Libr Assoc.* oktober 2018;106(4):531–41.
99. Haraldstad AMB, Christophersen E. Chapter 5 - Literature Searches and Reference Management. I: Laake P, Benestad HB, Olsen BR, redaktører. *Research in Medical and Biological Sciences (Second Edition)* [Internet]. Amsterdam: Academic Press; 2015 [henvist 3. marts 2023]. s. 125–65. Tilgængelig hos: <https://www.sciencedirect.com/science/article/pii/B9780127999432000057>
100. Zotero | Your personal research assistant [Internet]. [henvist 13. marts 2023]. Tilgængelig hos: <https://www.zotero.org/>
101. Patino CM, Ferreira JC. Inclusion and exclusion criteria in research studies: definitions and why they matter. *J Bras Pneumol.* 2018;44(2):84.
102. Rayyan [Internet]. [henvist 13. marts 2023]. Tilgængelig hos: <https://rayyan.ai/>
103. NVivo [Internet]. Lumivero. [henvist 15. april 2023]. Tilgængelig hos: <https://lumivero.com/products/nvivo/>
104. Brinkmann S. *Det kvalitative interview*. 1. udgave. Kbh: Hans Reitzel; 2014.
105. BI og Analyse [Internet]. Region Nordjylland. [henvist 17. maj 2023]. Tilgængelig hos: <https://rn.dk/om-region-nordjylland/organisation/patientforloeb-og-økonomi/bi-og-analyse>
106. Statistical software for data science | Stata [Internet]. [henvist 30. april 2023]. Tilgængelig hos: <https://www.stata.com/>
107. Validation of clinical prediction models: what does the “calibration slope” really measure? *Journal of Clinical Epidemiology*. 1. februar 2020;118:93–9.
108. Austin PC, Steyerberg EW. Interpreting the concordance statistic of a logistic regression model: relation to the variance and odds ratio of a continuous explanatory variable. *BMC Medical Research Methodology*. 20. juni 2012;12(1):82.
109. Vickers AJ, van Calster B, Steyerberg EW. A simple, step-by-step guide to interpreting decision curve analysis. *Diagnostic and Prognostic Research*. 4. oktober 2019;3(1):18.
110. Ramspeck CL, Jager KJ, Dekker FW, Zoccali C, van Diepen M. External validation of prognostic models: what, why, how, when and where? *Clinical Kidney Journal*. 1. januar 2021;14(1):49–58.
111. Aznar-Gimeno R, Esteban LM, Labata-Lezaun G, del-Hoyo-Alonso R, Abadia-Gallego D, Paño-Pardo JR, m.fl. A Clinical Decision Web to Predict ICU Admission or Death for Patients Hospitalised with COVID-19 Using Machine Learning Algorithms. *Int J Environ Res Public Health*. 17. august 2021;18(16):8677.
112. Lefevre PLC, Dulai PS, Wang Z, Guizzetti L, Feagan BG, Pop A, m.fl. A Clinical Prediction Model to Determine Probability of Response to Certolizumab Pegol for Crohn’s Disease. *BioDrugs*. januar 2022;36(1):85–93.
113. Li D, Gao J, Hong N, Wang H, Su L, Liu C, m.fl. A Clinical Prediction Model to Predict Heparin Treatment Outcomes and Provide Dosage Recommendations: Development and Validation Study. *J Med Internet Res*. 20. maj 2021;23(5):e27118.

114. Song SH, Kim H, Kim JK, Lee H, Oh JJ, Lee SC, m.fl. A smart, practical, deep learning-based clinical decision support tool for patients in the prostate-specific antigen gray zone: model development and validation. *Journal of the American Medical Informatics Association : JAMIA*. oktober 2022;29(11):1949–57.
115. Geng A, Xiao J, Dong B, Yuan S. Analysis of prognostic factors and construction of prognostic models for triple-positive breast cancer. *Front Oncol*. 1. februar 2023;13:1071076.
116. Lee SM, Lee G, Kim TK, Le T, Hao J, Jung YM, m.fl. Development and Validation of a Prediction Model for Need for Massive Transfusion During Surgery Using Intraoperative Hemodynamic Monitoring Data. *JAMA network open*. december 2022;5(12):e2246637.
117. Patzer RE, Basu M, Larsen CP, Pastan SO, Mohan S, Patzer M, m.fl. iChoose Kidney: a Clinical Decision Aid for Kidney Transplantation vs. Dialysis Treatment. *Transplantation*. marts 2016;100(3):630–9.
118. Shen J, Casie Chetty S, Shokouhi S, Maharjan J, Chuba Y, Calvert J, m.fl. Massive external validation of a machine learning algorithm to predict pulmonary embolism in hospitalized patients. *Thromb Res*. august 2022;216:14–21.
119. Gao Y, Xiong X, Jiao X, Yu Y, Chi J, Zhang W, m.fl. PRCTC: a machine learning model for prediction of response to corticosteroid therapy in COVID-19 patients. *Aging (Albany NY)*. 12. januar 2022;14(1):54–72.
120. Horta AB, Geraldes C, Salgado C, Vieira S, Xavier M, Papoila AL. A Multivariable Prediction Model to Select Colorectal Surgical Patients for Co-Management. *Acta medica portuguesa*. februar 2021;34(2):118–27.
121. Dulai PS, Singh S, Vande Casteele N, Meserve J, Winters A, Chablaney S, m.fl. Development and Validation of Clinical Scoring Tool to Predict Outcomes of Treatment With Vedolizumab in Patients With Ulcerative Colitis. *Clinical gastroenterology and hepatology : the official clinical practice journal of the American Gastroenterological Association*. december 2020;18(13):2952-2961.e8.
122. Dhillon RK, McLernon DJ, Smith PP, Fishel S, Dowell K, Deeks JJ, m.fl. Predicting the chance of live birth for women undergoing IVF: a novel pretreatment counselling tool. *Hum Reprod*. januar 2016;31(1):84–92.
123. Van Calster B, Steyerberg EW, Wynants L, van Smeden M. There is no such thing as a validated prediction model. *BMC Medicine*. 24. februar 2023;21(1):70.
124. Montomoli J, Romeo L, Moccia S, Bernardini M, Migliorelli L, Berardini D, m.fl. Machine learning using the extreme gradient boosting (XGBoost) algorithm predicts 5-day delta of SOFA score at ICU admission in COVID-19 patients. *Journal of Intensive Medicine*. 1. oktober 2021;1(2):110–6.
125. Harvey WT, Carabelli AM, Jackson B, Gupta RK, Thomson EC, Harrison EM, m.fl. SARS-CoV-2 variants, spike mutations and immune escape. *Nat Rev Microbiol*. 2021;19(7):409–24.
126. Knoop J, van Lankveld W, Beijer L, Geerdink FJB, Heymans MW, Hoogeboom TJ, m.fl. Development and internal validation of a machine learning prediction model for low back pain non-recovery in patients with an acute episode consulting a physiotherapist in primary care. *BMC Musculoskelet Disord*. 3. september 2022;23:834.
127. Sørensen MS, Gerds TA, Hindsø K, Petersen MM. External Validation and Optimization of the

- SPRING Model for Prediction of Survival After Surgical Treatment of Bone Metastases of the Extremities. Clinical orthopaedics and related research. august 2018;476(8):1591–9.
128. Pavic M, Bogowicz M, Kraft J, Vuong D, Mayinger M, Kroeze SGC, m.fl. FDG PET versus CT radiomics to predict outcome in malignant pleural mesothelioma patients. EJNMMI Research. 13. juli 2020;10(1):81.
 129. Perkins ZB, Yet B, Sharrock A, Rickard R, Marsh W, Rasmussen TE, m.fl. Predicting the Outcome of Limb Revascularization in Patients With Lower-extremity Arterial Trauma: Development and External Validation of a Supervised Machine-learning Algorithm to Support Surgical Decisions. Bd. 272, Annals of surgery. 2020. s. 564–72.
 130. Puntis S, Whiting D, Pappa S, Lennox B. Development and external validation of an admission risk prediction model after treatment from early intervention in psychosis services. Transl Psychiatry. 11. januar 2021;11(1):1–11.
 131. Wan B, Caffo B, Vedula SS. A Unified Framework on Generalizability of Clinical Prediction Models. Frontiers in Artificial Intelligence [Internet]. 2022 [henvist 10. maj 2023];5. Tilgængelig hos: <https://www.frontiersin.org/articles/10.3389/frai.2022.872720>
 132. Trivedi MH, Kern JK, Marcee A, Granneman B, Kleiber B, Bettinger T, m.fl. Development and Implementation of Computerized Clinical Guidelines: Barriers and Solutions [Internet]. Tilgængelig hos: <https://www.thieme-connect.de/products/ejournals/pdf/10.1055/s-0038-1634216.pdf>
 133. Mandrekar JN. Receiver Operating Characteristic Curve in Diagnostic Test Assessment. Journal of Thoracic Oncology. 1. september 2010;5(9):1315–6.
 134. Martin GP, Jenkins DA, Bull L, Sisk R, Lin L, Hulme W, m.fl. Toward a framework for the design, implementation, and reporting of methodology scoping reviews. J Clin Epidemiol. november 2020;127:191–7.
 135. Riley RD, Debray TPA, Collins GS, Archer L, Ensor J, van Smeden M, m.fl. Minimum sample size for external validation of a clinical prediction model with a binary outcome. Statistics in Medicine. 2021;40(19):4230–51.
 136. Collins GS, Ogundimu EO, Altman DG. Sample size considerations for the external validation of a multivariable prognostic model: a resampling study. Stat Med. 30. januar 2016;35(2):214–26.
 137. Van Calster B, Nieboer D, Vergouwe Y, De Cock B, Pencina MJ, Steyerberg EW. A calibration hierarchy for risk models was defined: from utopia to empirical data. J Clin Epidemiol. juni 2016;74:167–76.
 138. Moons KGM, Altman DG, Reitsma JB, Ioannidis JPA, Macaskill P, Steyerberg EW, m.fl. Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD): Explanation and Elaboration. Ann Intern Med. 6. januar 2015;162(1):W1–73.
 139. Rivera SC, Kyte DG, Aiyegebusi OL, Slade AL, McMullan C, Calvert MJ. The impact of patient-reported outcome (PRO) data from clinical trials: a systematic review and critical analysis. Health Qual Life Outcomes. 16. oktober 2019;17:156.
 140. McGee RG. How to Include Patient-Reported Outcome Measures in Clinical Trials. Curr Osteoporos Rep. oktober 2020;18(5):480–5.
 141. Riley RD, Ensor J, Snell KIE, Debray TPA, Altman DG, Moons KGM, m.fl. External validation of clinical prediction models using big datasets from e-health records or IPD meta-analysis: opportunities

and challenges. BMJ : British Medical Journal (Online) [Internet]. 22. juni 2016 [henvist 17. februar 2023];353. Tilgængelig hos: <https://www.proquest.com/docview/2247471399/abstract/71A600BD38604D9DPQ/1>

142. Bjerregaard HH. Datadefinition og kildesystemer og regneregler.
143. Paulsen A, Odgaard A, Overgaard S. Translation, cross-cultural adaptation and validation of the Danish version of the Oxford hip score: Assessed against generic and disease-specific questionnaires. Bone Joint Res. september 2012;1(9):225–33.
144. Odgaard A, Paulsen A. Translation and Cross-Cultural Adaptation of the Danish Version of Oxford Knee Score (OKS) [Internet]. [henvist 23. februar 2023]. Tilgængelig hos: <https://www.spoergeskemaer.dk/wp-content/uploads/OKS-abstract-2009.pdf>
145. Parkin DW, Do Rego B, Shaw R. EQ-5D-3L and quality of life in total knee arthroplasty (TKA) patients: beyond the index scores. Journal of Patient-Reported Outcomes. 30. august 2022;6(1):91.
146. Brooks GA, Kansagra AJ, Rao SR, Weitzman JI, Linden EA, Jacobson JO. A Clinical Prediction Model to Assess Risk for Chemotherapy-Related Hospitalization in Patients Initiating Palliative Chemotherapy. JAMA Oncol. juli 2015;1(4):441–7.
147. Ningrum DNA, Kung WM, Tzeng IS, Yuan SP, Wu CC, Huang CY, m.fl. A Deep Learning Model to Predict Knee Osteoarthritis Based on Nonimage Longitudinal Medical Record. J Multidiscip Healthc. 2021;14:2477–85.
148. Gao J, Merchant AM. A Machine Learning Approach in Predicting Mortality Following Emergency General Surgery. Am Surg. september 2021;87(9):1379–85.
149. Rankin S, Han L, Scherzer R, Tenney S, Keating M, Genberg K, m.fl. A Machine Learning Model for Predicting Mortality within 90 Days of Dialysis Initiation. Kidney360. 20. juli 2022;3(9):1556–65.
150. Na X, Martin AJ, Sethi S, Kyne L, Garey KW, Flores SW, m.fl. A Multi-Center Prospective Derivation and Validation of a Clinical Prediction Tool for Severe Clostridium difficile Infection. PLoS One. 23. april 2015;10(4):e0123405.
151. Lee JS, Kim CK, Kang J, Park JM, Park TH, Lee KB, m.fl. A Novel Computerized Clinical Decision Support System for Treating Thrombolysis in Patients with Acute Ischemic Stroke. J Stroke. maj 2015;17(2):199–209.
152. Wang L, Liu A, Wang Z, Xu N, Zhou D, Qu T, m.fl. A Prognostic Model of Non-Small Cell Lung Cancer With a Radiomics Nomogram in an Eastern Chinese Population. Frontiers in Oncology [Internet]. 2022 [henvist 4. april 2023];12. Tilgængelig hos: <https://www.frontiersin.org/articles/10.3389/fonc.2022.816766>
153. Keek SA, Wesseling FWR, Woodruff HC, van Timmeren JE, Nauta IH, Hoffmann TK, m.fl. A Prospectively Validated Prognostic Model for Patients with Locally Advanced Squamous Cell Carcinoma of the Head and Neck Based on Radiomics of Computed Tomography Images. Cancers (Basel). 29. juni 2021;13(13):3271.
154. Zeng X, Shi S, Sun Y, Feng Y, Tan L, Lin R, m.fl. A time-aware attention model for prediction of acute kidney injury after pediatric cardiac surgery. Bd. 30, Journal of the American Medical Informatics Association : JAMIA. 2022. s. 94–102.
155. Hernandes Rocha TA, Elahi C, Cristina da Silva N, Sakita FM, Fuller A, Mmbaga BT, m.fl. A

traumatic brain injury prognostic model to support in-hospital triage in a low-income country: a machine learning-based approach. Bd. 132, Journal of neurosurgery. 2019. s. 1961–9.

156. Sun W, Li G, Song Y, Zhu Z, Yang Z, Chen Y, m.fl. A web based dynamic MANA Nomogram for predicting the malignant cerebral edema in patients with large hemispheric infarction. BMC Neurology. 29. september 2020;20(1):360.
157. Horton WB, Hannah EE, Morales FL, Chaney CR, Krahn KN, Chernyavskiy P, m.fl. Accuracy of a Risk Alert Threshold for ICU Hypoglycemia: Retrospective Analysis of Alert Performance and Association With Clinical Deterioration Events. Critical Care Medicine. januar 2023;51(1):136.
158. Hu D, Huang Z, Chan TM, Dong W, Lu X, Duan H. Acute Coronary Syndrome Risk Prediction Based on GRACE Risk Score. Studies in health technology and informatics. 2017;245:398–402.
159. Ming DK, Hernandez B, Sangkaew S, Vuong NL, Lam PK, Nguyet NM, m.fl. Applied machine learning for the risk-stratification and clinical decision support of hospitalised patients with dengue in Vietnam. PLOS Digital Health. 18. januar 2022;1(1):e0000005.
160. Ding Gyu, Xu Jhua, He Jhong, Nie Zyu. Clinical scoring model based on age, NIHSS, and stroke-history predicts outcome 3 months after acute ischemic stroke. Front Neurol. 5. august 2022;13:935150.
161. Zhang X, Liang J, Du Z, Xie Q, Li T, Tang F. Comparison of nomogram with random survival forest for prediction of survival in patients with spindle cell carcinoma. J Cancer Res Ther. december 2022;18(7):2006–12.
162. Huang Z, Tong Y, Kong Q. Construction of a Tool to Predict Overall Survival of Patients With Primary Spinal Tumors After Surgical Resection: A Real-World Analysis Based on the Surveillance, Epidemiology, and End Results Database. Global Spine J. 26. marts 2022;21925682221086540.
163. Elahi C, Adil SM, Sakita F, Mmbaga BT, Hernandes Rocha TA, Fuller A, m.fl. Corticosteroid Randomization after Significant Head Injury and International Mission for Prognosis and Clinical Trials in Traumatic Brain Injury Models Compared with a Machine Learning-Based Predictive Model from Tanzania. Journal of neurotrauma. januar 2022;39(1):151–8.
164. Horng S, Sontag DA, Halpern Y, Jernite Y, Shapiro NI, Nathanson LA. Creating an automated trigger for sepsis clinical decision support at emergency department triage using machine learning. PloS one. 2017;12(4):e0174708.
165. Wang C, Xu X, Shao J, Zhou K, Zhao K, He Y, m.fl. Deep Learning to Predict EGFR Mutation and PD-L1 Expression Status in Non-Small-Cell Lung Cancer on Computed Tomography Images. J Oncol. 31. december 2021;2021:5499385.
166. Mirkhani A, Roshanpoor A, Pournik O, Haddadi H, Mirzaei J, Kaveh F. Developing a Prognostic Model to Predict Mortality in Patients with Acute Bacterial Meningitis. Stud Health Technol Inform. 27. maj 2021;281:774–8.
167. Nguyen M, Corbin CK, Eulalio T, Ostberg NP, Machiraju G, Marafino BJ, m.fl. Developing machine learning models to personalize care levels among emergency room patients for hospital admission. Journal of the American Medical Informatics Association : JAMIA. oktober 2021;28(11):2423–32.
168. Zeng S, Dai L, Yang J, Gao X, Yu X, Ren Q, m.fl. Development and external validation of a nomogram predicting prognosis of upper tract urothelial carcinoma after radical nephroureterectomy. Urologic oncology. april 2019;37(4):290.e17-290.e24.

169. Thio QCBS, Karhade AV, Ogink PT, Bramer JAM, Ferrone ML, Calderón SL, m.fl. Development and Internal Validation of Machine Learning Algorithms for Preoperative Survival Prediction of Extremity Metastatic Disease. *Clin Orthop Relat Res.* februar 2020;478(2):322–33.
170. Molassiotis A, Stamatakis Z, Kontopantelis E. Development and preliminary validation of a risk prediction model for chemotherapy-related nausea and vomiting. *Support Care Cancer.* oktober 2013;21(10):2759–67.
171. Li Y, Zhao L, Yang C, Yu Z, Song J, Zhou Q, m.fl. Development and Validation of a Clinical Prediction Model for Sleep Disorders in the ICU: A Retrospective Cohort Study. *Front Neurosci.* 16. april 2021;15:644845.
172. Feng X, Ren L, Xiang Y, Xu Y. Development and validation of a nomogram for evaluating the incident risk of carotid atherosclerosis in patients with type 2 diabetes. *Front Endocrinol (Lausanne).* 16. februar 2023;14:1131430.
173. Anley DT, Akalu TY, Merid MW, Tsegaye T. Development and Validation of a Nomogram for the Prediction of Unfavorable Treatment Outcome Among Multi-Drug Resistant Tuberculosis Patients in North West Ethiopia: An Application of Prediction Modelling. *Infection and drug resistance.* 2022;15:3887–904.
174. Yu WQ, Gao HJ, Shi GD, Tang JY, Wang HF, Hu SY, m.fl. Development and validation of a nomogram to predict anastomotic leakage after esophagectomy for esophageal carcinoma. *J Thorac Dis.* juni 2021;13(6):3549–65.
175. Khor S, Lavallee D, Cizik AM, Bellabarba C, Chapman JR, Howe CR, m.fl. Development and Validation of a Prediction Model for Pain and Functional Outcomes After Lumbar Spine Surgery. *JAMA surgery.* juli 2018;153(7):634–42.
176. Suh J, Yoo S, Park J, Cho SY, Cho MC, Son H, m.fl. Development and validation of an explainable artificial intelligence-based decision-supporting tool for prostate biopsy. *BJU international.* december 2020;126(6):694–703.
177. Osmanodja B, Stegbauer J, Kantauskaitė M, Rump LC, Heinzel A, Reindl-Schwaighofer R, m.fl. Development and validation of multivariable prediction models of serological response to SARS-CoV-2 vaccination in kidney transplant recipients. *Front Immunol.* 4. oktober 2022;13:997343.
178. Zhang J, Yang Y, Fu X, Guo W. Development and validation of nomograms for prediction of overall survival and cancer-specific survival of patients of colorectal cancer. *Japanese Journal of Clinical Oncology.* 9. marts 2020;50(3):261–9.
179. Obi Y, Nguyen DV, Zhou H, Soohoo M, Zhang L, Chen Y, m.fl. Development and Validation of Prediction Scores for Early Mortality at Transition to Dialysis. *Mayo Clin Proc.* september 2018;93(9):1224–35.
180. Müller D, Haschtmann D, Fekete TF, Kleinstück F, Reitmeir R, Loibl M, m.fl. Development of a machine-learning based model for predicting multidimensional outcome after surgery for degenerative disorders of the spine. *Eur Spine J.* august 2022;31(8):2125–36.
181. Bossarte RM, Ross EL, Liu H, Turner B, Bryant C, Zainal NH, m.fl. Development of a model to predict combined antidepressant medication and psychotherapy treatment response for depression among veterans. *Bd. 326, Journal of affective disorders.* 2023. s. 111–9.
182. Rich SN, Jun I, Bian J, Boucher C, Cherabuddi K, Morris JGJ, m.fl. Development of a Prediction

- Model for Antibiotic-Resistant Urinary Tract Infections Using Integrated Electronic Health Records from Multiple Clinics in North-Central Florida. *Infectious diseases and therapy*. oktober 2022;11(5):1869–82.
183. Choi GH, Yun J, Choi J, Lee D, Shim JH, Lee HC, m.fl. Development of machine learning-based clinical decision support system for hepatocellular carcinoma. *Sci Rep*. 9. september 2020;10(1):14855.
 184. Rho MJ, Park J, Moon HW, Lee C, Nam S, Kim D, m.fl. Dr. Answer AI for prostate cancer: Clinical outcome prediction model and service. *PLoS One*. 5. august 2020;15(8):e0236553.
 185. Raj R, Wennervirta JM, Tjerkaski J, Luoto TM, Posti JP, Nelson DW, m.fl. Dynamic prediction of mortality after traumatic brain injury using a machine learning algorithm. *npj Digit Med*. 18. juli 2022;5(1):1–8.
 186. Kim JY, Boyle L, Khashram M, Campbell D. Editor's Choice - Development and Validation of a Multivariable Prediction Model of Peri-operative Mortality in Vascular Surgery: The New Zealand Vascular Surgical Risk Tool (NZRISK-VASC). *Eur J Vasc Endovasc Surg*. april 2021;61(4):657–63.
 187. Davoudi A, Sajdeya R, Ison R, Hagen J, Rashidi P, Price CC, m.fl. Fairness in the prediction of acute postoperative pain using machine learning models. *Front Digit Health*. 11. januar 2023;4:970281.
 188. Zhou SN, Jv DW, Meng XF, Zhang JJ, Liu C, Wu ZY, m.fl. Feasibility of machine learning-based modeling and prediction using multiple centers data to assess intrahepatic cholangiocarcinoma outcomes. *Ann Med*. 2022;55(1):215–23.
 189. Karsten RT, Stuiver MM, van der Molen L, Navran A, de Boer JP, Hilgers FJM, m.fl. From reactive to proactive tube feeding during chemoradiotherapy for head and neck cancer: A clinical prediction model-based approach. *Oral Oncol*. januar 2019;88:172–9.
 190. Luo Y, Wang Z, Wang C. Improvement of APACHE II score system for disease severity based on XGBoost algorithm. *BMC Medical Informatics and Decision Making*. 6. august 2021;21(1):237.
 191. Forsberg JA, Potter BK, Wagner MB, Vickers A, Dente CJ, Kirk AD, m.fl. Lessons of War: Turning Data Into Decisions. *EBioMedicine*. 17. juli 2015;2(9):1235–42.
 192. Tarumi S, Takeuchi W, Chalkidis G, Rodriguez-Loya S, Kuwata J, Flynn M, m.fl. Leveraging Artificial Intelligence to Improve Chronic Disease Care: Methods and Application to Pharmacotherapy Decision Support for Type-2 Diabetes Mellitus. *Methods Inf Med*. juni 2021;60(Suppl 1):e32–43.
 193. Poly TN, Islam MM, Muhtar MS, Yang HC, Nguyen PAA, Li YCJ. Machine Learning Approach to Reduce Alert Fatigue Using a Disease Medication-Related Clinical Decision Support System: Model Development and Validation. *JMIR medical informatics*. november 2020;8(11):e19489.
 194. Mueller M, Wagner CC, Stanislaus R, Almeida JS. Machine learning to predict extubation outcome in premature infants. *The 2013 International Joint Conference on Neural Networks (IJCNN)*. august 2013;1–6.
 195. Cole J, Hughey S, Metzger A, Geiger P, Fluke L, Booth GJ. Machine Learning to Predict Fascial Dehiscence after Exploratory Laparotomy Surgery. *J Surg Res*. december 2021;268:514–20.
 196. Liu L, Shen F, Liang H, Yang Z, Yang J, Chen J. Machine Learning-Based Modeling of Ovarian Response and the Quantitative Evaluation of Comprehensive Impact Features. *Diagnostics (Basel)*. 14. februar 2022;12(2):492.
 197. Kim HM, Byun SS, Kim JK, Jeong CW, Kwak C, Hwang EC, m.fl. Machine learning-based pre-

diction model for late recurrence after surgery in patients with renal cell carcinoma. BMC Medical Informatics and Decision Making. 13. september 2022;22(1):241.

198. Parekh N, Ali K, Davies JG, Stevenson JM, Banya W, Nyangoma S, m.fl. Medication-related harm in older adults following hospital discharge: development and validation of a prediction tool. BMJ Qual Saf. 1. februar 2020;29(2):142–53.

199. Wu C, Nazha A, Bazeley P, Zein JG. Novel machine learning to predict asthma acute exacerbation. American Journal of Respiratory and Critical Care Medicine [Internet]. 2020;201(1). Tilgængelig hos: <https://www.embase.com/search/results?subaction=viewrecord&id=L632376536&from=export>

200. Cochran AR, Raub KM, Murphy KJ, Iannitti DA, Vrochides D. Novel use of REDCap to develop an advanced platform to display predictive analytics and track compliance with Enhanced Recovery After Surgery for pancreaticoduodenectomy. Int J Med Inform. november 2018;119:54–60.

201. Sniderman J, Stark RB, Schwartz CE, Imam H, Finkelstein JA, Nousiainen MT. Patient Factors That Matter in Predicting Hip Arthroplasty Outcomes: A Machine-Learning Approach. J Arthroplasty. juni 2021;36(6):2024–32.

202. Wells BJ, Lenoir KM, Diaz-Garelli JF, Futrell W, Lockerman E, Pantalone KM, m.fl. Predicting Current Glycated Hemoglobin Values in Adults: Development of an Algorithm From the Electronic Health Record. JMIR Medical Informatics [Internet]. december 2018 [henvist 4. april 2023];6(4). Tilgængelig hos: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6231807/>

203. Zhao J, Gu S, McDermaid A. Predicting outcomes of chronic kidney disease from EMR data based on Random Forest Regression. Math Biosci. april 2019;310:24–30.

204. Tarumi S, Takeuchi W, Qi R, Ning X, Ruppert L, Ban H, m.fl. Predicting pharmacotherapeutic outcomes for type 2 diabetes: An evaluation of three approaches to leveraging electronic health record data from multiple sources. J Biomed Inform. maj 2022;129:104001.

205. Allotey J, Fernandez-Felix BM, Zamora J, Moss N, Bagary M, Kelso A, m.fl. Predicting seizures in pregnant women with epilepsy: Development and external validation of a prognostic model. PLoS medicine. maj 2019;16(5):e1002802.

206. Hüssers J, Hafer G, Hegemann J, Wiemeyer S, John SM, Hübner U. Predicting the amputation risk for patients with diabetic foot ulceration – a Bayesian decision support tool. BMC Medical Informatics and Decision Making. 24. august 2020;20(1):200.

207. Tang H, Poynton MR, Hurdle JF, Baird BC, Koford JK, Goldfarb-Rumyantzev AS. Predicting three-year kidney graft survival in recipients with systemic lupus erythematosus. ASAIO J. 2011;57(4):300–9.

208. Qian Q, Wu J, Wang J, Sun H, Yang L. Prediction Models for AKI in ICU: A Comparative Study. Int J Gen Med. 25. februar 2021;14:623–32.

209. Hong S, Lee S, Lee J, Cha WC, Kim K. Prediction of Cardiac Arrest in the Emergency Department Based on Machine Learning and Sequential Characteristics: Model Development and Retrospective Clinical Validation Study. JMIR Med Inform. 4. august 2020;8(8):e15932.

210. Etu EE, Monplaisir L, Arslanturk S, Masoud S, Aguwa C, Markevych I, m.fl. Prediction of Length of Stay in the Emergency Department for COVID-19 Patients: A Machine Learning Approach. Bd. 10, IEEE Access. 2022. s. 42243–51.

211. Oosterhoff JHF, Karhade AV, Oberai T, Franco-Garcia E, Doornberg JN, Schwab JH. Prediction

- of Postoperative Delirium in Geriatric Hip Fracture Patients: A Clinical Prediction Model Using Machine Learning Algorithms. *Geriatr Orthop Surg Rehabil.* 13. december 2021;12:21514593211062276.
212. Likosky DS, Paugh TA, Harrington SD, Wu X, Rogers MAM, Dickinson TA, m.fl. Prediction of Transfusions after Isolated Coronary Artery Bypass Grafting Surgery. *Ann Thorac Surg.* marts 2017;103(3):764–72.
213. Eleuteri A, Taktak AFG, Coupland SE, Heimann H, Kalirai H, Damato B. Prognostication of metastatic death in uveal melanoma patients: A Markov multi-state model. Bd. 102, Computers in biology and medicine. 2018. s. 151–6.
214. Peng SY, Chuang YC, Kang TW, Tseng KH. Random forest can predict 30-day mortality of spontaneous intracerebral hemorrhage with remarkable discrimination. Bd. 17, European journal of neurology. 2010. s. 945–50.
215. Jiang W, Siddiqui S, Barnes S, Barouch LA, Korley F, Martinez DA, m.fl. Readmission Risk Trajectories for Patients With Heart Failure Using a Dynamic Prediction Approach: Retrospective Study. Bd. 7, JMIR medical informatics. 2019. s. e14756.
216. Kim K, Yang H, Yi J, Son HE, Ryu JY, Kim YC, m.fl. Real-Time Clinical Decision Support Based on Recurrent Neural Networks for In-Hospital Acute Kidney Injury: External Validation and Model Interpretation. Bd. 23, Journal of medical Internet research. 2021. s. e24120.
217. van den Boorn HG, Dijksterhuis WPM, van der Geest LGM, de Vos-Geelen J, Besselink MG, Wilmink JW, m.fl. SOURCE-PANC: A Prediction Model for Patients With Metastatic Pancreatic Ductal Adenocarcinoma Based on Nationwide Population-Based Data. Bd. 19, Journal of the National Comprehensive Cancer Network : JNCCN. 2021. s. 1045–53.
218. Kent M, Penson DF, Albertsen PC, Goodman M, Hamilton AS, Stanford JL, m.fl. Successful external validation of a model to predict other cause mortality in localized prostate cancer [Internet]. Bd. 14, BMC Medicine. 2016. Tilgængelig hos: <https://www.embase.com/search/results?subaction=viewrecord&id=L608161241&from=export>
219. Hargovan S, Gunnarsson R, Carter A, De Costa A, Brooks J, Groch T, m.fl. The 4-Hour Cairns Sepsis Model: A novel approach to predicting sepsis mortality at intensive care unit admission. Bd. 34, Australian critical care : official journal of the Confederation of Australian Critical Care Nurses. 2021. s. 552–60.
220. de Wit HAJM, Winkens B, Mestres Gonzalvo C, Hurkens KPGM, Mulder WJ, Janknegt R, m.fl. The development of an automated ward independent delirium risk prediction model. Bd. 38, International journal of clinical pharmacy. 2016. s. 915–23.
221. Du Y, Rafferty AR, McAuliffe FM, Mehegan J, Mooney C. Towards an explainable clinical decision support system for large-for-gestational-age births. Bd. 18, PloS one. 2023. s. e0281821.
222. Atsma F, Molenkamp O, Bouma H, Bolder SB, Stef Groenewoud A, Westert GP. Uniform criteria for total hip replacement surgery in patients with hip osteoarthritis: A decision tool to guide treatment decisions [Internet]. Bd. 33, International Journal for Quality in Health Care. 2021. Tilgængelig hos: <https://www.embase.com/search/results?subaction=viewrecord&id=L2012858062&from=export>
223. de Vries-Bouwstra J, Le Cessie S, Allaart C, Breedveld F, Huizinga T. Using predicted disease outcome to provide differentiated treatment of early rheumatoid arthritis. Bd. 33, The Journal of rheumatology. 2006. s. 1747–53.

224. Xu H, Pang J, Yang X, Li M, Zhao D. Using predictive process monitoring to assist thrombolytic therapy decision-making for ischemic stroke patients. Bd. 20, BMC medical informatics and decision making. 2020. s. 120.
225. LaFleur J, Nelson RE, Yao Y, Adler RA, Nebeker JR. Validated risk rule using computerized data to identify males at high risk for fracture. Bd. 23, Osteoporosis international : a journal established as result of cooperation between the European Foundation for Osteoporosis and the National Osteoporosis Foundation of the USA. 2012. s. 1017–27.
226. Haaskjold YL, Lura NG, Bjørneklett R, Bostad L, Bostad LS, Knoop T. Validation of two IgA nephropathy risk-prediction tools using a cohort with a long follow-up. Nephrology, dialysis, transplantation : official publication of the European Dialysis and Transplant Association - European Renal Association. 2022. s. gfac225.

11.0 Bilag

Bilag 1: Patient Reported Outcomes (PRO)

"Patient Reported Outcomes" (PRO) indsamles med "Patient Reported Outcomes Measures" (PROM) og giver indsigt, der kan forbedre behandlingen af patienter og befolkninger, mens de også forbedrer beslutningsprocessen for kliniker og patient(71,72). PRO rapporteres direkte af patienten, uden at en kliniker fortolker patientens svar, og omhandler patientens sundhed, livskvalitet eller funktionsniveau i relation til sundhedsvæsnet eller en behandling(72,139). Udfaldet kan f.eks. være smerteintensitet eller ændringer af funktion. Ved at anvende PRO, kan der opnås en mere omfattende forståelse for fordelene ved en behandling, end ved blot at anvende udfald som f.eks. overlevelse(71). Indsamlet PRO-data kan have sammenhæng med patienters sundhedsresultater på længere sigt(72). Der kan være flere udfordringer ved at indsamle og anvende PRO-data. En udfordring kan være at de ofte indeholder mange spørgsmål og det tager derfor lang tid at besvare for patienterne, mængden af spørgsmål øger også tiden det tager at indsamle, analysere og præsentere data(140).

Bilag 2: Valideringstyper

Validering af kliniske prognostiske beslutningsstøttesystemer

Validering af kliniske prognostiske beslutningsstøttesystemer omfatter flere metoder til at evaluere reproducertbarheden og generaliserbarheden af en model, som er designet til at forudsige et eller flere udvald baseret på patientdata(85). Til at evaluere prognostiske beslutningsstøttesystemers ydeevne anvendes som minimum kalibreringen og diskrimineringen(8,85,110). Der findes to strategier til at validere prædiktionsmodeller intern validering og ekstern validering(8).

Intern Validering:

Intern validering anvendes til at give information om reproducertbarheden af modellen og om modellen er overtilpasset til data, men intern validering kan ikke give informationer om generaliserbarheden af en model, da der ikke anvendes nye patienter fra en anden lokalitet(12,85). Intern validering foregår oftest ved at opdele data i træningsdata og testdata, som kan føre til overtilpasning til træningsdata(8,85,110). Risikoen for overtilpasning kan modvirkes ved at anvende teknikker som f.eks. bootstrapping og krydsvalidering til intern validering(12).

Ekstern validering:

Ekstern validering anvendes til at teste prædiktionsmodellens ydeevne og præcision på data, der ikke er anvendt til at udvikle modellen. Dette kan inkludere data fra en anden patientgruppe, en anden lokalitet (geografisk validering) eller fra en anden tidsperiode (temporal validering)(8,12,85,110). Formålet med ekstern validering, er at teste pålideligheden af modellen på nye uafhængige data(8,85). En udfordring ved ekstern validering er at det er udfordrende at opnå tilsvarende data, som modellen er trænet og bygget efter, i en anden lokalitet(141).

Bilag 3: Risici udfald

Funktion:

Til at bestemme patientens funktionsniveau anvendes Oxford Hip Score (OHS) til hoftepatienter og Oxford Knee Score OKS til knæpatienter(142). Begge spørgeskemaer er designet og udviklet til at vurdere patienternes funktion og smerte ud fra patiens perspektiv(143,144).

Livskvalitet:

Til at vurdere patienternes livskvalitet anvendes spørgeskemaet EQ-5D-3L(142). EQ-5D-3L er et mål for selv rapporteret helbredsrelateret livskvalitet. Det er designet til at kunne sammenlignes på tværs af forskellige patientdemografi, sundhedstilstande, sygdomsområder og behandlinger. Spørgeskemaet indsamler data fra fem dimension af helbredsrelateret livskvalitet; 1) mobilitet, 2) egenomsorg, 3) normale aktiviteter, 4) smerte og ubehag og 5) angst og depression. Patienterne kan vælge mellem tre niveauer, til at beskrive deres problemer i hver dimension(145).

Reoperation:

Re-opereret indenfor to år efter operationen uanset om operationen er akut eller elektiv, dog ikke de første 30 dage, der anses som korttidskomplikationer. Reoperationerne defineres ud fra procedurekoder som for knæ er: KNFC2, KNFC3, KNFC4 og KNFC5. For hofte er procedurekoderne: KNGC2, KNGC3 og KNGC4(142).

Korttidskomplikationer:

Korttidskomplikationer er akut indlæggelse med varighed på over 24 timer indenfor de første 30 dage efter udskrivelsen eller dyb venetrombose (DVT) behandlet ambulant eller lungeemboli behandlet ambulant. Reoperation indenfor de første 30 dage, både akutte og elektive, anses ligeledes som korttidskomplikationer. Det samme er gældende for hofte luksationer indenfor de første 30 dage uanset om de behandles i skadestuen eller ved indlæggelse(142).

Luksationer:

Luksationer indenfor to år efter operationen uanset om patienten behandles ambulant eller ved indlæggelse, dog ikke de første 30 dage, der anses som korttidskomplikationer. Luksationer kan kun ske ved THA og defineres ud fra procedurekoden KNFH20(142)

Bilag 4: Søgestreng

Embase

Klinisk beslutningsstøttesystem	Prognostisk prædiktion
"Clinical decision support system*"	"Prediction model*"
"Clinical decision support system"	"Prediction model" (Emtree)
"decision support systems, clinical"	"Prognostic model*"
CDSS	"Prognostic model" (Emtree)
"Clinical Decision Support*"	"Prediction rule*"
"Clinical Decision Support" (Emtree)	
"Medical decision support system*"	
"Decision-making tool for clinicians"	
"Clinical decision-making software"	
"Decision support technology"	
"Decision support technologies"	
"Clinical decision support software"	

IEEE

Klinisk beslutningsstøttesystem	Prognostisk prædiktion
"Clinical decision support system*"	"Prediction model*"
"decision support systems, clinical"	"Prognostic model*"
CDSS	"Prediction rule*"
"Clinical Decision Support*"	
"Medical decision support system*"	
"Decision-making tool for clinicians"	
"Clinical decision-making software"	
"Decision support technology"	
"Clinical decision support software"	

Bilag 5: Inkluderede studier

Navn	År	Forfatter(e)	Lande(e)	Emnefelt	Sample size	Dataindsamling	Lokalitet	Validering
A Clinical Decision Web to Predict ICU Admission or Death for Patients Hospitalised with COVID-19 Using Machine Learning Algorithms(111)	2021	Aznar-Gimeno, Rocío; Esteban, Luis M.; Labata-Lezaun, Gorka et al.	Spanien	Covid-19	3623	Retrospektivt	Multi	Eksternt valideret
A Clinical Prediction Model to Assess Risk for Chemotherapy-Related Hospitalization in Patients Initiating Palliative Chemotherapy(146)	2015	Brooks, Gabriel A.; Kansagra, Ankit J.; Rao, Sowmya R. et al.	USA	Kræft	1579	Retrospektivt	Single	Internt valideret
A Clinical Prediction Model to Determine Probability of Response to Certolizumab Pegol for Crohn's Disease(112)	2022	Lefevre, Pavine L.C.; Dulai, Parambir S.; Wang, Zhongya et al.	USA	Crohns sygdom	2157	Retrospektivt	Multi	Eksternt valideret
A Clinical Prediction Model to Predict Heparin Treatment Outcomes and Provide Dosage Recommendations: Development and Validation Study(113)	2021	Li, Dongkai; Gao, Jianwei; Hong, Na et al.	Kina	Udfald efter blodfortyndende	5156	Retrospektivt	Multi	Eksternt valideret
A Deep Learning Model to Predict Knee Osteoarthritis Based on Nonimage Longitudinal Medical Record(147)	2021	Ningrum, Dina Nur Anggraini; Kung, Woon-Man; Tzeng, I.-Shiang et al.	Taiwan	Knæartrose	132594	Retrospektivt	Multi	Internt valideret
A Machine Learning Approach in Predicting Mortality Following Emergency General Surgery(148)	2021	Gao, Jeff; Merchant, Aziz M.	USA	Akut general kirurgi	177872	Retrospektivt	Multi	Internt valideret
A Machine Learning Model for Predicting Mortality within 90 Days of Dialysis Initiation(149)	2022	Rankin, Summer; Han, Lucy; Scherzer, Rebecca; Tenney, Susan et al.	USA	Nyresygdom	188	Retrospektivt	Multi	Internt valideret
A Multi-Center Prospective Derivation and Validation of a	2015	Na, Xi; Martin, Alan J.; Sethi, Saurabh et al.	USA; Irland	Clostridioides difficile-infektion	638	Prospektivt	Multi	Internt valideret

Clinical Prediction Tool for Severe Clostridium difficile Infection(150)								
A Multivariable Prediction Model to Select Colorectal Surgical Patients for Co-Management(120)	2021	Horta, Alexandra Bayão; Geraldes, Carlos; Salgado, Cátia et al.	Portugal	Tarmkirugi	344	Retrospektivt	Single	Temporalt valideret
A Novel Computerized Clinical Decision Support System for Treating Thrombolysis in Patients with Acute Ischemic Stroke(151)	2015	Lee, Ji Sung; Kim, Chi Kyung; Kang, Jihoon et al.	Sydkorea	Apopleksi	958	Retrospektivt	Single	Internt valideret
A Prognostic Model of Non-Small Cell Lung Cancer With a Radiomics Nomogram in an Eastern Chinese Population(152)	2022	Wang, Lijie; Liu, Ailing; Wang, Zhiheng et al.	Kina	Kræft	1480	Retrospektivt	Single	Internt valideret
A Prospectively Validated Prognostic Model for Patients with Locally Advanced Squamous Cell Carcinoma of the Head and Neck Based on Radiomics of Computed Tomography Images(153)	2021	Keek, Simon A.; Wesseling, Frederik W. R.; Woodruff, Henry C et al.	Holland	Kræft	809	Retro- og prospektivt	Multi	Eksternt valideret
A smart, practical, deep learning-based clinical decision support tool for patients in the prostate-specific antigen gray zone: model development and validation(114)	2022	Song, Sang Hun; Kim, Hwanik; Kim, Jung Kwon et al.	Sydkorea	Prostata	18824	Retrospektivt	Multi	Eksternt valideret
A time-aware attention model for prediction of acute kidney injury after pediatric cardiac surgery(154)	2022	Zeng, Xian; Shi, Shanshan; Sun, Yuhan et al.	Kina	Nyresygdom	3386	Retrospektivt	Single	Internt valideret
A traumatic brain injury prognostic model to support in-hospital triage in a low-income country: a machine learning-based approach(155)	2019	Hernandes Rocha, Thiago Augusto; Elahi, Cyrus; Cristina da Silva, Núbia et al.	Tanzania	Hjerneskade	3138	Retrospektivt	Single	Internt valideret

A web based dynamic MANA Nomogram for predicting the malignant cerebral edema in patients with large hemispheric infarction(156)	2020	Sun, Wenzhe; Li, Guo; Song, Yang et al.	Kina	Apopleksi	142	Retrospektivt	Single	Internt valideret
Accuracy of a Risk Alert Threshold for ICU Hypoglycemia: Retrospective Analysis of Alert Performance and Association With Clinical Deterioration Events(157)	2023	Horton, William B.; Hannah, Elaine E.; Morales, Frances L. et al.	USA	Intensiv patienter	342	Retrospektivt	Single	Ingen beskrivelse af validering
Acute Coronary Syndrome Risk Prediction Based on GRACE Risk Score(158)	2017	Hu, Danqing; Huang, Zhengxing; Chan, Tak-Ming et al.	Kina	Åreforkalkning	3643	Retrospektivt	Single	Internt valideret
Analysis of prognostic factors and construction of prognostic models for triple-positive breast cancer(115)	2023	Geng, Anqi; Xiao, Jingjing; Dong, Bingyao; Yuan, Shifang	Kina	Kræft	5960	Retrospektivt	Multi	Eksternt valideret
Applied machine learning for the risk-stratification and clinical decision support of hospitalised patients with dengue in Vietnam(159)	2022	Ming, Damien K.; Hernandez, Bernard; Sangkaew, Sorawat et al.	Vietnam	Dengueshock syndrom	4131	Retrospektivt	Multi	Internt valideret
Clinical scoring model based on age, NIHSS, and stroke-history predicts outcome 3 months after acute ischemic stroke(160)	2022	Ding, Gang-yu; Xu, Jian-hua; He, Ji-hong et al.	Kina	Apopleksi	132	Prospektivt	Single	Internt valideret
Comparison of nomogram with random survival forest for prediction of survival in patients with spindle cell carcinoma(161)	2022	Zhang, Xiaoshuai; Liang, Jing; Du, Zhaohui et al.	Kina	Kræft	250	Retrospektivt	Multi	Internt valideret
Construction of a Tool to Predict Overall Survival of Patients With Primary Spinal Tumors After Surgical Resection: A Real-World Analysis Based	2022	Huang, Zhangheng; Tong, Yuexin; Kong, Qingquan	Kina	Kræft	695	Retrospektivt	Multi	Eksternt valideret

on the Surveillance, Epidemiology, and End Results Database(162)								
Corticosteroid Randomization after Significant Head Injury and International Mission for Prognosis and Clinical Trials in Traumatic Brain Injury Models Compared with a Machine Learning-Based Predictive Model from Tanzania(163)	2022	Elahi, Cyrus; Adil, Syed M.; Sakita, Francis et al.	Tanzania	Hovedskader	2972	Prospektivt	Single	Internt valideret
Creating an automated trigger for sepsis clinical decision support at emergency department triage using machine learning(164)	2017	Horng, Steven; Sontag, David A.; Halpern, Yoni et al.	USA	Sepsis	230936	Retrospektivt	Multi	Internt valideret
Deep Learning to Predict EGFR Mutation and PD-L1 Expression Status in Non-Small-Cell Lung Cancer on Computed Tomography Images(165)	2021	Wang, Chengdi; Xu, Xiuyuan; Shao, Jun et al.	Kina	Kræft	1262	Retrospektivt	Single	Internt valideret
Developing a Prognostic Model to Predict Mortality in Patients with Acute Bacterial Meningitis(166)	2021	Mirkhani, Atiehsadat; Roshanpoor, Arash; Pournik, Omid et al.	Iran	Hjerneskade (Meningitis)	3923	Retrospektivt	Multi	Internt valideret
Developing machine learning models to personalize care levels among emergency room patients for hospital admission(167)	2021	Nguyen, Minh; Corbin, Conor K.; Eu-lalio, Tiffany et al.	USA	Intensiv patienter	41654	Retrospektivt	Multi	Temporalt valideret
Development and external validation of a nomogram predicting prognosis of upper tract urothelial carcinoma after radical nephroureterectomy(168)	2019	Zeng, Shuxiong; Dai, Lihe; Yang, Jun et al.	Kina	Kræft	672	Retrospektivt	Multi	Eksternt valideret
Development and external validation of an admission risk	2021	Puntis, Stephen; Whiting, Daniel; Pappa, Sofia et al.	England	Psykose	2224	Retrospektivt	Multi	Eksternt valideret

prediction model after treatment from early intervention in psychosis services(130)								
Development and internal validation of a machine learning prediction model for low back pain non-recovery in patients with an acute episode consulting a physiotherapist in primary care(126)	2022	Knoop, J.; van Lankveld, W.; Beijer, L et al.	Holland	Lændesmerter	247	Prospektivt	Multi	Internt valideret
Development and Internal Validation of Machine Learning Algorithms for Preoperative Survival Prediction of Extremity Metastatic Disease(169)	2020	Thio, Quirina C. B. S.; Karhade, Aditya V.; Ogink, Paul T. et al.	USA	Kræft	1090	Retrospektivt	Multi	Internt valideret
Development and preliminary validation of a risk prediction model for chemotherapy-related nausea and vomiting(170)	2013	Molassiotis, A.; Stamatakis, Z.; Kontopantelis, E.	Storbritannien	Bivirkninger ved kemoterapi	336	Prospektivt	Multi	Internt valideret
Development and Validation of a Clinical Prediction Model for Sleep Disorders in the ICU: A Retrospective Cohort Study(171)	2021	Li, Yun; Zhao, Lina; Yang, Chenyi et al.	Kina	Søvnforstyrrelser	2082	Retrospektivt	Single	Internt valideret
Development and validation of a nomogram for evaluating the incident risk of carotid atherosclerosis in patients with type 2 diabetes(172)	2023	Feng, Xiao; Ren, Liying; Xiang, Yuping et al.	Kina	Arteriosklerose	1049	Retrospektivt	Single	Internt valideret
Development and Validation of a Nomogram for the Prediction of Unfavorable Treatment Outcome Among Multi-Drug Resistant Tuberculosis Patients in North West Ethiopia: An Application of Prediction Modelling(173)	2022	Anley, Denekew Tenaw; Akalu, Temesgen Yihunie; Merid, Mehari Woldemariam et al.	Etiopien	Tuberkulose	578	Retrospektivt	Multi	Internt valideret
Development and validation of a nomogram to predict anastomotic leakage after	2021	Yu, Wen-Quan; Gao, Hui-Jiang; Shi, Guo-Dong et al.	Kina	Anastomose-lækage	604	Retrospektivt	Single	Internt valideret

esophagectomy for esophageal carcinoma(174)								
Development and Validation of a Prediction Model for Need for Massive Transfusion During Surgery Using Intraoperative Hemodynamic Monitoring Data(116)	2022	Lee, Seung Mi; Lee, Garam; Kim, Tae Kyong et al.	Sydkorea	Blodtransfusion	18480	Retrospektivt	Multi	Eksternt valideret
Development and Validation of a Prediction Model for Pain and Functional Outcomes After Lumbar Spine Surgery(175)	2018	Khor, Sara; Lavallee, Danielle; Cizik, Amy M. et al.	USA	Ortopædkirugi	1965	Retrospektivt	Multi	Internt valideret
Development and validation of an explainable artificial intelligence-based decision-supporting tool for prostate biopsy(176)	2020	Suh, Jungyo; Yoo, Sangjun; Park, Juhyun et al.	Sydkorea	Prostata	3791	Retrospektivt	Single	Internt valideret
Development and Validation of Clinical Scoring Tool to Predict Outcomes of Treatment With Vedolizumab in Patients With Ulcerative Colitis(121)	2020	Dulai, Parambir S.; Singh, Siddharth; Vande Casteele, Niels et al.	USA; Canada	Collitis	819	Retrospektivt	Multi	Temporalt valideret
Development and validation of multivariable prediction models of serological response to SARS-CoV-2 vaccination in kidney transplant recipients(177)	2022	Osmanodja, Bilgin; Stegbauer, Johannes; Kantauskaitė, Marta et al.	Tyskland	Vaccinerespons	590	Retrospektivt	Multi	Eksternt valideret
Development and validation of nomograms for prediction of overall survival and cancer-specific survival of patients of colorectal cancer(178)	2020	Zhang, Jieyun; Yang, Yue; Fu, Xiaojian et al.	Kina	Kræft	239594	Retrospektivt	Multi	Eksternt valideret
Development and Validation of Prediction Scores for Early Mortality at Transition to Dialysis(179)	2018	Obi, Yoshitsugu; Nguyen, Danh V.; Zhou, Hui; Soohoo, Melissa et al.	USA	Tidlig død	40162	Retrospektivt	Multi	Eksternt valideret
Development of a machine-learning based model for pre-	2022	Müller, D.; Haschtmann, D.; Fekete, T. F. et al.	Schweiz	Udfald efter operation i ryggen	8374	Retrospektivt	Single	Internt valideret

dicting multidimensional outcome after surgery for degenerative disorders of the spine(180)								
Development of a model to predict combined antidepressant medication and psychotherapy treatment response for depression among veterans(181)	2023	Bossarte, Robert M.; Ross, Eric L.; Liu, Howard et al.	USA	Depression	658	Retrospektivt	Multi	Internt valideret
Development of a Prediction Model for Antibiotic-Resistant Urinary Tract Infections Using Integrated Electronic Health Records from Multiple Clinics in North-Central Florida(182)	2022	Rich, Shannan N.; Jun, Inyoung; Bian, Jiang et al.	USA	Urinvejsinfektion	9990	Retrospektivt	Multi	Eksternt valideret
Development of machine learning-based clinical decision support system for hepatocellular carcinoma(183)	2020	Choi, Gwang Hyeon; Yun, Jihye; Choi, Jonggi; Lee, Danbi et al.	Sydkorea	Kræft	1021	Retrospektivt	Single	Internt valideret
Dr. Answer AI for prostate cancer: Clinical outcome prediction model and service(184)	2020	Rho, Mi Jung; Park, Jihwan; Moon, Hyong Woo et al.	Sydkorea	Kræft	7128	Retrospektivt	Multi	Internt valideret
Dynamic prediction of mortality after traumatic brain injury using a machine learning algorithm(185)	2022	Raj, Rahul; Wennervirta, Jenni M.; Tjerkaski, Jonathan et al.	Finland; USA; Sverige	Tidlig død efter hjerneskade	1150	Retrospektivt	Multi	Eksternt valideret
Editor's Choice - Development and Validation of a Multivariable Prediction Model of Peri-operative Mortality in Vascular Surgery: The New Zealand Vascular Surgical Risk Tool (NZRISK-VASC)(186)	2021	Kim, Jee Young; Boyle, Luke; Khashram, Manar et al.	New Zealand	Peri-operativ dødlighed	21597	Retrospektivt	Multi	Internt valideret
External Validation and Optimization of the SPRING Model for Prediction of Survival After Surgical Treatment of Bone Metastases of the Extremities(127)	2018	Sørensen, Michala Skovlund; Gerds, Thomas Alexander; Hindsø, Klaus et al.	Danmark	Kræft	434	Retro- og prospektivt	Multi	Temporalt valideret

Fairness in the prediction of acute postoperative pain using machine learning models(187)	2023	Davoudi, Anis; Sajdeya, Ruba; Ison, Ron et al.	USA	Postoperativ smerte	14263	Retrospektivt	Single	Internt valideret
FDG PET versus CT radiomics to predict outcome in malignant pleural mesothelioma patients(128)	2020	Pavic, M.; Bogowicz, M.; Kraft, J. et al.	Schweiz	Kræft	72	Retrospektivt	Single	Temporalt valideret
Feasibility of machine learning-based modeling and prediction using multiple centers data to assess intrahepatic cholangiocarcinoma outcomes(188)	2022	Zhou, Shuang-Nan; Jv, Da-Wei; Meng, Xiang-Fei; Zhang, Jing-Jing et al.	Kina	Kræft	4902	Retrospektivt	Multi	Internt valideret
From reactive to proactive tube feeding during chemoradiotherapy for head and neck cancer: A clinical prediction model-based approach(189)	2019	Karsten, R. T.; Stuiver, M. M.; van der Molen, L. et al.	Holland	Sondeernæring	336	Retrospektivt	Single	Internt valideret
iChoose Kidney: a Clinical Decision Aid for Kidney Transplantation vs. Dialysis Treatment(117)	2016	Patzer, Rachel E.; Basu, Mohua; Larsen, Christian P. et al.	USA	Nyresygdom	721339	Retrospektivt	Multi	Eksternt valideret
Improvement of APACHE II score system for disease severity based on XGBoost algorithm(190)	2021	Luo, Yan; Wang, Zhiyu; Wang, Cong	Kina	Intensiv patienter	24777	Retrospektivt	Multi	Eksternt valideret
Lessons of War: Turning Data Into Decisions(191)	2015	Forsberg, Jonathan A.; Potter, Benjamin K.; Wagner, Matthew B et al.	USA	Krigsskader	73	Retrospektivt	Single	Internt valideret
Leveraging Artificial Intelligence to Improve Chronic Disease Care: Methods and Application to Pharmacotherapy Decision Support for Type-2 Diabetes Mellitus(192)	2021	Tarumi, Shinji; Takeuchi, Wataru; Chalkidis, George et al.	Japan; USA	Diabetes	27904	Retrospektivt	Multi	Temporalt valideret
Machine Learning Approach to Reduce Alert Fatigue Using a Disease Medication-Related	2020	Poly, Tahmina Nasrin; Islam, Md Mohaimenul; Muhtar, Muhammad Solihuddin et al.	Taiwan	Alarmtræthed	6453	Retrospektivt	Single	Internt valideret

Clinical Decision Support System: Model Development and Validation(193)								
Machine learning to predict extubation outcome in premature infants(194)	2013	Mueller, M.; Wagner, C. C.; Stanislaus, R. et al.	USA	Tubering	486	Retrospektivt	Single	Internt valideret
Machine Learning to Predict Fascial Dehiscence after Exploratory Laparotomy Surgery(195)	2021	Cole, Jacob; Hughey, Scott; Metzger, Alexander et al.	USA	Facial-løsning	93024	Retrospektivt	Multi	Temporalt valideret
Machine learning using the extreme gradient boosting (XGBoost) algorithm predicts 5-day delta of SOFA score at ICU admission in COVID-19 patients(124)	2021	Montomoli, Jonathan; Romeo, Luca; Moccia, Sara et al.	Italien; Schweiz; Frankrig, Holland	Covid-19	675	Ingen beskrivelse	Multi	Internt valideret
Machine Learning-Based Modeling of Ovarian Response and the Quantitative Evaluation of Comprehensive Impact Features(196)	2022	Liu, Liu; Shen, Fujin; Liang, Hua et al.	Kina	Fertilitet	1365	Prospektivt	Single	Temporalt valideret
Machine learning-based prediction model for late recurrence after surgery in patients with renal cell carcinoma(197)	2022	Kim, Hyung Min; Byun, Seok-Soo; Kim, Jung Kwon et al.	Sydkorea	Kræft	2956	Retrospektivt	Multi	Internt valideret
Massive external validation of a machine learning algorithm to predict pulmonary embolism in hospitalized patients(118)	2022	Shen, Jieru; Casie Chetty, Satish; Shokouhi, Sepideh et al.	USA	Lungeemboli	~ 330000	Retrospektivt	Multi	Eksternt valideret
Medication-related harm in older adults following hospital discharge: development and validation of a prediction tool(198)	2020	Parekh, Nikesh; Ali, Khalid; Davies, John Graham et al.	England	Medicinskade	818	Prospektivt	Multi	Internt valideret
Novel machine learning to predict asthma acute exacerbation(199)	2020	Wu, C.; Nazha, A.; Bazeley, P. et al.	USA	Astma	60302	Retrospektivt	Multi	Internt valideret

Novel use of REDCap to develop an advanced platform to display predictive analytics and track compliance with Enhanced Recovery After Surgery for pancreaticoduodenectomy(200)	2018	Cochran, Allyson R.; Raub, Kyle M.; Murphy, Keith J. et al.	USA	Helbredelsestid efter operation	400	Retrospektivt	Single	Internt valideret
Patient Factors That Matter in Predicting Hip Arthroplasty Outcomes: A Machine-Learning Approach(201)	2021	Sniderman, Jhase; Stark, Roland B.; Schwartz, Carolyn E. et al.	Canada; USA	Alloplastik	160	Prospektivt	Single	Internt valideret
PRCTC: a machine learning model for prediction of response to corticosteroid therapy in COVID-19 patients(119)	2022	Gao, Yue; Xiong, Xiaoming; Jiao, Xiaofei et al.	Kina	Covid-19	666	Retrospektivt	Multi	Eksternt valideret
Predicting Current Glycated Hemoglobin Values in Adults: Development of an Algorithm From the Electronic Health Record(202)	2018	Wells, Brian J.; Lenoir, Kristin M.; Diaz-Garelli, Jose-Franck et al.	USA	Hæmaglobin	22635	Retrospektivt	Single	Internt valideret
Predicting outcomes of chronic kidney disease from EMR data based on Random Forest Regression(203)	2019	Zhao, Jing; Gu, Shaopeng; McDermid, Adam	USA	Nyresygdom	61740	Retrospektivt	Multi	Eksternt valideret
Predicting pharmacotherapeutic outcomes for type 2 diabetes: An evaluation of three approaches to leveraging electronic health record data from multiple sources(204)	2022	Tarumi, Shinji; Takeuchi, Wataru; Qi, Rong et al.	Japan	Diabetes	27843	Retrospektivt	Multi	Internt valideret
Predicting seizures in pregnant women with epilepsy: Development and external validation of a prognostic model(205)	2019	Allotey, John; Fernandez-Felix, Borja M.; Zamora, Javier et al.	England; Spanien; Australien; Holland	Epilepsi	527	Prospektivt	Multi	Eksternt valideret
Predicting the amputation risk for patients with diabetic foot ulceration – a Bayesian decision support tool(206)	2020	Hüser, Jens; Hafer, Guido; Hegemann, Jan et al.	Tyskland	Fodsår	237	Prospektivt	Single	Ingen beskrivelse af validering

Predicting the chance of live birth for women undergoing IVF: a novel pretreatment counselling tool(122)	2016	Dhillon, R. K.; McLernon, D. J.; Smith, P. P. et al.	England	Fertilitet	9915	Retrospektivt	Multi	Temporalt valideret
Predicting the Outcome of Limb Revascularization in Patients With Lower-extremity Arterial Trauma: Development and External Validation of a Supervised Machine-learning Algorithm to Support Surgical Decisions(129)	2020	Perkins, Zane B.; Yet, Barbaros; Sharrock, Anna et al.	England; USA	Revaskularisering	3187	Retrospektivt	Multi	Eksternt valideret
Predicting three-year kidney graft survival in recipients with systemic lupus erythematosus(207)	2011	Tang, Hongying; Poynton, Mollie R.; Hurdle, John F. et al.	Israel; USA	Nyresygdom	4754	Retrospektivt	Multi	Internt valideret
Prediction Models for AKI in ICU: A Comparative Study(208)	2021	Qian, Qing; Wu, Jinming; Wang, Jiayang et al.	Kina	Nyresygdom	17205	Retrospektivt	Single	Temporalt valideret
Prediction of Cardiac Arrest in the Emergency Department Based on Machine Learning and Sequential Characteristics: Model Development and Retrospective Clinical Validation Study(209)	2020	Hong, Sungjun; Lee, Sungjoo; Lee, Jeonghoon et al.	Sydkorea	Hjertefejl	214307	Retrospektivt	Single	Temporalt valideret
Prediction of Length of Stay in the Emergency Department for COVID-19 Patients: A Machine Learning Approach(210)	2022	Etu, E.-E.; Monplaisir, L.; Arslanturk, S.; Masoud, S. et al.	USA	Indlæggelsestid	3301	Retrospektivt	Single	Internt valideret
Prediction of Postoperative Delirium in Geriatric Hip Fracture Patients: A Clinical Prediction Model Using Machine Learning Algorithms(211)	2021	Oosterhoff, Jacobien H. F.; Karhade, Aditya V.; Oberai, Tarandeep et al.	USA	Delirium	28207	Retrospektivt	Multi	Internt valideret
Prediction of Transfusions after Isolated Coronary Artery Bypass Grafting Surgery(212)	2017	Likosky, Donald S.; Paugh, Theron A.; Harrington, Steven D. et al.	USA	Blodtransfusion	20337	Retrospektivt	Multi	Temporalt valideret

Prognostication of metastatic death in uveal melanoma patients: A Markov multi-state model(213)	2018	Eleuteri, Antonio; Taktak, Azzam F. G.; Coupland, Sarah E. et al.	England; USA	Kræft	4161	Retrospektivt	Single	Internt valideret
Random forest can predict 30-day mortality of spontaneous intracerebral hemorrhage with remarkable discrimination(214)	2010	Peng, S.-Y.; Chuang, Y.-C.; Kang, T.-W. et al.	Taiwan	Hjerneblødning	423	Retrospektivt	Single	Internt valideret
Readmission Risk Trajectories for Patients With Heart Failure Using a Dynamic Prediction Approach: Retrospective Study(215)	2019	Jiang, Wei; Siddiqui, Sauleh; Barnes, Sean et al.	USA	Hjertefejl	534	Retrospektivt	Single	Internt valideret
Real-Time Clinical Decision Support Based on Recurrent Neural Networks for In-Hospital Acute Kidney Injury: External Validation and Model Interpretation(216)	2021	Kim, Kipyoo; Yang, Hyeonsik; Yi, Jin-yeong et al.	Sydkorea	Nyresygdom	69081	Retrospektivt	Multi	Eksternt valideret
SOURCE-PANC: A Prediction Model for Patients With Metastatic Pancreatic Ductal Adenocarcinoma Based on Nationwide Population-Based Data(217)	2021	van den Boorn, Héctor G.; Dijksterhuis, Willemieke P. M.; van der Geest, Lydia G. M. et al.	Holland	Kræft	4739	Retrospektivt	Single	Temporalt valideret
Successful external validation of a model to predict other cause mortality in localized prostate cancer(218)	2016	Kent, M.; Penson, D. F.; Albertsen, P. C. et al.	USA	Kræft	~ 10000	Retrospektivt	Multi	Eksternt valideret
The 4-Hour Cairns Sepsis Model: A novel approach to predicting sepsis mortality at intensive care unit admission(219)	2021	Hargovan, Satyen; Gunnarsson, Ronny; Carter, Angus et al.	Australien; Sverige	Sepsis	500	Retrospektivt	Single	Internt valideret
The development of an automated ward independent delirium risk prediction model(220)	2016	de Wit, Hugo A. J. M.; Winkens, Bjorn; Mestres Gonzalvo, Carlota et al.	Holland	Delirium	225	Retrospektivt	Single	Internt valideret

Towards an explainable clinical decision support system for large-for-gestational-age births(221)	2023	Du, Yuhan; Rafferty, Anthony R.; McAuliffe, Fionnuala M. et al.	USA	Fødsler	565	Retrospektivt	Ingen beskrivelse	Eksternt valideret
Uniform criteria for total hip replacement surgery in patients with hip osteoarthritis: A decision tool to guide treatment decisions(222)	2021	Atsma, F.; Molenkamp, O.; Bouma, H et al.	Holland	Alloplastik	364	Retrospektivt	Multi	Internt valideret
Using predicted disease outcome to provide differentiated treatment of early rheumatoid arthritis(223)	2006	de Vries-Bouwstra, Jeska; Le Cessie, Saskia; Allaart, Cornelis et al.	Holland	Rheumatoid arthritis	152	Retrospektivt	Single	Ingen beskrivelse af validering
Using predictive process monitoring to assist thrombolytic therapy decision-making for ischemic stroke patients(224)	2020	Xu, Haifeng; Pang, Jianfei; Yang, et al.	Kina	Apopleksi	1191	Retrospektivt	Single	Internt valideret
Validated risk rule using computerized data to identify males at high risk for fracture(225)	2012	LaFleur, J.; Nelson, R. E.; Yao, et al.	USA	Frakturer	84763	Retrospektivt	Multi	Internt valideret
Validation of two IgA nephropathy risk-prediction tools using a cohort with a long follow-up(226)	2022	Haaskjold, Yngvar Lunde; Lura, Njål Gjærde; Bjørneklett, Rune; Bostad, Leif; Bostad, Lars Sigurd; Knoop, Thomas	Norge	Nyresygdom	306	Retrospektivt	Multi	Eksternt valideret

Bilag 6: Begrænsninger ved udviklingen i de enkelte studier

Navn	BiAS	Datakvalitet	Generaliserbarhed	Mangler eks-tern validering	Begrænsninger ved udfald	Overtilpas-ning	Populations data shift	Problemer med prædiktorer	Rodekasse	Sample size
A Clinical Decision Web to Predict ICU Admission or Death for Patients Hospitalised with COVID-19 Using Machine Learning Algorithms(111)			X	X		X		X		
A Clinical Prediction Model to Assess Risk for Chemotherapy-Related Hospitalization in Patients Initiating Palliative Chemotherapy(146)	X		X					X	X	
A Clinical Prediction Model to Determine Probability of Response to Certolizumab Pegol for Crohn's Disease(112)				X	X					
A Clinical Prediction Model to Predict Heparin Treatment Outcomes and Provide Dosage Recommendations: Development and Validation Study(113)		X		X				X		X
A Deep Learning Model to Predict Knee Osteoarthritis Based on Nonimage Longitudinal Medical Record(147)								X		
A Machine Learning Approach in Predicting Mortality Following Emergency General Surgery(148)	X	X	X	X				X		
A Machine Learning Model for Predicting Mortality within 90 Days of Dialysis Initiation(149)			X							
A Multi-Center Prospective Derivation and Validation of a Clinical Prediction Tool for Severe Clostridium difficile Infection(150)		X		X				X		

A Multivariable Prediction Model to Select Colorectal Surgical Patients for Co-Management(120)			X	X						X
A Novel Computerized Clinical Decision Support System for Treating Thrombolysis in Patients with Acute Ischemic Stroke(151)			X	X						
A Prognostic Model of Non-Small Cell Lung Cancer With a Radiomics Nomogram in an Eastern Chinese Population(152)		X	X					X		
A Prospectively Validated Prognostic Model for Patients with Locally Advanced Squamous Cell Carcinoma of the Head and Neck Based on Radiomics of Computed Tomography Images(153)										X
A smart, practical, deep learning-based clinical decision support tool for patients in the prostate-specific antigen gray zone: model development and validation(114)	X	X		X		X		X		
A time-aware attention model for prediction of acute kidney injury after pediatric cardiac surgery(154)			X	X				X		
A traumatic brain injury prognostic model to support in-hospital triage in a low-income country: a machine learning-based approach(155)				X						X
A web based dynamic MANA Nomogram for predicting the malignant cerebral edema in patients with large hemispheric infarction(156)		X	X		X					

Accuracy of a Risk Alert Threshold for ICU Hypoglycemia: Retrospective Analysis of Alert Performance and Association With Clinical Deterioration Events(157)			X							
Acute Coronary Syndrome Risk Prediction Based on GRACE Risk Score(158)		X								
Analysis of prognostic factors and construction of prognostic models for triple-positive breast cancer(115)	X			X				X		
Applied machine learning for the risk-stratification and clinical decision support of hospitalised patients with dengue in Vietnam(159)		X	X					X		
Clinical scoring model based on age, NIHSS, and stroke-history predicts outcome 3 months after acute ischemic stroke(160)	X		X	X				X		X
Comparison of nomogram with random survival forest for prediction of survival in patients with spindle cell carcinoma(161)		X		X						X
Construction of a Tool to Predict Overall Survival of Patients With Primary Spinal Tumors After Surgical Resection: A Real-World Analysis Based on the Surveillance, Epidemiology, and End Results Database(162)	X	X						X		
Corticosteroid Randomization after Significant Head Injury and International Mission for Prognosis and Clinical Trials in Traumatic Brain Injury Models Compared with a Machine Learning-	X			X			X			

Based Predictive Model from Tanzania(163)									
Creating an automated trigger for sepsis clinical decision support at emergency department triage using machine learning(164)		X		X					
Deep Learning to Predict EGFR Mutation and PD-L1 Expression Status in Non-Small-Cell Lung Cancer on Computed Tomography Images(165)		X	X	X				X	
Developing a Prognostic Model to Predict Mortality in Patients with Acute Bacterial Meningitis(166)		X	X	X					
Developing machine learning models to personalize care levels among emergency room patients for hospital admission(167)		X	X	X					
Development and external validation of a nomogram predicting prognosis of upper tract urothelial carcinoma after radical nephroureterectomy(168)	X	X							X
Development and external validation of an admission risk prediction model after treatment from early intervention in psychosis services(130)						X			X
Development and internal validation of a machine learning prediction model for low back pain non-recovery in patients with an acute episode consulting a physiotherapist in primary care(126)				X		X		X	X

Development and Internal Validation of Machine Learning Algorithms for Preoperative Survival Prediction of Extremity Metastatic Disease(169)		X	X	X				X		
Development and preliminary validation of a risk prediction model for chemotherapy-related nausea and vomiting(170)	X		X					X		X
Development and Validation of a Clinical Prediction Model for Sleep Disorders in the ICU: A Retrospective Cohort Study(171)	X	X	X							
Development and validation of a nomogram for evaluating the incident risk of carotid atherosclerosis in patients with type 2 diabetes(172)	X	X	X	X						X
Development and Validation of a Nomogram for the Prediction of Unfavorable Treatment Outcome Among Multi-Drug Resistant Tuberculosis Patients in North West Ethiopia: An Application of Prediction Modeling(173)		X		X				X		
Development and validation of a nomogram to predict anastomotic leakage after esophagectomy for esophageal carcinoma(174)		X	X	X						
Development and Validation of a Prediction Model for Need for Massive Transfusion During Surgery Using Intraoperative Hemodynamic Monitoring Data(116)	X	X		X				X		
Development and Validation of a Prediction Model for Pain and Functional Outcomes After Lumbar Spine Surgery(175)		X	X		X			X		

Development and validation of an explainable artificial intelligence-based decision-supporting tool for prostate biopsy(176)		X	X	X						
Development and Validation of Clinical Scoring Tool to Predict Outcomes of Treatment With Vedolizumab in Patients With Ulcerative Colitis(121)		X		X						
Development and validation of multivariable prediction models of serological response to SARS-CoV-2 vaccination in kidney transplant recipients(177)					X					
Development and validation of nomograms for prediction of overall survival and cancer-specific survival of patients of colorectal cancer(178)		X		X						
Development and Validation of Prediction Scores for Early Mortality at Transition to Dialysis(179)		X	X				X			
Development of a machine-learning based model for predicting multidimensional outcome after surgery for degenerative disorders of the spine(180)		X	X	X						
Development of a model to predict combined antidepressant medication and psychotherapy treatment response for depression among veterans(181)	X									X
Development of a Prediction Model for Antibiotic-Resistant Urinary Tract Infections Using Integrated Electronic Health Records from Multiple Clinics in North-Central Florida(182)		X	X					X		

Development of machine learning-based clinical decision support system for hepatocellular carcinoma(183)			X	X						X
Dr. Answer AI for prostate cancer: Clinical outcome prediction model and service(184)		X								
Dynamic prediction of mortality after traumatic brain injury using a machine learning algorithm(185)		X	X							
Editor's Choice - Development and Validation of a Multivariable Prediction Model of Peri-operative Mortality in Vascular Surgery: The New Zealand Vascular Surgical Risk Tool (NZRISK-VASC)(186)		X	X	X						
External Validation and Optimization of the SPRING Model for Prediction of Survival After Surgical Treatment of Bone Metastases of the Extremities(127)			X	X		X				
Fairness in the prediction of acute postoperative pain using machine learning models(187)		X	X	X						X
FDG PET versus CT radiomics to predict outcome in malignant pleural mesothelioma patients(128)		X		X		X		X		X
Feasibility of machine learning-based modeling and prediction using multiple centers data to assess intrahepatic cholangiocarcinoma outcomes(188)		X	X							X
From reactive to proactive tube feeding during chemoradiotherapy for head and neck cancer: A		X	X	X				X	X	

clinical prediction model-based approach(189)									
iChoose Kidney: a Clinical Decision Aid for Kidney Transplantation vs. Dialysis Treatment(117)	X		X	X				X	
Improvement of APACHE II score system for disease severity based on XGBoost algorithm(190)		X		X				X	
Lessons of War: Turning Data Into Decisions(191)			X						X
Leveraging Artificial Intelligence to Improve Chronic Disease Care: Methods and Application to Pharmacotherapy Decision Support for Type-2 Diabetes Mellitus(192)	X			X	X				X
Machine Learning Approach to Reduce Alert Fatigue Using a Disease Medication-Related Clinical Decision Support System: Model Development and Validation(193)		X	X					X	
Machine learning to predict extubation outcome in premature infants(194)		X						X	
Machine Learning to Predict Facial Dehiscence after Exploratory Laparotomy Surgery(195)		X					X	X	
Machine learning using the extreme gradient boosting (XGBoost) algorithm predicts 5-day delta of SOFA score at ICU admission in COVID-19 patients(124)				X					
Machine Learning-Based Modeling of Ovarian Response and the			X					X	

Quantitative Evaluation of Comprehensive Impact Features(196)									
Machine learning-based prediction model for late recurrence after surgery in patients with renal cell carcinoma(197)			X	X					X
Massive external validation of a machine learning algorithm to predict pulmonary embolism in hospitalized patients(118)		X		X					
Medication-related harm in older adults following hospital discharge: development and validation of a prediction tool(198)				X					
Novel machine learning to predict asthma acute exacerbation(199)	X	X	X						
Novel use of REDCap to develop an advanced platform to display predictive analytics and track compliance with Enhanced Recovery After Surgery for pancreaticoduodenectomy(200)			X	X			X		
Patient Factors That Matter in Predicting Hip Arthroplasty Outcomes: A Machine-Learning Approach(201)		X	X						
PRCTC: a machine learning model for prediction of response to corticosteroid therapy in COVID-19 patients(119)		X		X				X	X
Predicting Current Glycated Hemoglobin Values in Adults: Development of an Algorithm From the Electronic Health Record(202)		X		X					

Predicting outcomes of chronic kidney disease from EMR data based on Random Forest Regression(203)	X								X		
Predicting pharmacotherapeutic outcomes for type 2 diabetes: An evaluation of three approaches to leveraging electronic health record data from multiple sources(204)	X		X	X							
Predicting seizures in pregnant women with epilepsy: Development and external validation of a prognostic model(205)			X						X		X
Predicting the amputation risk for patients with diabetic foot ulceration – a Bayesian decision support tool(206)			X	X					X		X
Predicting the chance of live birth for women undergoing IVF: a novel pretreatment counseling tool(122)	X		X	X				X	X		
Predicting the Outcome of Limb Revascularization in Patients With Lower-extremity Arterial Trauma: Development and External Validation of a Supervised Machine-learning Algorithm to Support Surgical Decisions(129)	X	X	X	X		X		X			
Predicting three-year kidney graft survival in recipients with systemic lupus erythematosus(207)		X									X
Prediction Models for AKI in ICU: A Comparative Study(208)		X	X								X
Prediction of Cardiac Arrest in the Emergency Department Based on Machine Learning and Sequential Characteristics:			X	X	X				X		

Model Development and Retrospective Clinical Validation Study(209)										
Prediction of Length of Stay in the Emergency Department for COVID-19 Patients: A Machine Learning Approach(210)		X	X					X		
Prediction of Postoperative Delirium in Geriatric Hip Fracture Patients: A Clinical Prediction Model Using Machine Learning Algorithms(211)	X	X	X	X				X		
Prediction of Transfusions after Isolated Coronary Artery Bypass Grafting Surgery(212)		X	X	X				X		
Prognostication of metastatic death in uveal melanoma patients: A Markov multi-state model(213)				X				X		
Random forest can predict 30-day mortality of spontaneous intracerebral hemorrhage with remarkable discrimination(214)			X					X		
Readmission Risk Trajectories for Patients With Heart Failure Using a Dynamic Prediction Approach: Retrospective Study(215)		X								X
Real-Time Clinical Decision Support Based on Recurrent Neural Networks for In-Hospital Acute Kidney Injury: External Validation and Model Interpretation(216)		X	X							
SOURCE-PANC: A Prediction Model for Patients With Metastatic Pancreatic Ductal Adenocarcinoma Based on Nationwide Population-Based Data(217)	X				X			X		

Successful external validation of a model to predict other cause mortality in localized prostate cancer(218)		X						X		
The 4-Hour Cairns Sepsis Model: A novel approach to predicting sepsis mortality at intensive care unit admission(219)		X	X	X						X
The development of an automated ward independent delirium risk prediction model(220)				X						
Towards an explainable clinical decision support system for large-for-gestational-age births(221)			X							
Uniform criteria for total hip replacement surgery in patients with hip osteoarthritis: A decision tool to guide treatment decisions(222)			X	X						
Using predicted disease outcome to provide differentiated treatment of early rheumatoid arthritis(223)		X		X						
Using predictive process monitoring to assist thrombolytic therapy decision-making for ischemic stroke patients(224)		X	X	X				X		X
Validated risk rule using computerized data to identify males at high risk for fracture(225)		X	X	X						X
Validation of two IgA nephropathy risk-prediction tools using a cohort with a long follow-up(226)		X	X							X