



Det Helsevitenskapelige fakultet

Bruk av naturlig språkprosessering i psykiatri: En systematisk kartleggingsoversikt

En systematisk kartleggingsoversikt over studier gjort på bruk av naturlig språkprosessering for analyse av elektronisk pasientjournal til pasienter med psykoselidelser og affektive lidelser.

Live Risnes Haglund

Masteroppgave i profesjonsstudiet medisin, MED-3950, mai 2023

Hovedveileder: Brita Elvevåg – Forskningsgruppe psykiatri, Institutt for klinisk medisin, UiT

Biveileder: Terje Bektesevic Holmlund – Forskningsgruppe psykiatri, Institutt for klinisk medisin, UiT

Forord

Hensikten med Masteroppgaven medisin ved UiT – Norges arktiske universitet er at vi som studenter skal initiere, planlegge og gjennomføre et vitenskapelig arbeid innen et valgt tema. Vi skal øve oss på datainnhenting, kunnskapsvurdering og sammenfatting av en medisinsk vitenskapelig rapport med fokus på informasjonsformidling, kildebruk og korrekt språk.

I et hav av tema vi kunne velge, hadde jeg et ønske om å skrive en oppgave innen psykiatri, da dette er noe jeg har hatt interesse for gjennom hele studiet. Jeg tok derfor kontakt med Brita Elvevåg ved *Forskningsgruppe psykiatri ved Institutt for klinisk medisin*, som sa ja til å være min hovedveileder. Hun tok så kontakt med Terje Holmlund ved samme institutt som ville være biveileder for meg. De to kom raskt med forslag til mulige prosjekter vi kunne gjøre sammen.

Opprinnelig var det planlagt at denne oppgaven skulle være en empirisk studie som så på kjønnsbias i skriftlig pasientjournal hos pasienter med affektive lidelser og psykoselidelser. Studien skulle se på antall ord i «status presens», et gitt segment i inntakjournal for psykiatriske pasienter, og sammenligne antallet i journaler fra kvinner og menn. Gjennomføringen av det opprinnelige prosjektet ble imidlertid vanskelig, hovedsakelig som følge av tiden det ville tatt å få tilgang til det nødvendige datamaterialet, spesielt grunnet behov for godkjenning fra etisk komité og personvernombud. I forlengelse av den opprinnelige planen ble det derfor besluttet å gjøre en litteraturstudie med hensikt å undersøke hvordan andre studier har brukt journaldata fra psykiatriske pasienter, og det er dette som vil presenteres i denne masteroppgaven.

Masteroppgaven min har ikke mottatt økonomisk finansiering.

Brita Elvevåg har vært til spesielt stor hjelp ved forming av oppgaven, korrekturlesning, og å se oppgaven i en større sammenheng. Terje B. Holmlund har vært til stor hjelp med litteratursøket og etterarbeidet med dette. Tusen takk til dere begge for at dere fra start har vært nysgjerrige og motiverende. Takk for tilliten dere har vist meg, og for all erfaring dere har delt. Til slutt, tusen takk for at dere har lært meg å stole på prosessen. Uten deres bidrag ville ikke dette vært mulig.

Live Risnes Haglund, 26. mai 2023

Live Risnes Haglund

Innholdsfortegnelse

Forord	1
Sammendrag	3
Forkortelser	4
1 Innledning	5
1.1 Kunstig intelligens, maskinl�ring og naturlig spr�kprosessering	5
1.2 Tekstformat i elektronisk pasientjournal	6
1.3 Psykoselidelser og affektive lidelser	7
1.4 Problemstilling og form�l med unders�kelsen	7
2 Materialer og metode	8
2.1 Forskningssp�rsm�l	8
2.2 S�kestrategi for identifisering av relevante studier	9
2.3 Inklusjonskriterier for seleksjon av relevante studier	10
2.4 Kartlegging av data fra relevante studier	11
2.5 Oppsummering av resultater og funn	12
3 Resultater	12
3.1 Lokalisering og tidspunkt for utgivelse av studiene	14
3.2 Pasientpopulasjoner og dokumenttyper som ble brukt	16
3.3 Hensikt med studiene og bruk av naturlig spr�kprosessering	17
3.4 Informasjon om programvare eller metode for naturlig spr�kprosessering	23
4 Diskusjon	25
4.1 Oppsummering av viktigste funn	25
4.2 Forst�else av funnene i forhold til form�l med unders�kelsen	26
4.3 Styrker og svakheter ved unders�kelsen	26
4.4 Mulige implikasjoner av funn	28
5 Konklusjon	32
Referanser	33
Vedlegg	37
Vedlegg 1 – Referanser alle resultater litteraturs�k PubMed	37

Sammendrag

Bakgrunn: Bruk av kunstig intelligens (AI) har et stadig økende fokus, også i helsevesenet. En metode som virker lovende, er naturlig språkprosessering (NLP), som kan brukes til analysering av skriftlig tekst, for eksempel tekst i elektroniske pasientjournaler. Denne undersøkelsen har som formål å undersøke forskning som er gjort på bruk av naturlig språkprosessering for analysering av elektroniske journaler fra pasienter med alvorlige psykiske lidelser, som affektive lidelser og psykoselidelser. Den overordnede hensikten med dette, er å få et inntrykk av om noe av forskningen som er gjort har fokus på forbedring av pasientenes helsesituasjon.

Materiale og metode: Det ble gjennomført en systematisk kartleggingsoversikt («scoping review»). Litteratursøket ble gjort i én database for medisinsk forskning, PubMed, med søketermene «psychiatry», «electronic medical records» og «natural language processing». Søket var ikke avgrenset i tid. For at en artikkel skulle bli inkludert i undersøkelsen måtte den være empirisk, ha utført analyser på journaldata i fritext, ha brukt elektroniske journaler fra psykiatriske pasienter med psykoselidelser og/eller affektive lidelser og være skrevet på engelsk språk.

Resultater: Litteratursøket resulterte i totalt 211 unike artikler, av disse oppfylte 37 artikler inklusjonskriteriene i kartleggingsoversikten, og ble undersøkt videre. De fleste av studiene var gjennomført i Storbritannia og USA. Størrelsen på studiepopulasjonen varierte mye, fra noen hundre til flere hundre tusen inkluderte pasienter i studiene. Det var lite av forskningen som var gjort på spesifikke dokumenttyper fra pasientjournal, som for eksempel epikriser eller inntakstjournaler. Hensikten for studiene varierte mye, men kunne deles inn i noen felles kategorier: 1) identifisering av informasjon fra journal, 2) kvantitative undersøkelser av populasjonen eller journalene, 3) seleksjon av pasienter til kohorter og 4) vurdering av risiko.

Fortolkning: Det trengs mer grunnforskning før teknologi for naturlig språkprosessering til analyse av elektronisk journal vil bidra med forbedring av psykiatriske pasienters helsesituasjon.

Forkortelser

AI	Kunstig intelligens (artificial intelligence)
NLP	Naturlig språkprosessering (natural language processing)
ML	Maskinlæring (machine learning)
SLaM	South London and Maudsley NHS Foundation Trust
CRIS	The Clinical Record Interactive Search
BERT	Bidirectional Encoder Representations from Transformers
GATE	General Architecture for Text Engineering
NILE	Narrative Information Linear Extraction
LDA	Latent Dirichlet Allocation
MedCAT	Medical Concept Annotation Tool
NLTK	Natural Language Toolkit
cTAKES	Clinical Text Analysis and Knowledge Extraction System
NER	Named Entity Recognition
TF-IDF	Term Frequency – Inverse Document Frequency
LSTM	Long Short-Term Memory
GloVe	Global Vectors for Word Representation
TH	Terje B. Holmlund
LH	Live R. Haglund

1 Innledning

Utredning og diagnostikk av psykiske lidelser er i all hovedsak basert på samtale mellom pasient og kliniker. (1, 2). I Norge skal informasjonen som kommer frem i slike samtaler dokumenteres i et journalsystem (3). Der utredningen av somatiske lidelser ofte kan støttes av mer objektive undersøkelser, som laboratorieprøver og bildediagnostikk, vil denne typen undersøkelser først og fremst bidra til å ekskludere organiske årsaker til sykdom hos psykiatriske pasienter, ikke til diagnostikk i seg selv. Det har i senere tid vært stadig mer fokus på at teknologiske verktøy som kunstig intelligens («artificial intelligence»: AI), maskinlæring («machine learning»: ML) og naturlig språkprosessering («natural language processing»: NLP) kan bli viktige bidragsytere for bedre og mer strukturert diagnostikk og behandling i psykiatri(4). Et område det har blitt knyttet særlig forventning til er analyse av tekst i pasientjournaler, denne nedtegnede informasjonen som oppstår i samtalen mellom pasient og kliniker. For å få et realistisk bilde av hvordan denne teknologien ligger an til å kunne bidra forbedring av helsetjenester i fremtiden, er det nødvendig med noen begrepsavklaringer og en systematisk gjennomgang av hvordan fagfeltet ser ut nå.

1.1 Kunstig intelligens, maskinlæring og naturlig språkprosessering

For å få et bilde av hva teknologiske verktøy som AI, ML og NLP kan bidra med i helsevesenet, er det viktig å ha et inntrykk av hva disse begrepene innebærer. Det finnes ikke en entydig konsensus for hvordan disse begrepene skal defineres, men de kan for eksempel defineres ut ifra teknologien de baserer seg på, eller ut ifra hva de brukes til. Forklaringene som følger, er altså bare en måte å definere og forklare uttrykkene på.

Kunstig intelligens (AI) kan forklares som et generelt konsept basert på dannelsen av teknologiske ekspertsystemer for å utføre ulike oppgaver. Systemene innehar intelligent atferd, og er i stand til å lære, forklare og hjelpe brukerne av systemet (5)¹. Maskinlæring (ML) er en samling metoder som kan regnes som en undergruppe av AI.

Maskinlæringsmodeller er statistiske systemer som enten kan lære seg egenskaper ved et datasett og forskjell i viktighet mellom disse egenskapene, som er predikative for en gitt variabel, eller som kan lære seg viktigheten av egenskaper som er definert av brukeren. Når egenskapene en bruker vil utforske er gitt, kan modellen predikere utfall eller klinisk

¹ Oversettelse basert på tabell 1 fra C. Chandler et al. «Using Machine Learning in Psychiatry: The Need to Establish a Framework That Nurtures Trustworthiness» (5)

klassifikasjon på nye og usette datasett (5)². Naturlig språkprosessering (NLP) er en form for maskinlæring som brukes til å analysere tekst og tale (6). Det er denne formen for kunstig intelligens som kan komme spesielt til nytte ved analyse av elektroniske pasientjournaler, og på sikt, utvikling og etablering av verktøy for diagnostikk og beslutningsstøtte i psykiatri.

1.2 Tekstformat i elektronisk pasientjournal

I utredningen av pasienter med psykiatrisk sykdom blir det ofte gjennomført en rekke samtaler og kartlegginger. I dagens praksis foreligger dokumentasjonen fra disse undersøkelsene i tekstformat i elektroniske pasientjournaler. Formatet til disse tekstene er i hovedsak det man kan kalle fritext, eller ustrukturert tekst. Dette vil si skriftlig tekst som en kliniker skriver, eller som blir diktert, inn i et journaldokument. Språket og teksten som brukes har ingen andre begrensninger enn at det skal forstås av andre klinikere og, i økende grad, av pasienten selv. Som motsetning til dokumenter med denne typen tekst, finnes det i journaler strukturerte data, for eksempel bestemte koder registrert for datoer, diagnoser og kjønn. Slik informasjon føres gjerne inn i journalsystemer ved avkryssing av bokser eller nedtrekksmenyer i elektroniske skjema. Mellom disse ytterpunktene finnes et spekter av ulike formater for hvordan informasjon blir registrert i journaler. Et av disse formatene er semistrukturert tekst, som er en mellomting mellom strukturert tekst og fritext. Semistrukturert tekst sees typisk i journaldokumenter med har en klar og forhåndsdefinert struktur. Et eksempel på et slikt dokument er sykepleienotater, som kan ha tydelige og nummererte seksjoner med informasjon om ting som sårstell, matinntak eller mobiliseringsevne. De ulike metodene for registrering i journal har alle sine fordeler og ulemper, men det er særlig i fritextformatet det har vært knyttet forventinger til at nye metoder for kunstig intelligens skal kunne komme til nytte.

For en kliniker er det flere fordeler med å kunne formulere seg i fritext. I motsetning til å forholde seg til et skjema eller dokumenter med avkryssing, gir fritext forfatteren mulighet til å være mer fleksibel i format og ordbruk, og gir også mulighet til å formulere mer nyanserte bilder av en situasjon. Det finnes likevel mange utfordringer knyttet til å gjøre videre analyser på fritext med for eksempel verktøy for naturlig språkprosessering. Eksempler på slike utfordringer kan blant annet være at det i teksten kan være brukt

² Oversettelse basert på tabell 1 fra C. Chandler et al. «*Using Machine Learning in Psychiatry: The Need to Establish a Framework That Nurtures Trustworthiness*» (5)

forkortelser eller synonymer for vanlige ord, det kan være skrivefeil, teksten kan være på ulike språk eller grammatikken kan være atypisk, noe som kan gjøre det vanskelig for NLP-programmer å fungere adekvat i ulike sammenhenger (7). I tillegg er også formatet for fritekstdokumenter fleksibelt med tanke på tegnsetting og struktur (7).

1.3 Psykoselidelser og affektive lidelser

Psykatri omfavner et bredt spekter av tilstander, noen mer alvorlige enn andre. I Norge benyttes diagnosesystemet ICD-10 for klassifisering og diagnostisering av psykiske lidelser og atferdsforstyrrelser (8, 9). Kodeverket er internasjonalt og gis ut av World Health Organization (WHO). I denne oppgaven vil fokuset primært være på bruk av elektronisk pasientjournal for pasienter med diagnoser kategorisert under *Schizofreni, schizotyp lidelse og paranoide lidelser* (F20-29), videre referert til som *psykoselidelser*, og *Affektive lidelser/stemningslidelser* (F30-39), videre referert til som *affektive lidelser*. En viktig årsak til at det er akkurat disse diagnosene som fokuseres på, er at det er diagnoser under disse kategoriene som typisk refereres til som alvorlige psykiske lidelser (10). Dette er lidelser med høy mortalitet (11, 12), noe som gjør tidlig og riktig diagnostikk svært viktig. For å få et inntrykk av hvordan kunstig intelligens og naturlig språkprosessering kan brukes til å gjøre endringer i helsetjenester for psykiatriske pasienter, er forskning som omhandler disse lidelsene et godt utgangspunkt.

1.4 Problemstilling og formål med undersøkelsen

Formålet med denne undersøkelsen var å gjøre en vurdering av litteraturen som finnes på forskning som bruker naturlig språkprosessering for analysering av elektroniske journaler til pasienter med alvorlig psykisk sykdom, som psykoselidelser og affektive lidelser. For å gjøre dette, ble metoden for systematiske kartleggingsoversikter («scoping reviews») benyttet. Det var ønskelig at undersøkelsen skulle gi en oversikt over litteraturen i feltet, og peke på de mest vanlige og lovende metodene innenfor bruk av naturlig språkprosessering og elektroniske pasientjournaler i psykiatri, se mer detaljert formålsbeskrivelse under *metode*. En slik oversikt er viktig, da dette er et fagfelt med en teknologi i rask utvikling. Begrepet «kunstig intelligens» brukes ofte om teknologien når den omtales i media, og kan føre til både entusiasme og bekymring. Å få et realistisk bilde av hva som *faktisk* har vært gjort av forskning og utvikling, vil kunne gi beslutningstakere i helsevesenet et bedre grunnlag i

vurderinger rundt fremtidig bruk av slik teknologi. Dette er særlig viktig i et felt som psykiatri, der teknologiske verktøy som utvikles kan få konsekvenser for særlig sårbare mennesker. Et overordnet mål med denne undersøkelsen var derfor å se på om noe av den forskningen som har vært gjort på fagfeltet hadde som hensikt å forbedre psykiatriske pasienters helsesituasjon. Med «forbedring av psykiatriske pasienters helsesituasjon» menes i denne sammenhengen forskning som har hatt fokus på å bedre enkeltpasienters helse, enten i form av tidligere og mer nøyaktig diagnostikk, eller i form av bedre behandling.

2 Materialer og metode

Undersøkelsen ble gjennomført som en systematisk kartleggingsoversikt (oversatt fra engelsk, «scoping review»). Valget falt på dette studiedesignet da undersøkelsen ikke først og fremst har som mål å vurdere evidens for behandling eller oppsummere forskning for å svare på et spesifikt forskningsspørsmål, som ved en eksempelvis en systematisk litteratur studie, men å *belyse litteraturen som foreligger i et gitt fagområde* (13-15). Det skal ikke gjøres en kritisk vurdering av de enkelte artiklene søket resulterer i, men heller *gi et overblikk over det valgte fagområdet for å kunne beskrive trender i forskningen som er gjort*. En systematisk kartleggingsoversikt er mer fleksibel enn en systematisk litteraturstudie, og muligheten for å tilpasse metoden underveis i arbeidet ut ifra resultatene er større (14). En av årsakene til å velge systematisk kartleggingsoversikt som studiedesign, er at disse kan brukes der en ønsker å få oversikt, belyse eller kartlegge et fagområde. Metoden for denne undersøkelsen tar utgangspunkt i Arksey og O'Malleys metodologiske rammeverk for «scoping reviews» (14), men følger ikke PRISMA-ScR sjekklister for «Scoping reviews» (16).

2.1 Forskningsspørsmål

Det første som ble gjort var å etablere et forskningsspørsmål. Det viktigste målet for undersøkelsen, og dermed forskningsspørsmålet, var å få en oversikt over hva slags forskning som er gjort på *bruk av naturlig språkprosessering i elektroniske pasientjournaler fra psykiatriske pasienter med affektive lidelser eller psykoselidelser*.

Det var i utgangspunktet ønskelig å finne ut *hvor* disse studiene var gjort, *hvor store* datamengder eller pasientpopulasjoner som var brukt i studiene, *hvilke* dokumenter fra elektronisk pasientjournal som var brukt og *hvilken* programvare og metode for naturlig språkprosessering de ulike studiene hadde brukt. I tillegg var det et ønske å få et overblikk

over hva informasjonen i pasientjournal brukes til i forskningen, altså hva slags type forskningsspørsmål som var stilt, og hva som var målet med bruken av språkteknologisk programvare.

2.2 Søkestrategi for identifisering av relevante studier

Litteratursøket ble gjort i den elektroniske databasen PubMed. Valget falt på denne databasen av flere årsaker. Den viktigste årsaken var at litteraturen som finnes i denne databasen er medisinsk relevant (17), noe som var viktig for søket, da det var ønskelig å vurdere studier gjort i det medisinske fagområdet psykiatri. Litteraturen som finnes i PubMed er ofte fagfellevurdert av personer med forståelse og interesse for medisinsk forskning. Det ble i tillegg vurdert å også søke i andre databaser for litteratur med medisinsk bakgrunn, for eksempel MEDLINE, EMBASE eller PsycINFO, men det ble tatt et bevisst valg i å ikke benytte disse, for å gjøre en tidlig avgrensning i søket, da både tid og personressurser for gjennomgang av litteratur var begrenset.

For å likevel få med mest mulig relevant litteratur fra den valgte databasen ble det gjort et bredt søk. Søketermene som ble valgt var derfor få, men generelle, og tok utgangspunkt i det formulerte forskningsspørsmålet. Da fokuset skulle være på forskning gjort på journalene til psykiatriske pasienter, var det naturlig at et av søkeordene ble «psykiatri» («psychiatry»). Videre var det viktig at søket fanget opp studier som var gjort på bruk av elektroniske pasientjournaler, og neste søkeord ble derfor «elektroniske pasientjournaler» («electronic medical records»). Det siste søkeordet som ble valgt var basert på at studiene skulle ha brukt programvare for å analysere journaldata i fritekst. En vanlig metode for dette er naturlig språkprosessering («natural language processing») og dette ble derfor vår siste søkeord. Det endelige søket ble som følger: «((psychiatry) AND (electronic medical records) AND (natural language processing))».

Litteratursøket ble gjennomført den 29. september 2022³. Søket ble ikke avgrenset i tid, og gikk fra og med 1.januar 1999 til og med dato for søket. Dette var også fordi det var ønskelig at det skulle være bredt, i tillegg til at en ved å unngå tidsavgrensning kunne få et inntrykk av hvordan utviklingen på fagområdet hadde vært over tid. Det ble heller ikke gjort avgrensning i

³ Det ble den 23.05.23 gjennomført et søk i PubMed med samme søkeord som i det opprinnelige litteratursøket, men for perioden fra 29.09.22-23.05.23. Dette søket resulterte i 22 treff, fem treff fra 2022 og 17 treff fra 2023. Disse treffene er ikke inkludert i denne undersøkelsen.

form av hvor i kildene søketermene måtte befinne seg i, de kunne altså befinne seg i både tittel, abstrakt og fulltekst.

2.3 Inklusjonskriterier for seleksjon av relevante studier

Alle resultatene fra litteratursøket i PubMed ble satt inn i Microsoft Excel for videre seleksjon. Gjennomgang av alle artiklene for å avgjøre hvilke som var relevante for studien ble gjort av en person (LH). I de tilfellene der det var tvil om hvilken av kriteriene en artikkel oppfylte, ble det gjort en gjennomgang av den enkelte artikkelen i fellesskap med en annen person (TH). Dette for å sikre korrekt kategorisering av artiklene slik at resultatene ble korrekte. Seleksjonen skulle ta utgangspunkt i tittel og abstrakt i artiklene for å selektere de relevante studiene, men det ble underveis i selekteringen tydelig at disse delene av artiklene ikke gav nok informasjon for kategorisering. Det ble derfor i de fleste artiklene behov for å lese gjennom metode-delen i tillegg.

For å sikre at litteraturen skulle være relevant for å besvare forskningsspørsmålet som var bestemt, ble det satt en rekke inklusjonskriterier før seleksjonen. En systematisk kartleggingsoversikt kan vurdere litteratur uavhengig av studiedesign (14), men i denne studien var det ønskelig å se kun på empiriske studier, og ikke reviews, oversiktsartikler eller andre typer studier. At en studie var empirisk, ble derfor et inklusjonskriterium. Det var noen tilfeller der det var tvil om en studie skulle bli regnet som empirisk eller ikke, disse tilfellene ble gjennomgått av to personer sammen (LH og TH). For at en studie ble kategorisert som empirisk måtte den inneholde kvantitative analyser og observasjoner av pasienter, eller karakteristikker og funksjon av teknologien som var brukt. Teoretiske utregninger eller rene litteraturstudier ble ikke regnet som empiriske.

Det neste inklusjonskriteriet var at artikkelen presenterte forskning på journaldata i fritekst, ikke strukturert tekst. Også her var det noen gråsoner der artiklene ble vurdert av to personer (LH og TH). Det var i noen tilfeller tvil om teksten som var analysert var journaldata eller ikke, for eksempel ved korrespondanser mellom pasient og behandler. Disse ble vurdert til å være fritekst, da det er sannsynlig at en slik korrespondanse vil være lagret i journal.

Videre var det avgjørende for om en artikkel ble selektert til de utvalgte studiene at studiepopulasjonen i forskningen var psykiatriske pasienter, og mer spesifikt pasienter med affektive lidelser eller psykoselidelser. At pasientene hadde en diagnose under en av disse kategoriene kunne være presisert med enten diagnosenavn eller diagnosekode.

Diagnosekodene kunne være kode F20-29 *Schizofreni, schizotyp lidelse og paranoide lidelser* eller F30-39 *Affektive lidelser/stemningslidelse* fra ICD-10 eller tilsvarende koder i ICD-8, ICD-9 eller DSM-systemet.

Underveis i kategoriseringen ble det tydelig at ikke alle studiepopulasjoner var selektert på bakgrunn av diagnose, og det ble derfor behov for å utvise noe skjønn for å kunne selektere de riktige artiklene. Dette ble avklart mellom to personer (LH og TH) slik at det skulle gi likt utfall for studiene der diagnose ikke var presisert. I noen av artiklene var studiepopulasjon selektert på bakgrunn av eksempelvis innleggelsesårsak eller dødsårsak. Der denne sannsynliggjorde psykiatrisk sykdom ble artikkelen inkludert videre. Eksempel på dette er innleggelsesårsak «suicidalforsøk» eller dødsårsak «suicid» som begge kan tyde på mulig bakenforliggende psykiatrisk sykdom. I andre artikler var studiepopulasjon selektert på bakgrunn av forskrevne medikamenter. Der det forskrevne medikamentet sannsynliggjorde bakenforliggende psykiatrisk sykdom ble artiklene inkludert. Eksempel på dette er at bruk av antidepressiva gjør det sannsynlig at en pasient har en affektiv lidelse, og at bruk av antipsykotika øker sannsynligheten for at en pasient har en psykoselidelse.

Artikler som var publisert på andre språk enn engelsk ble ikke inkludert.

Under arbeidet med seleksjon av relevante artikler ble det også samlet noe informasjon for å forenkle det videre arbeidet med resultatene. Dette var informasjon om opprinnelsesland til studiene. I tillegg ble underveis i seleksjonen tydelig at mange av de relevante studiene var gjort på felles datamateriale, noe som derfor var ønskelig å notere.

2.4 Kartlegging av data fra relevante studier

Alle de utvalgte studiene ble så gjennomgått i sin helhet for å gjøre en vurdering av innhold med tanke på spørsmålene vi ønsket å besvare (se avsnitt *forskningsspørsmål* over). Studiene ble satt inn i en tabell i, representert av *Tabell 1*, *Tabell 2* og *Tabell 3*, og informasjonen ble skrevet inn i denne underveis i gjennomgangen. Gjennomgangen ble hovedsakelig gjennomført av én person (LH), men der det var utfordring knyttet til å finne den ønskede informasjonen, ble studiene gått gjennom sammen med en annen person (TH).

I tabellen ble notert studienes opprinnelsesland, hvilke diagnose(r) studiene hadde fokusert på, størrelsen på studiepopulasjonen, hovedhensikten med studien, hvilke type dokumenter fra journal som var brukt og metode for NLP som var benyttet. Det ble også notert om studiene

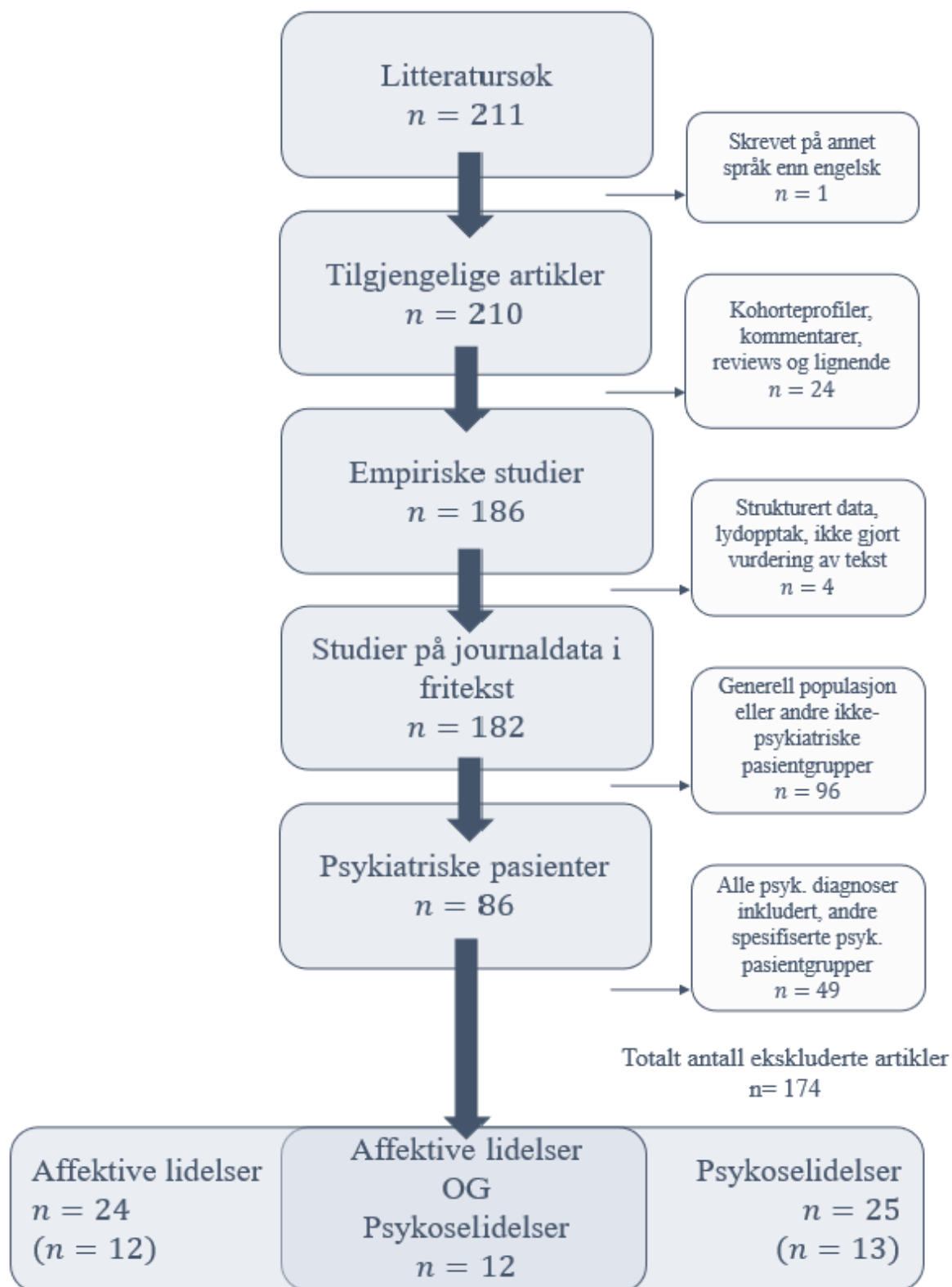
hadde populasjon fra et bestemt utvalg fra Storbritannia, (SLaM), da dette var et utvalg som tidlig skilte seg ut i gjennomgangen av de relevante studiene.

2.5 Oppsummering av resultater og funn

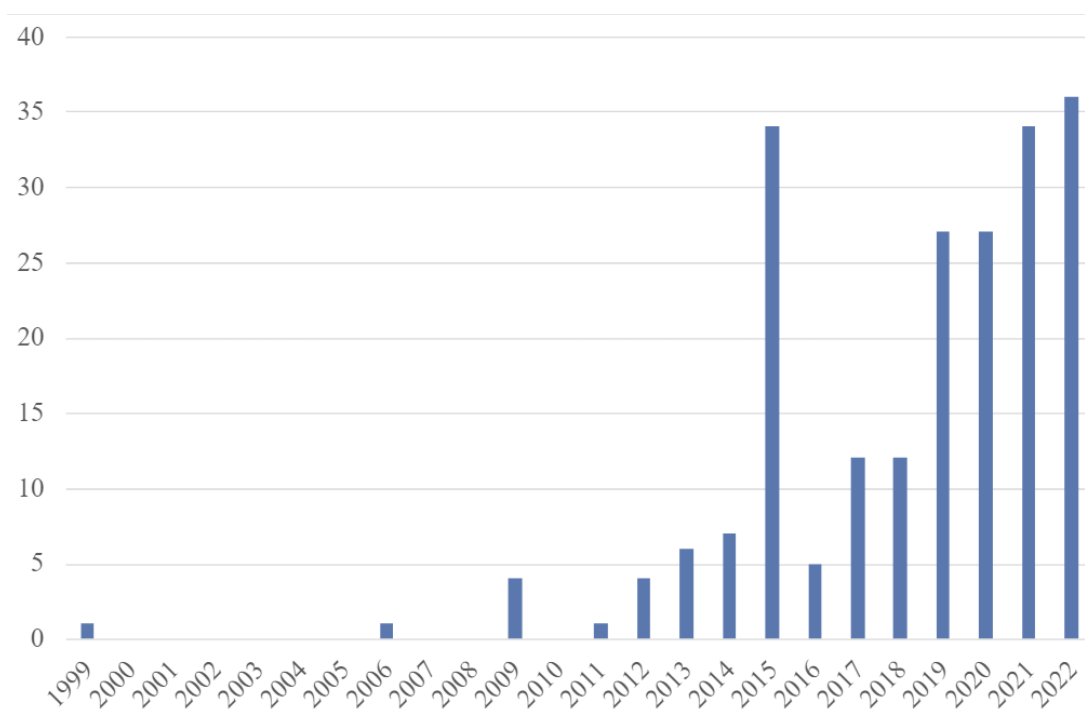
I oppsummeringen av resultater og funn, var det første som ble gjort å gi numeriske analyser av litteratursøket og seleksjon av relevante artikler. Det ble laget diagrammer for å forklare og synliggjøre prosessen. Videre ble andre relevante funn som kunne bidra til å svare på forskningsspørsmålet presentert. Dette var funn som omhandlet pasientpopulasjonen, hensikten med studiene og hva slags programvare som var benyttet for analyse av journaler.

3 Resultater

Litteratursøket resulterte i 211 unike artikler. Disse artiklene ble så gjennomgått for å vurdere hvilke som var relevante for den systematiske kartleggingsoversikten, etter beskrivelsen i metodedelene. Se *Figur 1* for oversikt over seleksjonsprosessen. Kun én artikkel, Menger et al. (18), var skrevet på annet språk enn engelsk og ble ekskludert som følge av dette. Av de 210 artiklene på engelsk språk ble 186 kategorisert som empiriske studier. Studier som ble ekskludert var systematiske oversikter, oversiktsartikler, kommentarer og kohorteprofiler. I 182 av de 186 empiriske artiklene var det brukt journaldata i fritekst i gjennomføring av studiene. I de ekskluderte studiene var det brukt strukturert data, lydopptak eller det var ikke brukt journaldata i det hele tatt. Videre omhandlet 86 av de 182 fritekststudiene psykiatriske pasienter. Av de 86 studiene gjort på journaler til psykiatriske pasienter var det 24 som omhandlet pasienter med affektive lidelser og 25 som omhandlet pasienter med psykoselidelser. Av disse omhandlet 12 artikler både pasienter med affektive lidelser og psykoselidelser, det var altså totalt 37 unike studier på journaldata til pasienter med psykoselidelser og/eller affektive lidelser. Det var disse 37 studiene som til slutt ble de utvalgt for nærmere undersøkelser, og som de videre resultatene og diskusjon tar utgangspunkt i. En oversikt over alle utvalgte studier finnes i *Tabell 1*, *Tabell 2* og *Tabell 3*. *Tabell 1* viser studiene som inkluderte pasienter med både affektive lidelser og psykoselidelser, *Tabell 2* viser studiene som var gjort på pasienter med psykoselidelser, og *Tabell 3* som viser studiene som hadde fokusert på pasienter med affektive lidelser.



Figur 1 - Figuren viser en oversikt over prosessen som ledet fram til de «utvalgte» artiklene. Litteratursøket resulterte i 211 treff som så ble gjennomgått etter inklusjonskriterier beskrevet i metod delen. Totalt 37 artikler oppfylte kriteriene som var satt, og disse ble undersøkt videre for å svare på forskningsspørsmålet.



Figur 2 – Figuren viser antall artikler per utgivelsesår for de 211 artiklene fra litteratursøket i PubMed. Det var en klar økning av utgivelser fra 2011 og til 2022. Året 2015 skilte seg ut med et høyt antall publiserte artikler.

3.1 Lokalisering og tidspunkt for utgivelse av studiene

Et av målene med undersøkelsen var å se på hvor de ulike studiene var gjort. De fleste av de 37 utvalgte studiene hadde sitt opphav i Storbritannia og USA. 22 av artiklene var fra Storbritannia og 13 av artiklene var fra USA. De resterende to artiklene hadde sitt opphav i henholdsvis Taiwan (19) og Sør-Korea (20).

Av de 13 studiene som hadde populasjon med kun psykoselidelser var 11 gjort i Storbritannia, én var gjort i USA (21) og én i Sør-Korea (20). Av studiene med populasjon med kun affektive lidelser var 11 av de 12 gjort i USA, og den siste var gjort i Storbritannia (22). Av de 12 studiene som var gjort på pasientpopulasjon med psykoselidelser og affektive lidelser var åtte gjort i Storbritannia (23-30), tre gjort i USA (31-33) og én i Taiwan (19). Dette viser at det i Storbritannia har vært mer fokus på bruk av journaldata hos pasienter med psykoselidelser, mens det i USA har blitt fokusert mer på bruk av journaldata til pasienter med affektive lidelser. Dette har noe å si for overførbarheten og generaliserbarheten til studiene.

Totalt 17 av de utvalgte studiene fra litteratursøket, er gjennomført i England på et datamateriale fra datamateriale South London and Maudsley NHS Foundation Trust (SLaM). Dette er en organisering av fire sykehus basert i London, som tilbyr tjenester for psykisk

helse- og rustilstander (34). SLaM har samlet alle pasientjournaler i en digital plattform, og gjort denne tilgjengelig for en rekke forskere for analyse. Denne plattformen kalles 'The Clinical Record Interactive Search' (CRIS) (35). Metodologiske beskrivelser av innsamling og organisering av denne plattformen finnes beskrevet i Stewart et al. (36) og Fernandes et al. (37).

Det var også av interesse å si noe om utgivelsesår for de utvalgte artiklene. Av de 37 utvalgte artiklene ble 18 (49%) utgitt i perioden 2020-2022. Åtte (22%) av artiklene var gitt ut i 2017-2019 (26, 27, 33, 38-42), ni (24%) var utgitt i 2014-2016 (21, 28-30, 43-47), og kun to (5%) artikler var gitt ut før 2014 (48, 49). Dette viser en økning i antall publikasjoner på området de siste årene, noe som også stemmer med fordelingen av utgivelser per år for alle treff i litteratursøket som ble gjort i PubMed, se *Figur 2*.

For litteratursøket som ble gjort i PubMed skiller året 2015 seg ut med et høyt antall publikasjoner sammenlignet med årene før og etter dette året. Ved gjennomgang av alle treffene i søket under utvelgelsesprosessen til de relevante studiene, ble det tydelig av at den sannsynlige årsaken til dette er at det i 2014 ble lagt ut en utfordring («challenge») knyttet til et datasett fra 'Integrating Biology and the Bedside' (i2b2) prosjektet, i samarbeid med University of Texas Health Science Center at Houston (UTHealth) (50, 51). Denne utfordringen gjør at mange av artiklene som er publisert i påfølgende år har felles datamateriale og nokså like mål for studiene. De fleste artiklene som ble publisert som følge av dette var imidlertid ikke relevante for denne undersøkelsen, da pasientpopulasjonen ikke var psykiatriske pasienter.

Det var nokså stor variasjon i hvilke tidsskrift de 37 utvalgte artiklene var publisert i, og antallet unike tidsskrift for publikasjon var 26. Noen av tidsskriftene hadde publisert flere artikler, blant annet «BMJ Open», som sto bak publikasjon av syv artikler (38, 46, 52-56). Et annet tidsskrift som hadde publisert flere artikler var «Frontiers in Psychiatry» som hadde publisert tre av artiklene (19, 20, 25). Blant de mest kjente tidsskriftene var «American Journal of Psychiatry» der artikkelen til Castro et al. (45) var publisert og «The Lancet» der artikkelen til Patel et al. (21) var publisert. Eksempler på andre tidsskrift som var representert, var Biological Psychiatry som hadde publisert studiene til McCoy et al. (33) og O'Dushlaine et al. (44), og Journal of Pain som hadde publisert artikkelen til Dobscha et al. (31).

3.2 Pasientpopulasjoner og dokumenttyper som ble brukt

Et annet mål for denne undersøkelsen var å få et inntrykk av datamengdene som ble brukt i de utvalgte studiene, enten ved antall dokumenter og/eller størrelse på dokumenter som er brukt, eller ved å få et inntrykk av størrelse på pasientpopulasjonen studiene tok utgangspunkt i. Å få inntrykk av antall og størrelsen på dokumentene som var brukt var vanskelig, da dette sjeldent var spesifisert i artiklene. Størrelse på pasientpopulasjon var mer tilgjengelig i artiklene, men også dette viste seg å være utfordrende å få et konkret svar på i enkelte artikler. Antallet analyserte journaler, eller størrelse på studiepopulasjon, er lagt inn i tabell 1-3. Der det var nevnt flere ulike populasjonsstørrelser, eller det var usikkert hvilken av de nevnte populasjonene som faktisk hadde fått sine journaler analysert med NLP, ble det høyeste nevnte tallet satt inn i tabellen.

I en del tilfeller var det uklart ut ifra metodedelen i studiene hva som var antallet pasienter i det totale tilgjengelige datamaterialet, og hva som var det faktiske antallet pasienter som fikk sine journalnotater analysert med NLP. Et eksempel på dette er Yang et al. (57), hvor det totale antallet i kohorten lot til å være 149000 pasienter, men det kan tolkes til at den utvalgte pasientpopulasjonen i studien var 13391.

I de utvalgte artiklene var det stor variasjon av antall pasienter eller pasientjournaler som var inkludert i studiene. Det var Panaite et al. (58) som hadde den minste pasientpopulasjonen på bare 109 pasienter. I den andre enden av skalaen er Patel et al. (56) med den største pasientpopulasjonen på hele 562940 pasienter. Det var ikke nøyaktige tall på hvor mange av pasientene i denne populasjonen som hadde affektive lidelse og psykoselidelser, men basert på Figur 3 i Patel et al. (56) kan en estimere at summen av schizofreni, schizoaffektive lidelser, bipolare lidelser og alvorlig depresjon til ca. 250000 individer. Ti studier hadde gjort analyser på mindre enn 1000 pasientjournaler, og i tillegg til Panaite et al. (58) er eksempler på noen av disse Dai et al. (19) og Downs et al. (39). 14 av studiene hadde en studiepopulasjon på mellom 1000 og 10000 pasienter, som eksempelvis Irving et al. (52), Kadra et al. (30) og Vaci et al. (22). Antall studier med studiepopulasjon mellom 10000 og 100000 var elleve og blant disse var studiene til Patel et al. (55) og Jackson et al. (27). I tillegg til studien til Patel et al. (56) hadde én annen studie brukt journaler fra mer enn 100000 pasienter, denne var Yang et al. (57) som trolig hadde 149000 pasienter i den totale kohorten. Oversikt over studiepopulasjoner i de ulike studiene finnes i *Tabell 1*, *Tabell 2* og *Tabell 3*.

I gjennomgangen av artiklene var det også ønskelig å se på hva slags kilder eller dokumenter fra elektronisk journal som var brukt som utgangspunkt for analyse med naturlig språkprosessering. Flere studier spesifiserte ikke tydelig hvilke typer dokumenter som var brukt, kun at de hadde brukt data fra elektronisk pasientjournal, eller at alle dokumenter i elektronisk journal var brukt. Eksempler på studier som ikke spesifiserte noe annet enn at det var brukt data fra elektronisk journal er Kadra et al. (29), Magrangeas et al. (53), Yang et al. (57), Mascio et al. (59) og O'Dushlaine et al. (44). Studier som hadde presisert at alle dokumenter i elektronisk journal var benyttet, men ikke spesifisert noe mer enn det, var blant annet Senior et al. (25), Das-Munshi et al. (28), Mason et al. (54) og Perlis et al. (48). Andre studier hadde brukt spesifikke deler eller dokumenter i journalen som utgangspunkt. Et eksempel er Rumshisky et al. (47) og Farran et al. (23) som begge hadde brukt epikriser som utgangspunkt for analysering. Et annet eksempel er Hart et al. (32) som hadde gjort analysering av inkomstjournaler, og Viani et al. (60) som blant annet hadde brukt første henvisning en pasient hadde til spesialisthelsetjenesten som utgangspunkt for analysen. Panaite et al. (58) oppgav å ha brukt psykiatriske journalnotater og Patel et al. (56) beskrev at det var brukt semistrukturerte fritekstfelt med beskrivelser av mental status og behandlingsplan som utgangspunkt for studien. Se for øvrig *Tabell 1*, *Tabell 2* og *Tabell 3* for oversikt over hvilke dokumenttyper de utvalgte studiene oppgav at var brukt. De ulike studiene har altså tatt utgangspunkt i noe ulike journaldokumenter, men felles for alle er at de inneholder journaldata i fritekst.

3.3 Hensikt med studiene og bruk av naturlig språkprosessering

Det var ønskelig å få et overblikk over hva de utvalgte studiene hadde som hensikt å undersøke, og hva som var målet med å bruke naturlig språkprosessering. Den systematiske kartleggingsoversikten skulle ikke gi en inngående oppsummering av dette, men en overordnet oversikt. Det var en del variasjon mellom studiene, men likevel var det noen fellestrekk som gikk igjen når det gjaldt hensikt og bruk av naturlig språkprosessering.

Flere av studiene hadde som mål å identifisere ulik informasjon fra elektronisk journal, og brukte naturlig språkprosessering som et verktøy for dette. Eksempler på informasjon som ble hentet ut fra de elektroniske journalene var bruk av medikamenter, symptomer på sykdom eller uttrykk som beskriver ulike tilstander. Et eksempel er Patel et al. (21) som hadde undersøkt assosiasjon mellom cannabisbruk og behandlingsresistens ved psykose, og har brukt naturlig språkprosessering for å identifisere bruk av cannabis fra journal. Et annet og

noe tilsvarende eksempel, er Kadra et al. (30) som hadde til hensikt å hente ut informasjon om antipsykotisk polyfarmasi fra journaldata, og brukte naturlig språkprosessering til dette.

Noen av studiene brukte naturlig språkprosessering mer for å gjøre kvantitative beskrivelser av pasientpopulasjonen. Hart et al. (32) var et eksempel på dette, med en studie som ønsket å undersøke prevalens og distribusjon av agitasjon og agitasjonsrelatert oppførsel. I denne studien ble naturlig språkprosessering brukt til å identifisere tegn på agitasjon i journaldokumenter. Jackson et al. (27) hadde som mål å kvantifisere og karakterisere transdiagnostiske fenotyper, og her ble naturlig språkprosessering brukt til å identifisere symptomdimensjoner basert på spesifikke kriterier.

Studien til Farran et al. (23), hadde som mål å bruke NLP til beregning av tromboserisiko og blødningsrisiko. Dette viser en annen måte å bruke naturlig språkprosessering på enn de studiene som kun har brukt det til å identifisere spesifikke ord for symptomer eller medikamenter. Å bruke naturlig språkprosessering med et slikt formål kan komme til nytte for mange pasienter, ikke bare dem med psykisk sykdom.

I noen tilfeller hadde NLP-verktøy blitt brukt til å selektere ut pasienter fra en større populasjon for videre undersøkelser. Dette gjaldt studier som Yang et al. (57) som har brukt naturlig språkprosessering til å hente ut journaler til pasienter som hadde fått forskrevet antipsykotika. O'Dushlaine et al. (44) hadde gjort noe tilsvarende, og brukt NLP til å velge ut pasienter med alvorlig depresjon til en av kohortene i studien.

Det var også noen av studiene som hadde til hensikt å utvikle modeller for prediksjon, risikovurdering eller diagnostikk innen psykiatri. Lee et al. (20) hadde som mål å utvikle en modell for prediksjon av tilbakefall av psykose, og brukte naturlig språkprosessering for å hente ut prediktive variabler for psykose i journal. En annen og noe lignende studie var Senior et al. (25) som undersøkte muligheten for å bruke naturlig språkprosessering til å identifisere spesifikke risikofaktorer for suicidalitet (OxMIS-variabler) i journal. Khapre et al. (24) hadde også gjort noe i denne retningen, og undersøkte om ulike symptomer på alvorlig psykisk sykdom kunne brukes for å vurdere risiko for tilbakefall i perinatalperioden hos kvinner med kjent psykisk sykdom, her ble naturlig språkprosessering brukt til å identifisere symptomer fra journal. Castro et al. (45) hadde som mål å bruke elektronisk journal til diagnostikk av bipolar lidelse, og brukte naturlig språkprosessering til å trekke ut ord knyttet til bipolaritet. Disse er altså eksempler på studier som har jobbet i retning av å utvikle beslutningsstøtteverktøy for

psykiatri, i tråd med en tanke om at naturlig språkprosessering skal kunne brukes for å hjelpe klinikere til å få mer nytte av de store informasjonsmengdene som finnes i pasientjournaler.

Antall (n)	Land	SLaM	Forskningsspørsmål/hensikt	Dokumenttyper	System/verktøy/metode	Referanse	Ref.nr.	DOI-lenke
399	UK	Ja	Å undersøke symptomdomener på alvorlig psykisk sykdom som prediktorer for risiko for tilbakefall under perinatalperioden hos kvinner med kjent alvorlig psykisk sykdom. NLP ble brukt for å identifisere symptomer fra journal.	Journalnotater (case notes)	CRIS	Khapre et al. (2021)	(24)	https://doi.org/10.1192/j.eurpsy.2021.18
500	Taiwan	Nei	Å undersøke bruk av NLP og "transfer learning" for fasilitering av pasientcreening i psykiatri. NLP ble brukt til klassifisering av psykisk sykdom basert på journaldokumenter.	Epikriser	BERT-Based Model Fine-Tuning, flere ulike typer nevralt nettverk.	Dai et al. (2021)	(19)	https://doi.org/10.3389/fpsy.2020.533949
558	UK	Ja	Å forstå årsaker til tidlig død hos etniske minoriteter med alvorlig psykisk lidelse. NLP ble brukt til å innhente informasjon om medikasjon, diagnose og sosioøkonomisk status fra journal.	Alle dokumenter i EMR	CRIS, GATE	Das-Munshi et al. (2016)	(28)	https://doi.org/10.1007/s00127-016-1185-8
900	UK	Ja	Å indentifisere "obsessive compulsive symptoms" (OCS) hos pasienter med schizofreni, schizoaftaktiv lidelse eller bipolar lidelse. NLP ble brukt til å identifisere symptomer assosiert med OCS i journal.	Journalnotater og korrespondanse	CRIS, GATE.	Chandran et al. (2019)	(26)	https://doi.org/10.1038/s41598-019-49165-2
3619	USA	Nei	Å karakterisere og kvantifisere transdiagnostiske fenotyper fra NIMH RDoC. NLP ble brukt til å identifisere symptomdimensjoner i journal.	Innkøstjournaler og epikriser	NILE	McCoy et al. (2018)	(33)	https://doi.org/10.1016/j.iopsy.2018.01.011
4548	USA	Nei	Å undersøke prevalens og distribusjon av agitasjon og agitasjonsrelatert oppførsel, samt assosiasjon til innleggelseslengde hos psykiatriske pasienter. NLP ble brukt for å identifisere agitasjon i journaldokumenter.	Innkøstjournaler	LDA	Hart et al. (2021)	(32)	https://doi.org/10.1016/j.genhosppsy.2020.11.003
4558	UK	Nei	Å undersøke muligheten for å indentifisere prediktorer for suicid for til validering av OxMIS (Oxford Mental Illness and Suicide Tool). NPL ble brukt for å identifisere risikofaktorer for suicidalitet (OxMIS-variabler) i journal.	Alle dokumenter i EMR	CRIS. Basert på Med7-modell, implementert med thinc-bibliotek (python).	Senior et al. (2020)	(25)	https://doi.org/10.3389/fpsy.2020.00268
7201	UK	Ja	Å presentere og evaluere en prosess for å hente ut data om antipsykotisk polyfarmasi fra fritekst og strukturert tekst i journal. NLP ble brukt til å hente informasjon om antipsykotisk medikasjon i journal.	Medikamentdok., journalnotater, korrespondanser	CRIS, GATE	Kadra et al. (2015)	(30)	https://doi.org/10.1186/s12888-015-0557-z
12879	UK	Ja	Å utvikle en applikasjon for identifisering av uønskede medikamentelle hendelser (ADEs) hos psykiatriske pasienter. NLP ble brukt til å identifisere tilfeller av ADEs i journal.	EMR, ikke spesifisert	CRIS, GATE	Iqbal et al. (2015)	(29)	https://doi.org/10.1371/journal.pone.0134208
16916	UK	Nei	Å evaluere forskrivning av antikoagulantia til pasienter med atrieflimmer, med og uten komorbid alvorlig psykisk sykdom. NLP ble brukt for å beregne CHA2DS2-VASc og HASBLED.	Epikriser	MedCAT	Farran et al. (2022)	(23)	https://doi.org/10.1016/j.psychires.2022.06.044
20472	UK	Ja	Å identifisere preferanser i språkbruk som reflekterer klinisk relevante symptomer på alvorlig psykisk sykdom og sammenligne dette med UK SNOMED CT. Det ble brukt NLP til å representere forholdet mellom SNOMED CT og ord brukt i journal.	Journaldokumenter, ikke spesifisert	CRIS, NLTK og Gensim, Continous Bag-of-Words	Jackson et al. (2018)	(27)	https://doi.org/10.12688/f1000research.13830.2
62721	USA	Nei	Å bestemme om psykisk sykdom er assosiert med indikatorer for kvalitet av smertebehandling. NLP ble brukt for å finne disse indikatorene i pasientjournaler.	Journalnotater om "Pain Care Quality"	Regelbasert NLP-algoritme konstruert i Python	Dobscha et al. (2022)	(31)	https://doi.org/10.1016/j.pain.2022.08.009

Tabell 1 – Tabellen viser en oversikt over de utvalgte artiklene som omhandler bruk av pasientjournaler fra pasienter med **affektive lidelser og psykoselidelser**. Tabellen er sortert etter antallet pasienter som har fått sine journaler analysert med NLP fra lavest til høyest. Tabellen viser også utgivelsesland for studiene, om studiepopulasjonen kom fra SLaM, forskningsspørsmål eller hensikt med studiene, hvilke dokumenttyper det er gjort analyser av og hvilke verktøy/system eller metode for NLP som er brukt i analysene.

Antall (n)	Land	SLaM	Forskningsspørsmål/hensikt	Dokumenttyper	System/verktøy/metode	Referanse	Ref.nr.	DOI-lenke
239	UK	Ja	Å identifisere symptomstart og varighet av ubehandlet psykose. NLP ble brukt til identifisering av tidsreferanser for psykosesymptomer i journaldokumenter.	Dokumenter fra "tidlig intervensjon" ved psykose	CRIS. SUTime - time expression extraction system	Viani et al. (2019)	(42)	https://doi.org/10.3233/shi190255
330	Sør-Korea	Nei	Å utvikle prediksjonsmodeller for tilbakefall av psykose ved å bruke journalnotater og strukturert data. NLP ble brukt til å hente ut prediktive variabler for tilbakefall av psykose i journal.	Psykologiske tester, inntakstjerner, sykepleienotater ved inntakst	LDA for å finne klynger av symptomer.	Lee et al. (2022)	(20)	https://doi.org/10.3389/fpsy.2022.844442
638	UK	Ja	Å undersøke assosiasjonen mellom negative symptomer i "early-onset" psykose og behandlingssvikt med antipsykotika. NLP ble brukt til identifisering av "Marder Factor NS" og antipsykotikabruk fra journal.	Progresjonsnotater, vurderinger av mental status, epikriser, korrespondanser	CRIS, GATE og TextHunter	Downs et al. (2019)	(39)	https://doi.org/10.1093/sc-hbul/sbx197
2026	UK	Ja	Å undersøke assosiasjon mellom cannabisbruk og behandlingsresistens ved førstegangpsykose. NLP ble brukt til identifisering av cannabisbruk i journal.	Journalnotater, ikke spesifisert	CRIS	Patel et al. (2015)	(21)	https://doi.org/10.1016/s0140-6736(15)60394-4
2157	USA	Nei	Å utvikle en metode for fenotypebeskrivelser som bruker kontekst for å finne konstellasjoner av symptomer ved psykose. NLP ble brukt til identifisering av symptomer knyttet til schizofreni og schizofrenirelaterte lidelser.	Journalnotater, ikke spesifisert	cTAKES,	Liu et al. (2019)	(40)	https://doi.org/10.1016/j.jbi.2019.103274
2579	UK	Ja	Å identifisere gjennomføring av kognitiv atferdsterapi for psykose (CBTp), og se på hvor mange pasienter med schizofreni eller psykose som hadde mottatt CBTp. NLP ble brukt for å identifisere gjennomføring av CBTp i journal.	"Case notes" og korrespondanser	CRIS, GATE	Colling et al. (2017)	(38)	https://doi.org/10.1136/bmjopen-2016-015297
3340	UK	Ja	Å undersøke kjønnsforskjeller i symptompresentasjon ved førstegangpsykose og samtidig bruk av illegale substanser. NLP ble brukt til identifisering av symptomer og bruk av illegale substanser i journal.	EMR, ikke spesifisert	CRIS NLP Applications library	Irving et al. (2021)	(52)	https://doi.org/10.1136/bmjopen-2020-042949
3433	UK	Ja	Å identifisere varighet av ubehandlet psykose. NLP ble brukt til identifisering av tidsreferanser for tidlige symptomer på psykose i journaldokumenter.	Første henvisning til spesialisthelsetjenesten, dokumenter fra "tidlig intervensjon" ved psykose	CRIS. SUTime - time expression extraction system	Viani et al. (2020)	(60)	https://doi.org/10.1186/s13326-020-00220-2
7678	UK	Ja	Å identifisere negative symptomer i journal hos pasienter med schizofreni og vurdere dette i forhold til kliniske utfall. NLP ble brukt for identifikasjon av negative symptomer i journal.	Alle dokumenter i EMR.	CRIS, GATE	Patel et al. (2015)	(46)	https://doi.org/10.1136/bmjopen-2015-007619
9323	UK	Ja	Å utvikle og evaluere NLP algoritmer for å oppdage journalføring av undergrupper av tankeinterferens og somatisk passivitet, og vurdere assosiasjon med prognose. NLP ble brukt til identifisering av symptomer på somatisk passivitet og tankeinterferens i journal.	EMR, ikke spesifisert	CRIS, TextHunter	Magrangeas et al. (2022)	(53)	https://doi.org/10.1136/bmjopen-2021-057433
20078	UK	Ja	Å undersøke om depressive symptomer predikerer at pasienter med psykose tar imot kognitiv atferdsterapi. NLP ble brukt til å hente ut informasjon om spesifikke psykiatriske symptomer og kognitiv atferdsterapi i journal.	Alle dokumenter i EMR	CRIS NLP Applications library	Mason et al. (2022)	(54)	https://doi.org/10.1136/bmjopen-2021-051873
24614	UK	Ja	Å identifisere kognitive utfordringer hos pasienter med schizofreni og se på disse i assosiasjon med kliniske utfall. NLP ble brukt til identifisering av symptomer på kognitiv dysfunksjon i journal.	EMR, ikke spesifisert	CRIS, BioBERT	Mascio et al. (2014)	(59)	https://doi.org/10.3389/fdgth.2021.711941
149000	UK	Nei	Å identifisere tilfeller av antipsykotisk polyfarmasi (APP) for å undersøke sammenhengen mellom APP og spesifikke negative bivirkninger. NLP ble brukt til å identifisere antipsykotiske legemidler i journal.	EMR, ikke spesifisert	CRIS NLP Applications library	Yang et al. (2022)	(57)	https://doi.org/10.1017/s003291722000952

Tabell 2 - Tabellen viser en oversikt over de utvalgte artiklene som omhandler bruk av pasientjournaler fra pasienter med **psykoselidelser**. Tabellen er sortert etter antallet pasienter som har fått sine journaler analysert med NLP fra lavest til høyest. Tabellen viser også utgivelsesland for studiene, om studiepopulasjonen kom fra SLaM, forskningsspørsmål eller hensikt med studiene, hvilke dokumenttyper det er gjort analyser av og hvilke verktøy/system eller metode for NLP som er brukt i analysene.

Antall (n)	Land	SLaM	Forskningsspørsmål/hensikt	Dokumenttyper	System/verktøy/metode	Referanse	Ref.nr.	DOI-lenke
109	USA	Nei	Vurdere gjennomførbarhet og reliabilitet ved identifisering av affektive tilstander fra journaldokumenter. NLP ble brukt til å sjekke tilstedeværelse av ord eller fraser knyttet til "stemningsleie" eller "affekt" i journalnotater.	Psykiatriske journalnotater	Python, brukt NER-tilnærming	Panaite et al. (2022)	(58)	https://doi.org/10.2196/344
150	USA	Nei	Å beskrive hvit substans hos pasienter med alvorlig depresjon. NLP ble brukt til å verifisere diagnose og MR-beskrivelser.	Journaldokumenter, ikke spesifisert	Ingen detaljer.	Hoogenboom et al. (2014)	(43)	https://doi.org/10.3109/156669499
324	USA	Nei	Å identifisere faktorer som er dokumentert i "medical chart progress notes" som er assosiert med suicid blant pasienter som har fått behandling for depresjon. NLP ble brukt til å identifisere symptomer og tegn på sykdom i journal.	Journalnotater, ikke spesifisert	"Chart abstraction", usikkert om dette er programvarebasert.	Kim et al. (2012)	(49)	https://doi.org/10.4088/JCP
1263	USA	Nei	Å undersøke for sjeldne "copy number variations" hos pasienter med behandlingsresistent alvorlig depressiv lidelse. NLP ble brukt til identifisering av pasienter med behandlingsresistent depresjon eller med fenotype som responderte på SSRI.	EMR, ikke spesifisert	HiTex	O'Dushlaine et al. (2014)	(44)	https://doi.org/10.1016/j.bic.2014.10.028
4687	USA	Nei	Å bruke epikriser for å predikere om pasienter blir re-innlagt. NLP ble brukt til å identifisere de mest informative ord og tema i epikriser.	Epikriser	LDA, TF-IDF, sci-kit learn.	Rumshisky et al. (2016)	(47)	https://doi.org/10.1038/tp.2016.10
5198	USA	Nei	Å bruke EMR for å vurdere utfall ved behandling av pasienter med depresjon. NLP ble brukt til å klassifisere klinisk status fra journal.	Alle dokumenter i EMR	HiTex	Perlis et al. (2012)	(48)	https://doi.org/10.1017/s0022389212000997
13000	UK	Nei	Å bruke NLP for å hente ut klinisk informasjon og vurdere hvordan dette samsvarer med vurdering fra eksperter. NLP ble brukt til å hente ut informasjon om sosial kontekst og historikk fra journal.	Journaldokumenter, ikke spesifisert	CRIS, NER, GloVe vektorer, LSTM, GATE og Prodigy (for annotering)	Vaci et al. (2020)	(22)	https://doi.org/10.1136/bmj-2020-300134
19707	UK	Ja	Å undersøke assosiasjoner mellom symptomer for mani/depresjon og kliniske utfall hos pasienter med unipolar depresjon. NLP ble brukt til uthenting av symptomer fra journal.	Journalnotater, kliniske tester, vurdering av mental status	CRIS NLP Applications library	Patel et al. (2022)	(55)	https://doi.org/10.1136/bmj-2022-056541
28184	UK	Nei	Å identifisere pasienter som har behandlingsresistent depresjon. NLP ble brukt til å innhente informasjon om ansettelsesstatus, psykiatrisk historikk og tidligere henvisninger til psykiatrisk behandling fra journal.	Journalnotater, brev, dokumenter	CRIS, Med7	Costa et al. (2022)	(61)	https://doi.org/10.1177/0269881122110628
41713	USA	Nei	Å kvantifisere effekten av å kombinere et prodrug opioid og SSRI på postoperative smerter hos deprimerede pasienter som undergikk kirurgi. NLP ble brukt til uthenting av informasjon relatert til depresjon fra journal.	Tidligere sykehistorie, somatiske progresjonsnotater, epikriser.	NLTK	Parthipan et al. (2019)	(41)	https://doi.org/10.1371/journal.pone.0215755
52235	USA	Nei	Å validere bruk av EHR for diagnostikk av bipolar lidelse. NLP ble brukt til ekstraksjon av termer knyttet til bipolar sykdom.	Diagnostiske notater, journalnotater, epikriser	HiTex	Castro et al. (2015)	(45)	https://doi.org/10.1176/appi.ps.2015.030423
562940	USA	Nei	Å utvikle NeuroBlu, en plattform for deidentifiserte EMR data for psykiatriske institusjoner i USA som bruker MindLinc journalsystem. NLP ble brukt til uthenting av info om mental status fra journalnotater.	Semistrukturerte fritekstfelt om mental status, behandlingsplan	LSTM	Patel et al. (2022)	(56)	https://doi.org/10.1136/bmj-2022-057227

Tabell 3 - Tabellen viser en oversikt over de utvalgte artiklene som omhandler bruk av pasientjournaler fra pasienter med **affektive lidelser**. Tabellen er sortert etter antallet pasienter som har fått sine journaler analysert med NLP fra lavest til høyest. Tabellen viser også utgivelsesland for studiene, om studiepopulasjonen kom fra SLaM, forskningsspørsmål eller hensikt med studiene, hvilke dokumenttyper det er gjort analyser av og hvilke verktøy/system eller metode for NLP som er brukt i analysene.

3.4 Informasjon om programvare eller metode for naturlig språkprosessering

Det siste målet med studien var å undersøke hva slags programmer eller metoder for naturlig språkprosessering som var brukt i de utvalgte studiene. Hvor mye, og hvor tilgjengelig denne informasjonen var i de ulike studiene, varierte. I noen studier var det kun nevnt at det var brukt naturlig språkprosessering, og ikke mer utfyllende informasjon, eller det var henvist til andre kilder. Andre studier nevnte program og metode for naturlig språkprosessering, og noen hadde utdypende informasjon om hvilken metode for naturlig språkprosessering som var brukt. Tabellene over utvalgte artikler (*Tabell 1*, *Tabell 2* og *Tabell 3*) viser derfor til en kombinasjon av programvare, metoder og infrastrukturelementer, alt ut ifra hvor mye informasjon som var oppgitt i artiklene.

En av studiene med få eller ingen detaljer om metode for naturlig språkprosessering var Hoogenboom et al. (43) som ikke nevnte noe om hvilket program eller metode som var brukt for naturlig språkprosessering, kun hva det var brukt til. Dobscha et al. (31) oppgir også lite informasjon om programvare og metode, kun at det var benyttet en NLP-algoritme som var konstruert i programmeringsspråket Python (<https://www.python.org/>). I andre studier var det ikke oppgitt program eller metode for naturlig språkprosessering, men referert til forfatterens tidligere erfaring med metodene, dette gjelder for eksempel McCoy et al. (33). Ved videre undersøkelse av referansen det henvises til i denne studien, var det sannsynlig at metoden som var brukt var «Narrative Information Linear Extraction (NILE)». I studien til Kim et al. (49) er det usikkert om det i det hele tatt var benyttet naturlig språkprosessering, eller om det kun var vurdert at dette kunne vært et nyttig verktøy. Studien er likevel tatt med blant de utvalgte artiklene, da naturlig språkprosessering var nevnt.

Noen studier var fortsettelse på tidligere studier som hadde blitt gjort, og de har derfor brukt teknikker som er beskrevet i andre artikler. Et eksempel på en slik studie var Patel et al. (56) som oppgir at tilnærmingen til naturlig språkprosessering som er brukt var en såkalt «deep learning, long-short-term memory (LSTM) metode», og henviser til en annen referanse for å få vite mer om metoden.

Ett system var brukt i vesentlig flere studier enn de andre. Dette var CRIS-plattformen, som er utviklet ved SLAM (beskrevet over). CRIS-plattformen har også en innebygd NLP-funksjon, og det er denne mange av studiene har benyttet seg av. NLP-funksjonen er i stor grad basert på «Generalised Architecture for Text Engineering» (GATE) og «TextHunter» (62). Denne funksjonen var brukt i totalt 21 av de 37 utvalgte studiene, og alle 21 var studier fra Storbritannia. Altså hadde 21 av de 22 britiske studiene brukt samme system for naturlig

språkprosessering. 17 av studiene var gjort på SLaM-populasjonen (se *Tabell 1*, *Tabell 2* og *Tabell 3* for hvilke artikler dette gjelder). De fire resterende studiene var gjort på andre populasjoner, men som også hadde elektroniske journaler lagt inn i CRIS-plattformen. Senior et al. (25) og Costa et al. (61) hadde hentet pasientpopulasjonen sin fra Oxford Health NHS Foundation Trust, Yang et al. (57) hadde hentet pasientpopulasjonen sin fra Camden & Islington NHS Foundation Trust og Vaci et al. (22) oppgir at pasientpopulasjonen er hentet fra 12 britiske nasjonale «Mental Health Trusts», uten å presisere videre hvilke sykehus som er inkludert. Den siste av de 22 britiske studiene, Farran et al. (23) oppgir å ha benyttet seg av NLP-verktøyet «Medical Concept Annotation Tool» (MedCAT).

Noen av studiene som hadde brukt CRIS-plattformen oppgav imidlertid mer spesifikke programmer eller metoder som var brukt. Chandran et al. (26), Das-Munshi et al. (28), Iqbal et al. (29), Kadra et al. (30), Colling et al. (38) og Patel et al. (46) oppgav å ha brukt «GATE» i sine studier. Magrangeas et al. (53) hadde i tillegg oppgitt at det var brukt «TextHunter», og Downs et al. (39) oppgav at det var brukt både «GATE» og «TextHunter» for naturlig språkprosessering i studien. Jackson et al. (27) hadde i tillegg til å oppgi CRIS også brukt «Natural Language ToolKit» (NLTK) og «Gensim», og Vaci et al. (22) oppgav «GATE» og «Prodigy» i tillegg til CRIS. Costa et al. (61) hadde brukt «Med7» sammen med CRIS. Viani et al. (60) og Viani et al. (42) hadde sammen med CRIS også brukt «SUTime – time expression extraction system».

Et program som var brukt til naturlig språkprosessering i flere av de amerikanske studiene var programmet «HiTex» som var brukt av både Perlis et al. (48), Castro et al. (45) og O'Dushlaine et al. (44). Ingen av disse studiene oppgav mer om metoden som var brukt.

Parthipan et al. (41) hadde brukt Python-modulen «NLTK», og Liu et al. (41) hadde brukt programvaren «Apache cTAKES» (Clinical Text Analysis and Knowledge Extraction System). Panaite et al. (58) hadde brukt Python, og metoden «Named entity recognition approach».

Blant artiklene som kun hadde nevnt metoder, og ikke NLP-programmer spesifikt, var Hart et al. (32) og Lee et al. (20) som begge hadde oppgitt at metoden for naturlig språkprosessering som er brukt er «Latent Dirichlet allocation» (LDA). Denne metoden er også brukt av Rumshisky et al. (47) som i tillegg hadde oppgitt at det var brukt «Term Frequency - Inverse Document Frequency» (TF-IDF) og Python-modulen «scikit-learn». Dai et al. (19) oppgav at det var brukt «Tuning» og flere ulike typer nevralt nettverk i deres studie.

Totalt sett viste det seg altså vanskelig å finne utdypende informasjon om bruk av programmer og metoder for naturlig språkprosessering i de utvalgte artiklene. Overordnet fikk en likevel et inntrykk av at forskningsprosjektene hadde brukt betydelige ressurser på å utvikle systemer for bearbeiding av tekstdata. Verktøyene brukte ofte åpne programmeringsverktøy, men det virket ikke til å være en utstrakt bruk av kommersielle internett-baserte tjenester. Journaldata ble altså tilsynelatende analysert innenfor institusjonenes infrastruktur, noe som vil kunne øke datasikkerheten.

4 Diskusjon

I denne systematiske kartleggingsoversikten en var målet å undersøke bruk av naturlig språkprosessering på elektroniske pasientjournaler til psykiatriske pasienter med alvorlige psykiatriske lidelser, som affektive lidelser og psykoselidelser. Det overordnede målet med dette var å få et inntrykk av om noe av den forskningen som var gjort på dette feltet faktisk bidrar til å forbedre pasientens situasjon ved å bedre helsetjenesten for de psykiatriske pasientene. For å kunne besvare dette ble det undersøkt flere aspekter ved de utvalgte studiene, blant annet hvor studiene var gjort, hvor store studiepopulasjoner som var brukt, hvilke dokumenter fra elektronisk journal som var analysert, hensikten med studiene og bruk av naturlig språkprosessering, og til slutt hvilke systemer eller metoder for naturlig språkprosessering som var brukt.

4.1 Oppsummering av viktigste funn

Litteratursøket resulterte i 211 artikler som ble gjennomgått for å finne relevante artikler, antall artikler som ble valgt ut til videre undersøkelse var 37. Flesteparten av studiene var gjennomført i Storbritannia og USA. Av de britiske studiene kom mange av studiene fra det samme pasientutvalget, SLaM. Størrelsen på pasientpopulasjonene varierte mye, fra minste populasjon på $n=109$ (58), til største populasjon på $n=562940$ (56). Det var få studier som hadde gjort forskning på spesifikke dokumenter i elektronisk journal, i de fleste tilfeller var det ikke presisert hvilke dokumenter som var brukt, eller så var alle dokumenter i elektronisk journal brukt. Overordnet kan det sies at hensikten for de fleste studiene og for bruk av naturlig språkprosessering kunne deles inn i noen felles kategorier: 1) identifisering av ulik informasjon fra journal, 2) kvantitative undersøkelser av populasjonen eller journalene, 3) seleksjon av pasienter til forskningskohorter og 4) vurdering av risiko. Når det gjaldt

programvare eller metode for naturlig språkprosessering, hadde dette ofte lite fokus i studiene. Det systemet for naturlig språkprosessering var brukt i flest studier var NLP-funksjonene knyttet til CRIS-plattformen.

4.2 Forståelse av funnene i forhold til formål med undersøkelsen

Formålet med litteraturundersøkelsen var å få en oversikt over hva som konkret har vært gjort i forskningsprosjekter som omhandler bruk av kunstig intelligens og naturlig språkprosessering i psykiatri, og hva slags nytte dette har for psykiatriske pasienter i helsevesenet. Overordnet kan det sies at funnene peker i retning av at verktøyene til nå ikke er i klinisk bruk, eller er direkte til nytte for enkeltpasienter. Dette vil si at fagfeltet fortsatt er på grunnforskningsstadium, og det potensielt er lenge til verktøyene det forskes på kan bli brukt i kliniske situasjoner. Disse funnene kan sies å i hovedsak være i tråd med andre systematiske undersøkelser av litteraturen, som for eksempel Le Glaz et al. (63). Dersom en ser på alle de 211 artiklene som litteratursøket resulterte i, gir disse et inntrykk av at fagfeltet i stor grad hatt fokus på utvikling av teknologiske verktøy som forenkler videre forskningsarbeid. Et godt eksempel på dette er utstrakt forskning på de-identifisering av journalldokumenter som for eksempel Lee et al. (64), Carell et al. (65) og Jian et al. (66). Dersom en ser på funnene fra de 37 utvalgte artiklene fra søket som omhandlet pasientgruppene med affektive lidelser og psykoselidelser, var det en noe større andel av studiene som er enklere å forestille seg at kan være forløpere til utvikling av kliniske verktøy. Eksempler på dette er studien som omhandlet innhenting av informasjon til beregning av risikoscore for selvmord (25) og trombose (23). Psykiatriske tilstander som affektive lidelser og psykoselidelser er tilstander med potensielt alvorlige kliniske konsekvenser, og at fagfeltet beveger seg i retning av å utvikle verktøy som kan ha klinisk betydning for disse pasientgruppene, er svært positivt.

4.3 Styrker og svakheter ved undersøkelsen

Det er flere styrker og svakheter ved den gjennomførte undersøkelsen. En viktig styrke er at litteratursøket som ble gjennomført viser en god bredde av fagfeltet og gir et overordnet inntrykk av hvordan fagfeltet ligger an, dette til tross for at resultatene ble spisset mot to spesifikke pasientgrupper. I gjennomgangen av studier som omhandlet pasienter med affektive lidelser og psykoselidelser så man også en stor variasjon i ulike forskningsformål og

metodevalg. Denne variasjonen gir også et inntrykk av hvilke ulike utfordringer fagfeltet står ovenfor.

En annen styrke ved undersøkelsen er at den ikke rettet seg mot spesielle deler av helsetjenesten, men mot alle ledd, både spesialisthelsetjenesten og primærhelsetjenesten, det eneste kravet var at det må være brukt elektronisk journalsystem. Søket kunne for eksempel være snevret inn til kun å gjelde studier som omhandlet data fra pasienter som ble behandlet av spesialister i psykiatri, med den hensikt å kartlegge bruk av naturlig språkprosessering i sammenhenger der fagkompetansen var spesielt høy. En slik tilnærming ville være lite realistisk, da det er en stor andel av pasientgrunnet som ikke er på et så høyt behandlingsnivå, men får tilstrekkelige oppfølging i primærhelsetjenesten.

I gjennomgangen av resultatene fra litteratursøket var inklusjonen av artikler som ble valgt ut til videre vurdering ganske bred. Der det var tvil om en artikkel oppfylte kriterier eller ikke, selv etter at to personer (LH og TH) hadde gått gjennom artiklene sammen, ble de aktuelle artiklene oftere selektert inn blant de relevante enn ut. Dette gjorde at de utvalgte artiklene fikk noe mer bredde enn de ville hatt dersom seleksjonen var strengere, noe som er en styrke ved undersøkelsen.

Litteratursøket ble gjennomført i én database, PubMed. Bakgrunnen for dette er begrunnet i metodedelen, men kan likevel regnes som en svakhet ved undersøkelsen. Ved å kun gjøre søket i en database kan flere relevante eller viktige artikler være utelatt. Andre databaser som kunne vært brukt er for eksempel PsycInfo, EMBASE, MEDLINE, ScienceDirect eller Scopus. Det kan også argumenteres for at i et fagfelt med såpass mye fokus på teknologi som det forskning på bruk av naturlig språkprosessering i elektronisk journal er, kunne søket vært gjort i mer teknologirettede søkemotorer som for eksempel IEEE explore. I tillegg er fagfeltet i svært rask utvikling, og det kunne derfor vært aktuelt å bruke sider som inneholder upubliserte artikler, som for eksempel arxiv.org eller medRxiv.org. Gitt at undersøkelsen har hovedfokus på studiers faktiske betydning for helsevesenet og pasienter, kan det likevel argumenteres for at å bruke PubMed var riktig for dens formål.

Det er i tillegg verdt å kommentere på søket i seg selv, og hvilke søketermer som er brukt, selv om også dette er begrunnet i metodedelen. Søket ble gjennomført med bruk av kun tre søketermer, «psychiatry», «electronic medical records» og «natural language processing». Det ble ikke brukt synonymer. Etersom målet med studien var å undersøke bruk av journaler fra pasienter med spesifikke diagnoser, affektive lidelser og psykoselidelser, kan det

argumenteres for at dette burde vært mer tydelig i valg av søketermer. Å bruke flere, eller mer detaljerte søketermer kunne ført til at litteratursøket fikk flere relevante resultater, eller at søkeresultatene ble mer spesifikke for den undersøkelsen som skulle bli gjort.

En annen svakhet ved undersøkelsen er at gjennomgangen av artiklene som litteratursøket resulterte i, i hovedsak er gjort av en person (LH). Dette kan ha hatt påvirkning på hvilke av artiklene som ble valgt ut for videre undersøkelse og ikke. Dersom alle artiklene hadde vært gjennomgått av flere enn en person, ikke bare de der det var tvil om kategorisering, kunne resultatene blitt annerledes. Dette gjelder også for gjennomgangen av de 37 utvalgte artiklene, og dersom alle disse ble gjennomgått av to personer kunne resultatene vært fremstilt annerledes enn når kun tvilstilfellene ble det.

Undersøkelsen valgte å fokusere på en relativt spesifikk gruppe psykiatriske pasienter, de med affektive lidelser og psykoselidelser. Dette er regnet for å være blant de mest alvorlige psykiske lidelsene, og dermed er fokus på disse viktig. Det kan likevel argumenteres for at ved å ekskludere forskning gjort på journaler til pasienter med andre psykiske lidelser, eller på pasienter med somatisk sykdom, kan viktige og relevante artikler ha blitt ekskludert. Av artiklene fra resultatet til litteratursøket hadde en del gjort forskning på journaldata til psykiatriske pasienter, men ikke presisert diagnosegrupper ut over dette. Disse ble ikke tatt med blant de utvalgte artiklene. Det er likevel sannsynlig at populasjonen i en del av disse artiklene innehold pasienter med affektive lidelser eller psykoselidelser, og dermed kan relevante artikler ha blitt selektert ut i prosessen med å finne de utvalgte artiklene.

4.4 Mulige implikasjoner av funn

Et overordnet mål for undersøkelsen var å vurdere nytten av den forskningen som er gjort med tanke på pasientenes helsesituasjon. Vi identifiserte ingen tydelige funn som direkte ville påvirke pasienters helsesituasjon. Med det menes at vi ikke fant dokumentasjon på NLP-baserte verktøy som var implementert i reelle pasientforløp, men heller at undersøkelsene var gjort på historiske data. Likevel kan studiene tenkes å ha sekundære effekter på fremtidige kliniske verktøy, men det å vurdere konkrete implikasjoner for pasienter og helsepersonell ut ifra en samling forskningsartikler er utfordrende, spesielt når variasjonen mellom de utvalgte artiklene var så stor. Det var i tillegg utfordrende å gjøre vurderinger av delmålene for undersøkelsen, da informasjon om metode, populasjon, dokumenter som var brukt og metode for naturlig språkprosessering, i flere av artiklene var vanskelig å finne informasjon om.

Hensikten for de ulike studiene hadde stor variasjon, og mange av studiene hadde hensikt og resultater uten åpenbar klinisk overføringsverdi, og som heller ikke var rettet mot forbedring av pasienters helsesituasjon. Et eksempel på en type studie som ikke har som hensikt å bruke naturlig språkprosessering på forbedring av enkeltpasienters helsesituasjon direkte, er studiene som brukte NLP-verktøy til å selektere ut kohorter til videre studier. Det kan argumenteres for at det er viktig med god seleksjon av kohorter for at forskningsresultatene skal bli spesifikke for en gitt gruppe, men dette er likevel forskning som pasientene fra de utvalgte populasjonene ikke selv vil ha nytte av.

Studiene som brukte NLP-verktøy for å gjøre kvantitative undersøkelser som for eksempel å beregne prevalens av ulike symptomer i en kohorte, som Hart et al. (32), er forskning som ikke er tydelig nyttig for forbedring av pasientenes helsesituasjon. Dette er studier som kan sies å være viktige for å få et inntrykk av populasjonen og si noe om denne på et gruppe-nivå, men som ikke er med på å forbedre pasientenes helsesituasjon på individbasis.

Da alle de 211 artiklene fra litteratursøket ble gjennomgått for å vurdere hvilke som var relevante for denne undersøkelsen, var det tydelig at ett av målene som gikk igjen var forskning på de-identifisering eller anonymisering av journaler. Flesteparten av disse studiene ble ikke valgt ut til videre vurdering, da de ikke oppfylte kriteriene, for eksempel fordi de ikke omhandlet psykiatriske pasienter eller pasienter med psykoselidelser eller affektive lidelser. Det er likevel relevant å si noe om disse, da de viser et godt eksempel på forskning som ikke har som hensikt å være til nytte for pasientene i første omgang. De-identifisering er først og fremst viktig dersom informasjon skal flyttes ut av et sikkert journalsystem. Dersom det lages NLP-verktøy som fungerer innad i et journalsystem, er ikke de-identifisering nødvendig. Utvikling av NLP-verktøy krever imidlertid store data, og forskning utenfor et journalsystem, noe som igjen gir et behov for de-identifisering. Studier på denne delen av fagfeltet er altså et godt eksempel på forskning er viktig for å muliggjøre annen forskning, men ikke som ikke har fokus på pasientens helse direkte.

På den andre siden hadde flere av studiene som hensikt å bruke naturlig språkprosessering til vurdering av risiko. Selv om studiene fremdeles er på grunnforskningsnivå, er dette viktige steg mot at forskningen faktisk kan komme til nytte for pasienter. Dersom det for eksempel er mulig å beregne risiko for suicid basert på journal som i studien til Senior et al. (25) kan dette være verktøy som bidrar til forbedring av helse på individnivå. Det samme kan verktøy som beregner risiko for trombose som i studien til Farran et al. (23). utfordringen for disse på sikt

er å overføre studiene fra forskning på historiske journaler og til journalene til nåværende pasienter.

4.4.1 Engelskspråklige land overrepresentert i resultater

I resultatet fra litteratursøket som ble gjort i PubMed, var det stor overvekt av artikler fra USA og Storbritannia, og bare noen få studier var gjort i andre land som for eksempel Frankrike, Spania, Kina og Danmark. Bare ett av treffene var skrevet på et annet språk enn engelsk, og blant de 37 artiklene som ble valgt ut til videre vurdering var bare to artikler gjort i andre land enn USA og Storbritannia. Dette utelukker ikke at andre land også gjør mye forskning på fagfeltet, men kan muligens si noe om at forskningen som gjøres på bruk av naturlig språkprosessering i pasientjournal kan være nokså spesifikk for et gitt land, og at publikasjoner i store internasjonale tidsskrift ikke alltid vil være relevant. Som i Norge, der norsk er hovedspråk for skriving av journal, vil andre land bruke sitt hovedspråk når journal blir skrevet. Når disse landene utvikler NLP-verktøy for analysering av journal, vil disse også være trent spesifikt for det gitte språket, og i liten grad ha overføringsverdi til andre land med andre språk. Det kan derfor tenkes at det å publisere forskningen internasjonalt ikke prioriteres. Dette kan også være med på å forklare hvorfor det var så få resultater fra andre land. Det er mulig å argumentere for at et litteratursøk som ikke fanger fagutvikling i alle verdens land risikerer å gå glipp av viktige bidrag. Likevel kan det også argumenteres for at fagfolk i psykiatri, og finansieringsinstitusjoner som gir dem bevilgninger, har oppmerksomhet rettet mot de ledende tidsskriftene og konferansene i fagfeltet, og disse bruker engelsk som språk. Overordnet bør derfor en helse-rettet søkeplattform som PubMed gi god oversikt over trender og innhold i forskningsproduksjonen som denne undersøkelsen skulle kartlegge.

4.4.2 Bruk av spesifikke dokumenter fra elektronisk journal

Et av funnene i undersøkelsen som var overraskende, var at det i de utvalgte studiene var lite fokus på bruk av spesifikke dokumenter fra elektronisk journal. Flere av studiene hadde ikke presisert hvilke typer dokumenter som var brukt, eller de hadde benyttet alle dokumenter fra pasientenes elektroniske journal. Noen få studier, som Rumshisky et al. (47), Viani et al. (60) og Hart et al. (32) hadde brukt spesifikke dokumenter som epikriser eller innkomstjournaler, men disse var unntaket heller en regelen. Det er vanskelig å se for seg hvordan det skal gå an

å utvikle gode, spesifikke, verktøy for diagnostikk eller risikovurdering dersom disse ikke er trent på én type dokumenter, men på all den informasjonen en elektronisk pasientjournal kan inneholde. Dersom en for eksempel skal vurdere aktuell suicidalfare ved innleggelse til sykehus, vil det være viktig å ha verktøy som er spesielt utviklet til å fange opp nyanser i hvordan klinisk status beskrives ved innleggelse, som i en inntakstjournal, i stedet for å bruke journalinformasjon som omhandler tidligere innleggelser. Kanskje kan en retning videre i forskningen være å tenke mer på bruk av enkelte dokumenttyper for utvikling av verktøy, før disse eventuelt overføres til andre dokumenter eller hele journalen.

4.4.3 Generaliserbarhet av de utvalgte studiene

Blant de 37 utvalgte studiene, er det et tydelig skille på hvilke land som har fokusert på forskning på de ulike pasientgruppene. Nesten alle studiene som er gjort på elektronisk journaldata fra pasienter med psykoselidelser er gjort i Storbritannia, nærmere bestemt London på pasientpopulasjonen fra South London and Maudsley NHS Foundation Trust (SLaM). At så mange studier er gjort på det same utvalget, taler imot generaliserbarhet av disse. Samtidig kan en argumentere for at dette materialet er generaliserbart, da populasjonen i SLaM er preget av stor diversitet. Londons befolkning er kjent for å være svært heterogen, med mennesker bosatt der fra hele verden. SLaM-populasjonen dekker alle disse, og dekker i tillegg både spesialisthelsetjeneste og primærhelsetjeneste, noe som vil si at pasienter fra flere grupper i samfunnet vil være inkludert i dette utvalget. Organisasjonen oppgir på nettsiden sin at de har mye fokus på like rettigheter («equality») for pasientene sine uavhengig av etnisk opphav og lignende (34), noe som også taler for generaliserbarhet.

Av studiene som er gjort på pasienter med affektive lidelser, er flesteparten gjort i USA. Dette taler også imot generaliserbarhet. Noen av studiene er gjort på svært store populasjoner, som Patel et al. (56), noe som øker sannsynligheten for at resultatene er generaliserbare.

4.4.4 Veien fra forskning til klinikk – «harmonisering»

Mange dokumenter er hybrider mellom strukturert tekst, semistrukturert tekst og fritekst. I tillegg lagres dokumenter i ulike filformater. Selv innad i Norge benytter ulike helseforetak ulike journalsystemer, og en kan se for seg at modeller for naturlig språkprosessering som ble utviklet i for eksempel HelseNord der journalsystemet DIPS benyttes, ikke automatisk vil være overførbart til Trøndelag, der Helseplattformen benyttes som journalsystem. I tillegg er

kulturen for dokument skriving ulik mellom ulike sykehus og i primærhelsetjenesten. Det er vanskelig å se for seg hvordan bruk og utvikling av kunstig intelligens og naturlig språkprosessering skal kunne harmoniseres i helsetjenesten. Kanskje kan det bare utvikles lokale modeller for naturlig språkprosessering som ikke vil være overførbare internasjonalt, og kanskje ikke heller nasjonalt.

5 Konklusjon

Denne undersøkelsen er en systematisk kartleggingsoversikt som hadde som formål å gi et overblikk over forskning som er gjort på bruk av naturlig språkprosessering i elektroniske journaler til pasienter med alvorlige psykiske lidelser, som affektive lidelser og psykoselidelser. Det ble gjennom undersøkelsen klart at hensikt og metoder for de artiklene som ble undersøkt ikke pekte i en tydelig retning hva gjaldt fokus på forbedring av helsetjenesten for enkeltpasienter. Studiene som tilsynelatende presenterte de mest velutviklede systemene for analysing elektroniske journaler fra psykiatriske pasienter, var gjort i Storbritannia, og var basert på en stor database med pasientjournaler fra South London and Maudsley NHS Foundation Trust (SLaM). Størrelsen på datamaterialene som var brukt i de kartlagte studiene varierte stort, fra data tilhørende populasjoner på noen hundre pasienter, til flere hundre tusen pasienter. De utvalgte artiklene hadde lite fokus på spesifikke metoder for naturlig språkprosessering, noe som muligens kan settes i sammenheng med at dette er en teknologi i rask utvikling. Basert på undersøkelsen sees det ikke noen åpenbare konsekvenser av studiene for helsetjenesten til psykiatriske pasienter slik vi kjenner den her i Norge i dag. For å se slike konsekvenser virker det som at teknologien er for langt unna stadiet der forskningsprosjekter kan bli praktisk implementert. Det kan derfor konkluderes med at grunnforskningen må fortsette i lengre tid før den vil bli klinisk relevant. Fremtidige metoder bør sørge for inkludering av pasientpopulasjoner med mangfold i språk og sosiokulturelle bakgrunner, for å øke generaliserbarheten av resultatene. Dette er vanskelig av mange årsaker, og det bør derfor arbeides mot utvikling av metoder som teknisk sett kan være overførbare mellom institusjoner og land. En slik harmonisering av tekniske fremgangsmåter kan være med på å løfte hele fagfeltet, og la selv mindre nasjoner være med på utviklingen.

Referanser

1. Andreassen OA, Malt UF, Malt EA, Melle I. Den psykiatriske undersøkelsen. Lærebok i psykiatri. 1. utg. Oslo: Gyldendal Norsk Forlag AS; 2020. s. 35-67.
2. Dahl AA, Aarre TF. Den psykiatriske undersøkelsen. I: Dahl AA, Aarre TF, red. Praktisk psykiatri. Bergen: Fagbokforlaget Vigmostad & Bjørke AS; 2012. s. 110-24.
3. Lov om helsepersonell m.v. (helsepersonelloven) LOV-1999-07-02-64 [hentet 16. mai 2023]. Tilgjengelig fra: <https://lovdata.no/lov/1999-07-02-64>/§39.
4. Zhang T, Schoene AM, Ji S, Ananiadou S. Natural language processing applied to mental illness detection: a narrative review. *npj Digital Medicine*. 2022;5(1):46.
5. Chandler C, Foltz PW, Elvevåg B. Using Machine Learning in Psychiatry: The Need to Establish a Framework That Nurtures Trustworthiness. *Schizophr Bull*. 2020;46(1):11-4.
6. Esteva A, Robicquet A, Ramsundar B, Kuleshov V, DePristo M, Chou K, et al. A guide to deep learning in healthcare. *Nat Med*. 2019;25(1):24-9.
7. Leaman R, Khare R, Lu Z. Challenges in clinical natural language processing for automated disorder normalization. *J Biomed Inform*. 2015;57:28-37.
8. Direktoratet for e-helse. Kodeverket ICD-10 (og ICD-11) [hentet 18. oktober 2021]. Tilgjengelig fra: <https://www.ehelse.no/kodeverk/kodeverket-icd-10-og-icd-11>.
9. The ICD-10 Classification of Mental and Behavioural Disorders [Internett]. World Health Organization (WHO). 1992. Tilgjengelig fra: <https://finnkode.ehelse.no/#icd10/0/1/1/2595808>.
10. American Psychiatric Association. What is serious mental illness? [Internettside]. [hentet 8. mai 2023]. Tilgjengelig fra: <https://smiadviser.org/about/serious-mental-illness>.
11. Charlson FJ, Ferrari AJ, Santomauro DF, Diminic S, Stockings E, Scott JG, et al. Global Epidemiology and Burden of Schizophrenia: Findings From the Global Burden of Disease Study 2016. *Schizophr Bull*. 2018;44(6):1195-203.
12. Høye A, Nesvåg R, Reichborn-Kjennerud T, Jacobsen BK. Sex differences in mortality among patients admitted with affective disorders in North Norway: a 33-year prospective register study. *Bipolar Disord*. 2016;18(3):272-81.
13. Pham MT, Rajić A, Greig JD, Sargeant JM, Papadopoulos A, McEwen SA. A scoping review of scoping reviews: advancing the approach and enhancing the consistency. *Research Synthesis Methods*. 2014;5(4):371-85.
14. Arksey H, O'Malley L. Scoping studies: towards a methodological framework. *International Journal of Social Research Methodology*. 2005;8(1):19-32.
15. Munn Z, Peters MDJ, Stern C, Tufanaru C, McArthur A, Aromataris E. Systematic review or scoping review? Guidance for authors when choosing between a systematic or scoping review approach. *BMC Med Res Methodol*. 2018;18(1):143.
16. Tricco AC, Lillie E, Zarin W, O'Brien KK, Colquhoun H, Levac D, et al. PRISMA Extension for Scoping Reviews (PRISMA-ScR): Checklist and Explanation. *Ann Intern Med*. 2018;169(7):467-73.
17. National Library of Medicine. PubMed Overview [Internettside]. [hentet 18. april 2023]. Tilgjengelig fra: <https://pubmed-ncbi-nlm-nih-gov.mime.uit.no/about/>.
18. Menger VJ, Spruit MR, Scheepers FE. [Knowledge discovery in clinical psychiatry: learning from electronic health records]. *Tijdschr Psychiatr*. 2021;63(4):294-300.
19. Dai HJ, Su CH, Lee YQ, Zhang YC, Wang CK, Kuo CJ, et al. Deep Learning-Based Natural Language Processing for Screening Psychiatric Patients. *Front Psychiatry*. 2020;11:533949.
20. Lee DY, Kim C, Lee S, Son SJ, Cho SM, Cho YH, et al. Psychosis Relapse Prediction Leveraging Electronic Health Records Data and Natural Language Processing Enrichment Methods. *Front Psychiatry*. 2022;13:844442.

21. Patel R, Wilson R, Jackson R, Ball M, Shetty H, Broadbent M, et al. Cannabis use and treatment resistance in first episode psychosis: a natural language processing study. *Lancet*. 2015;385 Suppl 1:S79.
22. Vaci N, Liu Q, Kormilitzin A, De Crescenzo F, Kurtulmus A, Harvey J, et al. Natural language processing for structuring clinical text data on depression using UK-CRIS. *Evid Based Ment Health*. 2020;23(1):21-6.
23. Farran D, Bean D, Wang T, Msosa Y, Casetta C, Dobson R, et al. Anticoagulation for atrial fibrillation in people with serious mental illness in the general hospital setting. *J Psychiatr Res*. 2022;153:167-73.
24. Khapre S, Stewart R, Taylor C. An evaluation of symptom domains in the 2 years before pregnancy as predictors of relapse in the perinatal period in women with severe mental illness. *Eur Psychiatry*. 2021;64(1):e26.
25. Senior M, Burghart M, Yu R, Kormilitzin A, Liu Q, Vaci N, et al. Identifying Predictors of Suicide in Severe Mental Illness: A Feasibility Study of a Clinical Prediction Rule (Oxford Mental Illness and Suicide Tool or OxMIS). *Front Psychiatry*. 2020;11:268.
26. Chandran D, Robbins DA, Chang CK, Shetty H, Sanyal J, Downs J, et al. Use of Natural Language Processing to identify Obsessive Compulsive Symptoms in patients with schizophrenia, schizoaffective disorder or bipolar disorder. *Sci Rep*. 2019;9(1):14146.
27. Jackson R, Patel R, Velupillai S, Gkotsis G, Hoyle D, Stewart R. Knowledge discovery for Deep Phenotyping serious mental illness from Electronic Mental Health records. *F1000Res*. 2018;7:210.
28. Das-Munshi J, Ashworth M, Gaughran F, Hull S, Morgan C, Nazroo J, et al. Ethnicity and cardiovascular health inequalities in people with severe mental illnesses: protocol for the E-CHASM study. *Soc Psychiatry Psychiatr Epidemiol*. 2016;51(4):627-38.
29. Iqbal E, Mallah R, Jackson RG, Ball M, Ibrahim ZM, Broadbent M, et al. Identification of Adverse Drug Events from Free Text Electronic Patient Records and Information in a Large Mental Health Case Register. *PLoS One*. 2015;10(8):e0134208.
30. Kadra G, Stewart R, Shetty H, Jackson RG, Greenwood MA, Roberts A, et al. Extracting antipsychotic polypharmacy data from electronic health records: developing and evaluating a novel process. *BMC Psychiatry*. 2015;15:166.
31. Dobscha SK, Luther SL, Kerns RD, Finch DK, Goulet JL, Brandt CA, et al. Mental Health Diagnoses are Not Associated With Indicators of Lower Quality Pain Care in Electronic Health Records of a National Sample of Veterans Treated in Veterans Health Administration Primary Care Settings. *J Pain*. 2023;24(2):273-81.
32. Hart KL, Pellegrini AM, Forester BP, Berretta S, Murphy SN, Perlis RH, et al. Distribution of agitation and related symptoms among hospitalized patients using a scalable natural language processing method. *Gen Hosp Psychiatry*. 2021;68:46-51.
33. McCoy TH, Jr., Yu S, Hart KL, Castro VM, Brown HE, Rosenquist JN, et al. High Throughput Phenotyping for Dimensional Psychopathology in Electronic Health Records. *Biol Psychiatry*. 2018;83(12):997-1004.
34. South London and Maudsley NHS Foundation Trust. South London and Maudsley NHS Foundation Trust [Internet]. South London and Maudsley; [hentet 13.april 2023]. Tilgjengelig fra: <https://slam.nhs.uk/>.
35. NIHR Maudsley Biomedical Research Centre. Clinical Record Interactive Search (CRIS) [Internet]. [hentet 11. april 2023]. Tilgjengelig fra: <https://www.maudsleybrc.nihr.ac.uk/facilities/clinical-record-interactive-search-cris/>.
36. Stewart R, Soremekun M, Perera G, Broadbent M, Callard F, Denis M, et al. The South London and Maudsley NHS Foundation Trust Biomedical Research Centre (SLAM BRC) case register: development and descriptive data. *BMC Psychiatry*. 2009;9(1):51.

37. Fernandes AC, Cloete D, Broadbent MTM, Hayes RD, Chang C-K, Jackson RG, et al. Development and evaluation of a de-identification procedure for a case register sourced from mental health electronic records. *BMC Medical Informatics and Decision Making*. 2013;13(1):71.
38. Colling C, Evans L, Broadbent M, Chandran D, Craig TJ, Kolliakou A, et al. Identification of the delivery of cognitive behavioural therapy for psychosis (CBTp) using a cross-sectional sample from electronic health records and open-text information in a large UK-based mental health case register. *BMJ Open*. 2017;7(7):e015297.
39. Downs J, Dean H, Lechler S, Sears N, Patel R, Shetty H, et al. Negative Symptoms in Early-Onset Psychosis and Their Association With Antipsychotic Treatment Failure. *Schizophr Bull*. 2019;45(1):69-79.
40. Liu Q, Woo M, Zou X, Champaneria A, Lau C, Mubbashar MI, et al. Symptom-based patient stratification in mental illness using clinical notes. *J Biomed Inform*. 2019;98:103274.
41. Parthipan A, Banerjee I, Humphreys K, Asch SM, Curtin C, Carroll I, et al. Predicting inadequate postoperative pain management in depressed patients: A machine learning approach. *PLoS One*. 2019;14(2):e0210575.
42. Viani N, Kam J, Yin L, Verma S, Stewart R, Patel R, et al. Annotating Temporal Relations to Determine the Onset of Psychosis Symptoms. *Stud Health Technol Inform*. 2019;264:418-22.
43. Hoogenboom WS, Perlis RH, Smoller JW, Zeng-Treitler Q, Gainer VS, Murphy SN, et al. Limbic system white matter microstructure and long-term treatment outcome in major depressive disorder: a diffusion tensor imaging study using legacy data. *World J Biol Psychiatry*. 2014;15(2):122-34.
44. O'Dushlaine C, Ripke S, Ruderfer DM, Hamilton SP, Fava M, Iosifescu DV, et al. Rare copy number variation in treatment-resistant major depressive disorder. *Biol Psychiatry*. 2014;76(7):536-41.
45. Castro VM, Minnier J, Murphy SN, Kohane I, Churchill SE, Gainer V, et al. Validation of electronic health record phenotyping of bipolar disorder cases and controls. *Am J Psychiatry*. 2015;172(4):363-72.
46. Patel R, Jayatilleke N, Broadbent M, Chang CK, Foskett N, Gorrell G, et al. Negative symptoms in schizophrenia: a study in a large clinical sample of patients using a novel automated method. *BMJ Open*. 2015;5(9):e007619.
47. Rumshisky A, Ghassemi M, Naumann T, Szolovits P, Castro VM, McCoy TH, et al. Predicting early psychiatric readmission with natural language processing of narrative discharge summaries. *Transl Psychiatry*. 2016;6(10):e921.
48. Perlis RH, Iosifescu DV, Castro VM, Murphy SN, Gainer VS, Minnier J, et al. Using electronic medical records to enable large-scale studies in psychiatry: treatment resistant depression as a model. *Psychol Med*. 2012;42(1):41-50.
49. Kim HM, Smith EG, Ganoczy D, Walters H, Stano CM, Ilgen MA, et al. Predictors of suicide in patient charts among patients with depression in the Veterans Health Administration health system: importance of prescription drug and alcohol abuse. *J Clin Psychiatry*. 2012;73(10):e1269-75.
50. Stubbs A, Kotfila C, Uzuner Ö. Automated systems for the de-identification of longitudinal clinical narratives: Overview of 2014 i2b2/UTHealth shared task Track 1. *J Biomed Inform*. 2015;58 Suppl(Suppl):S11-s9.
51. Stubbs A, Kotfila C, Xu H, Uzuner Ö. Identifying risk factors for heart disease over time: Overview of 2014 i2b2/UTHealth shared task Track 2. *J Biomed Inform*. 2015;58 Suppl(Suppl):S67-s77.

52. Irving J, Colling C, Shetty H, Pritchard M, Stewart R, Fusar-Poli P, et al. Gender differences in clinical presentation and illicit substance use during first episode psychosis: a natural language processing, electronic case register study. *BMJ Open*. 2021;11(4):e042949.
53. Magrangeas TT, Kolliakou A, Sanyal J, Patel R, Stewart R. Investigating the relationship between thought interference, somatic passivity and outcomes in patients with psychosis: a natural language processing approach using a clinical records search platform in south London. *BMJ Open*. 2022;12(8):e057433.
54. Mason A, Irving J, Pritchard M, Sanyal J, Colling C, Chandran D, et al. Association between depressive symptoms and cognitive-behavioural therapy receipt within a psychosis sample: a cross-sectional study. *BMJ Open*. 2022;12(5):e051873.
55. Patel R, Irving J, Brinn A, Taylor M, Shetty H, Pritchard M, et al. Associations of presenting symptoms and subsequent adverse clinical outcomes in people with unipolar depression: a prospective natural language processing (NLP), transdiagnostic, network analysis of electronic health record (EHR) data. *BMJ Open*. 2022;12(4):e056541.
56. Patel R, Wee SN, Ramaswamy R, Thadani S, Tandi J, Garg R, et al. NeuroBlu, an electronic health record (EHR) trusted research environment (TRE) to support mental healthcare analytics with real-world data. *BMJ Open*. 2022;12(4):e057227.
57. Yang JC, Thygesen JH, Werbeloff N, Hayes JF, Osborn DPJ. Antipsychotic polypharmacy and adverse drug reactions among adults in a London mental health service, 2008-2018. *Psychol Med*. 2022:1-8.
58. Panaite V, Devendorf AR, Finch D, Bouayad L, Luther SL, Schultz SK. The Value of Extracting Clinician-Recorded Affect for Advancing Clinical Research on Depression: Proof-of-Concept Study Applying Natural Language Processing to Electronic Health Records. *JMIR Form Res*. 2022;6(5):e34436.
59. Mascio A, Stewart R, Botelle R, Williams M, Mirza L, Patel R, et al. Cognitive Impairments in Schizophrenia: A Study in a Large Clinical Sample Using Natural Language Processing. *Front Digit Health*. 2021;3:711941.
60. Viani N, Kam J, Yin L, Bittar A, Dutta R, Patel R, et al. Temporal information extraction from mental health records to identify duration of untreated psychosis. *J Biomed Semantics*. 2020;11(1):2.
61. Costa T, Menzat B, Engelthaler T, Fell B, Franarin T, Roque G, et al. The burden associated with, and management of, difficult-to-treat depression in patients under specialist psychiatric care in the United Kingdom. *J Psychopharmacol*. 2022;36(5):545-56.
62. CRIS NLP SERVICE. Library of production-ready applications 2023 [oppdatert 1. mars 2023; hentet 4. mai 2023]. Tilgjengelig fra: <https://maudsleybrc.nihr.ac.uk/media/509890/applications-library-v22.pdf>.
63. Le Glaz A, Haralambous Y, Kim-Dufor DH, Lenca P, Billot R, Ryan TC, et al. Machine Learning and Natural Language Processing in Mental Health: Systematic Review. *J Med Internet Res*. 2021;23(5):e15708.
64. Lee YQ, Wang BH, Su CH, Chen PT, Lin WQ, Wu CS, et al. Protected Health Information Recognition of Unstructured Code-Mixed Electronic Health Records in Taiwan. *Stud Health Technol Inform*. 2022;290:627-31.
65. Carrell DS, Malin BA, Cronkite DJ, Aberdeen JS, Clark C, Li MR, et al. Resilience of clinical text de-identified with "hiding in plain sight" to hostile reidentification attacks by human readers. *J Am Med Inform Assoc*. 2020;27(9):1374-82.
66. Jian Z, Guo X, Liu S, Ma H, Zhang S, Zhang R, et al. A cascaded approach for Chinese clinical text de-identification with less annotation effort. *J Biomed Inform*. 2017;73:76-83.

Vedlegg

Vedlegg 1 – Referanser alle resultater litteratursøk PubMed

1. Abdalla M, Abdalla M, Rudzicz F, Hirst G. Using word embeddings to improve the privacy of clinical notes. *J Am Med Inform Assoc.* 2020;27(6):901-7.
2. Abu-El-Rub N, Urbain J, Kowalski G, Osinski K, Spaniol R, Liu M, et al. Natural Language Processing for Enterprise-scale De-identification of Protected Health Information in Clinical Notes. *AMIA Annu Symp Proc.* 2022;2022:92-101.
3. Afshar M, Joyce C, Dligach D, Sharma B, Kania R, Xie M, et al. Subtypes in patients with opioid misuse: A prognostic enrichment strategy using electronic health record data in hospitalized patients. *PLoS One.* 2019;14(7):e0219717.
4. Afshar M, Phillips A, Karnik N, Mueller J, To D, Gonzalez R, et al. Natural language processing and machine learning to identify alcohol misuse from the electronic health record in trauma patients: development and internal validation. *J Am Med Inform Assoc.* 2019;26(3):254-61.
5. Afshar M, Sharma B, Bhalla S, Thompson HM, Dligach D, Boley RA, et al. External validation of an opioid misuse machine learning classifier in hospitalized adult patients. *Addict Sci Clin Pract.* 2021;16(1):19.
6. Al-Harrasi AM, Iqbal E, Tsamakidis K, Lasek J, Gadelrab R, Soysal P, et al. Motor signs in Alzheimer's disease and vascular dementia: Detection through natural language processing, co-morbid features and relationship to adverse outcomes. *Exp Gerontol.* 2021;146:111223.
7. Alassia LN, Benítez S, Luna DR, Bernaldo de Quiros FG. Validating the Access to an Electronic Health Record: Classification and Content Analysis of Access Logs. *Stud Health Technol Inform.* 2015;216:3-6.
8. Ananthakrishnan AN, Lieberman D. Patient Electronic Health Records as a Means to Approach Genetic Research in Gastroenterology. *Gastroenterology.* 2015;149(5):1134-7.
9. Ayre K, Bittar A, Kam J, Verma S, Howard LM, Dutta R. Developing a Natural Language Processing tool to identify perinatal self-harm in electronic healthcare records. *PLoS One.* 2021;16(8):e0253809.
10. Barr PJ, Ryan J, Jacobson NC. Precision Assessment of COVID-19 Phenotypes Using Large-Scale Clinic Visit Audio Recordings: Harnessing the Power of Patient Voice. *J Med Internet Res.* 2021;23(2):e20545.
11. Barroilhet SA, Bieling AE, McCoy TH, Jr., Perlis RH. Association between DSM-5 and ICD-11 personality dimensional traits in a general medical cohort and readmission and mortality. *Gen Hosp Psychiatry.* 2020;64:63-7.
12. Barroilhet SA, Pellegrini AM, McCoy TH, Perlis RH. Characterizing DSM-5 and ICD-11 personality disorder features in psychiatric inpatients at scale using electronic health records. *Psychol Med.* 2020;50(13):2221-9.
13. Beckwith BA, Mahaadevan R, Balis UJ, Kuo F. Development and evaluation of an open source software tool for deidentification of pathology reports. *BMC Med Inform Decis Mak.* 2006;6:12.
14. Bejan CA, Ripperger M, Wilimitis D, Ahmed R, Kang J, Robinson K, et al. Improving ascertainment of suicidal ideation and suicide attempt with natural language processing. *Sci Rep.* 2022;12(1):15146.
15. Bhavsar V, Sanyal J, Patel R, Shetty H, Velupillai S, Stewart R, et al. The association between neighbourhood characteristics and physical victimisation in men and women with mental disorders. *BJPsych Open.* 2020;6(4):e73.
16. Bittar A, Velupillai S, Downs J, Sedgwick R, Dutta R. Reviewing a Decade of Research Into Suicide and Related Behaviour Using the South London and Maudsley NHS

- Foundation Trust Clinical Record Interactive Search (CRIS) System. *Front Psychiatry*. 2020;11:553463.
17. Bittar A, Velupillai S, Roberts A, Dutta R. Text Classification to Inform Suicide Risk Assessment in Electronic Health Records. *Stud Health Technol Inform*. 2019;264:40-4.
 18. Bittar A, Velupillai S, Roberts A, Dutta R. Using General-purpose Sentiment Lexicons for Suicide Risk Assessment in Electronic Health Records: Corpus-Based Analysis. *JMIR Med Inform*. 2021;9(4):e22397.
 19. Blease C, Torous J, Hägglund M. Does Patient Access to Clinical Notes Change Documentation? *Front Public Health*. 2020;8:577896.
 20. Botelle R, Bhavsar V, Kadra-Scalzo G, Mascio A, Williams MV, Roberts A, et al. Can natural language processing models extract and classify instances of interpersonal violence in mental healthcare electronic records: an applied evaluative study. *BMJ Open*. 2022;12(2):e052911.
 21. Bouhaddou O, Davis M, Donahue M, Mallia A, Griffin S, Teal J, et al. Automated Detection of Privacy Sensitive Conditions in C-CDAs: Security Labeling Services at the Department of Veterans Affairs. *AMIA Annu Symp Proc*. 2016;2016:332-41.
 22. Cardinal RN. Clinical records anonymisation and text extraction (CRATE): an open-source software system. *BMC Med Inform Decis Mak*. 2017;17(1):50.
 23. Carrell D, Malin B, Aberdeen J, Bayer S, Clark C, Wellner B, et al. Hiding in plain sight: use of realistic surrogates to reduce exposure of protected health information in clinical text. *J Am Med Inform Assoc*. 2013;20(2):342-8.
 24. Carrell DS, Cronkite DJ, Li MR, Nyemba S, Malin BA, Aberdeen JS, et al. The machine giveth and the machine taketh away: a parrot attack on clinical text deidentified with hiding in plain sight. *J Am Med Inform Assoc*. 2019;26(12):1536-44.
 25. Carrell DS, Cronkite DJ, Shea M, Oliver M, Luce C, Matson TE, et al. Clinical documentation of patient-reported medical cannabis use in primary care: Toward scalable extraction using natural language processing methods. *Subst Abus*. 2022;43(1):917-24.
 26. Carrell DS, Malin BA, Cronkite DJ, Aberdeen JS, Clark C, Li MR, et al. Resilience of clinical text de-identified with "hiding in plain sight" to hostile reidentification attacks by human readers. *J Am Med Inform Assoc*. 2020;27(9):1374-82.
 27. Carson NJ, Mullin B, Sanchez MJ, Lu F, Yang K, Menezes M, et al. Identification of suicidal behavior among psychiatrically hospitalized adolescents using natural language processing and machine learning of electronic health records. *PLoS One*. 2019;14(2):e0211116.
 28. Castro VM, Minnier J, Murphy SN, Kohane I, Churchill SE, Gainer V, et al. Validation of electronic health record phenotyping of bipolar disorder cases and controls. *Am J Psychiatry*. 2015;172(4):363-72.
 29. Castro VM, Rosand J, Giacino JT, McCoy TH, Perlis RH. Case-control study of neuropsychiatric symptoms in electronic health records following COVID-19 hospitalization in 2 academic health systems. *Mol Psychiatry*. 2022;27(9):3898-903.
 30. Chandran D, Robbins DA, Chang CK, Shetty H, Sanyal J, Downs J, et al. Use of Natural Language Processing to identify Obsessive Compulsive Symptoms in patients with schizophrenia, schizoaffective disorder or bipolar disorder. *Sci Rep*. 2019;9(1):14146.
 31. Chang NW, Dai HJ, Jonnagaddala J, Chen CW, Tsai RT, Hsu WL. A context-aware approach for progression tracking of medical concepts in electronic medical records. *J Biomed Inform*. 2015;58 Suppl(Suppl):S150-s7.
 32. Chaturvedi J, Mascio A, Velupillai SU, Roberts A. Development of a Lexicon for Pain. *Front Digit Health*. 2021;3:778305.

33. Chazard E, Mouret C, Ficheur G, Schaffar A, Beuscart JB, Beuscart R. Proposal and evaluation of FASDIM, a Fast And Simple De-Identification Method for unstructured free-text clinical records. *Int J Med Inform.* 2014;83(4):303-12.
34. Chen Q, Li H, Tang B, Wang X, Liu X, Liu Z, et al. An automatic system to identify heart disease risk factors in clinical texts over time. *J Biomed Inform.* 2015;58 Suppl(Suppl):S158-s63.
35. Chen T, Cullen RM, Godwin M. Hidden Markov model using Dirichlet process for de-identification. *J Biomed Inform.* 2015;58 Suppl(Suppl):S60-s6.
36. Chilman N, Song X, Roberts A, Tolani E, Stewart R, Chui Z, et al. Text mining occupations from the mental health electronic health record: a natural language processing approach using records from the Clinical Record Interactive Search (CRIS) platform in south London, UK. *BMJ Open.* 2021;11(3):e042274.
37. Clapp MA, Kim E, James KE, Perlis RH, Kaimal AJ, McCoy TH, Jr. Natural language processing of admission notes to predict severe maternal morbidity during the delivery encounter. *Am J Obstet Gynecol.* 2022;227(3):511.e1-.e8.
38. Cliffe C, Seyedsalehi A, Vardavoulia K, Bittar A, Velupillai S, Shetty H, et al. Using natural language processing to extract self-harm and suicidality data from a clinical sample of patients with eating disorders: a retrospective cohort study. *BMJ Open.* 2021;11(12):e053808.
39. Colling C, Evans L, Broadbent M, Chandran D, Craig TJ, Kolliakou A, et al. Identification of the delivery of cognitive behavioural therapy for psychosis (CBTp) using a cross-sectional sample from electronic health records and open-text information in a large UK-based mental health case register. *BMJ Open.* 2017;7(7):e015297.
40. Colling C, Khondoker M, Patel R, Fok M, Harland R, Broadbent M, et al. Predicting high-cost care in a mental health setting. *BJPsych Open.* 2020;6(1):e10.
41. Colling C, Mueller C, Perera G, Funnell N, Sauer J, Harwood D, et al. 'Real time' monitoring of antipsychotic prescribing in patients with dementia: a study using the Clinical Record Interactive Search (CRIS) platform to enhance safer prescribing. *BMJ Open Qual.* 2020;9(1).
42. Cormack J, Nath C, Milward D, Raja K, Jonnalagadda SR. Agile text mining for the 2014 i2b2/UTHealth Cardiac risk factors challenge. *J Biomed Inform.* 2015;58 Suppl(0):S120-s7.
43. Costa T, Menzat B, Engelthaler T, Fell B, Franarin T, Roque G, et al. The burden associated with, and management of, difficult-to-treat depression in patients under specialist psychiatric care in the United Kingdom. *J Psychopharmacol.* 2022;36(5):545-56.
44. Dai HJ, Su CH, Lee YQ, Zhang YC, Wang CK, Kuo CJ, et al. Deep Learning-Based Natural Language Processing for Screening Psychiatric Patients. *Front Psychiatry.* 2020;11:533949.
45. Dai HJ, Su CH, Wu CS. Adverse drug event and medication extraction in electronic health records via a cascading architecture with different sequence labeling models and word embeddings. *J Am Med Inform Assoc.* 2020;27(1):47-55.
46. Dai HJ, Wang FD, Chen CW, Su CH, Wu CS, Jonnagaddala J. Cohort selection for clinical trials using multiple instance learning. *J Biomed Inform.* 2020;107:103438.
47. Dalton-Locke C, Thygesen JH, Werbeloff N, Osborn D, Killaspy H. Using de-identified electronic health records to research mental health supported housing services: A feasibility study. *PLoS One.* 2020;15(8):e0237664.
48. Danielsen AA, Fenger MHJ, Østergaard SD, Nielbo KL, Mors O. Predicting mechanical restraint of psychiatric inpatients by applying machine learning on electronic health data. *Acta Psychiatr Scand.* 2019;140(2):147-57.

49. Das-Munshi J, Ashworth M, Gaughran F, Hull S, Morgan C, Nazroo J, et al. Ethnicity and cardiovascular health inequalities in people with severe mental illnesses: protocol for the E-CHASM study. *Soc Psychiatry Psychiatr Epidemiol.* 2016;51(4):627-38.
50. Dawkins E, Cruden-Smith L, Carter B, Amad A, Zandi MS, Lewis G, et al. Catatonia Psychopathology and Phenomenology in a Large Dataset. *Front Psychiatry.* 2022;13:886662.
51. Deferio JJ, Breiting S, Khullar D, Sheth A, Pathak J. Social determinants of health in mental health care and research: a case for greater inclusion. *J Am Med Inform Assoc.* 2019;26(8-9):895-9.
52. Dehghan A, Kovacevic A, Karystianis G, Keane JA, Nenadic G. Combining knowledge- and data-driven methods for de-identification of clinical narratives. *J Biomed Inform.* 2015;58 Suppl(Suppl):S53-s9.
53. Deleger L, Molnar K, Savova G, Xia F, Lingren T, Li Q, et al. Large-scale evaluation of automated clinical note de-identification and its impact on information extraction. *J Am Med Inform Assoc.* 2013;20(1):84-94.
54. Dobscha SK, Luther SL, Kerns RD, Finch DK, Goulet JL, Brandt CA, et al. Mental Health Diagnoses are Not Associated With Indicators of Lower Quality Pain Care in Electronic Health Records of a National Sample of Veterans Treated in Veterans Health Administration Primary Care Settings. *J Pain.* 2023;24(2):273-81.
55. Downs J, Dean H, Lechler S, Sears N, Patel R, Shetty H, et al. Negative Symptoms in Early-Onset Psychosis and Their Association With Antipsychotic Treatment Failure. *Schizophr Bull.* 2019;45(1):69-79.
56. Downs J, Velupillai S, George G, Holden R, Kikoler M, Dean H, et al. Detection of Suicidality in Adolescents with Autism Spectrum Disorders: Developing a Natural Language Processing Approach for Use in Electronic Health Records. *AMIA Annu Symp Proc.* 2017;2017:641-9.
57. Du L, Xia C, Deng Z, Lu G, Xia S, Ma J. A machine learning based approach to identify protected health information in Chinese clinical text. *Int J Med Inform.* 2018;116:24-32.
58. Edgcomb JB, Zima B. Machine Learning, Natural Language Processing, and the Electronic Health Record: Innovations in Mental Health Services Research. *Psychiatr Serv.* 2019;70(4):346-9.
59. Ehrenfeld JM, Gottlieb KG, Beach LB, Monahan SE, Fabbri D. Development of a Natural Language Processing Algorithm to Identify and Evaluate Transgender Patients in Electronic Health Record Systems. *Ethn Dis.* 2019;29(Suppl 2):441-50.
60. Enevoldsen KC, Danielsen AA, Rohde C, Jepsen OH, Nielbo KL, Østergaard SD. Monitoring of COVID-19 pandemic-related psychopathology using machine learning. *Acta Neuropsychiatr.* 2022;34(3):148-52.
61. Ermers NJ, Hagoort K, Scheepers FE. The Predictive Validity of Machine Learning Models in the Classification and Treatment of Major Depressive Disorder: State of the Art and Future Directions. *Front Psychiatry.* 2020;11:472.
62. Farran D, Bean D, Wang T, Msosa Y, Casetta C, Dobson R, et al. Anticoagulation for atrial fibrillation in people with serious mental illness in the general hospital setting. *J Psychiatr Res.* 2022;153:167-73.
63. Fernandes AC, Dutta R, Velupillai S, Sanyal J, Stewart R, Chandran D. Identifying Suicide Ideation and Suicidal Attempts in a Psychiatric Clinical Research Database using Natural Language Processing. *Sci Rep.* 2018;8(1):7426.
64. Ferrández Ó, South BR, Shen S, Friedlin FJ, Samore MH, Meystre SM. Generalizability and comparison of automatic clinical text de-identification methods and resources. *AMIA Annu Symp Proc.* 2012;2012:199-208.

65. Ferrández O, South BR, Shen S, Friedlin FJ, Samore MH, Meystre SM. BoB, a best-of-breed automated text de-identification system for VHA clinical documents. *J Am Med Inform Assoc.* 2013;20(1):77-83.
66. Filannino M, Uzuner Ö. Advancing the State of the Art in Clinical Natural Language Processing through Shared Tasks. *Yearb Med Inform.* 2018;27(1):184-92.
67. Fonferko-Shadrach B, Lacey AS, Roberts A, Akbari A, Thompson S, Ford DV, et al. Using natural language processing to extract structured epilepsy data from unstructured clinic letters: development and validation of the ExECT (extraction of epilepsy clinical text) system. *BMJ Open.* 2019;9(4):e023232.
68. Friberg JE, Qazi AH, Boyle B, Franciscus C, Vaughan-Sarrazin M, Westerman D, et al. Ankle- and Toe-Brachial Index for Peripheral Artery Disease Identification: Unlocking Clinical Data Through Novel Methods. *Circ Cardiovasc Interv.* 2022;15(3):e011092.
69. Goodday SM, Kormilitzin A, Vaci N, Liu Q, Cipriani A, Smith T, et al. Maximizing the use of social and behavioural information from secondary care mental health electronic health records. *J Biomed Inform.* 2020;107:103429.
70. Goodwin TR, Maldonado R, Harabagiu SM. Automatic recognition of symptom severity from psychiatric evaluation records. *J Biomed Inform.* 2017;75s:S71-s84.
71. Graham S, Depp C, Lee EE, Nebeker C, Tu X, Kim HC, et al. Artificial Intelligence for Mental Health and Mental Illnesses: an Overview. *Curr Psychiatry Rep.* 2019;21(11):116.
72. Graham SA, Lee EE, Jeste DV, Van Patten R, Twamley EW, Nebeker C, et al. Artificial intelligence approaches to predicting and detecting cognitive decline in older adults: A conceptual review. *Psychiatry Res.* 2020;284:112732.
73. Grouin C, Moriceau V, Zweigenbaum P. Combining glass box and black box evaluations in the identification of heart disease risk factors and their temporal relations from clinical records. *J Biomed Inform.* 2015;58 Suppl(Suppl):S133-s42.
74. Grouin C, Zweigenbaum P. Automatic de-identification of French clinical records: comparison of rule-based and machine-learning approaches. *Stud Health Technol Inform.* 2013;192:476-80.
75. Han S, Zhang RF, Shi L, Richie R, Liu H, Tseng A, et al. Classifying social determinants of health from unstructured electronic health records using deep learning-based natural language processing. *J Biomed Inform.* 2022;127:103984.
76. Hanauer D, Aberdeen J, Bayer S, Wellner B, Clark C, Zheng K, et al. Bootstrapping a de-identification system for narrative patient records: cost-performance tradeoffs. *Int J Med Inform.* 2013;82(9):821-31.
77. Hansen L, Enevoldsen KC, Bernstorff M, Nielbo KL, Danielsen AA, Østergaard SD. The PSYchiatric clinical outcome prediction (PSYCOP) cohort: leveraging the potential of electronic health records in the treatment of mental disorders. *Acta Neuropsychiatr.* 2021;33(6):323-30.
78. Hart KL, Pellegrini AM, Forester BP, Berretta S, Murphy SN, Perlis RH, et al. Distribution of agitation and related symptoms among hospitalized patients using a scalable natural language processing method. *Gen Hosp Psychiatry.* 2021;68:46-51.
79. Hart KL, Perlis RH, McCoy TH. Mapping of Transdiagnostic Neuropsychiatric Phenotypes Across Patients in Two General Hospitals. *J Acad Consult Liaison Psychiatry.* 2021;62(4):430-9.
80. He B, Guan Y, Cheng J, Cen K, Hua W. CRFs based de-identification of medical records. *J Biomed Inform.* 2015;58 Suppl(Suppl):S39-s46.
81. Henriksson A, Kvist M, Dalianis H. Prevalence Estimation of Protected Health Information in Swedish Clinical Text. *Stud Health Technol Inform.* 2017;235:216-20.

82. Holden R, Mueller J, McGowan J, Sanyal J, Kikoler M, Simonoff E, et al. Investigating Bullying as a Predictor of Suicidality in a Clinical Sample of Adolescents with Autism Spectrum Disorder. *Autism Res.* 2020;13(6):988-97.
83. Hoogenboom WS, Perlis RH, Smoller JW, Zeng-Treitler Q, Gainer VS, Murphy SN, et al. Limbic system white matter microstructure and long-term treatment outcome in major depressive disorder: a diffusion tensor imaging study using legacy data. *World J Biol Psychiatry.* 2014;15(2):122-34.
84. Iqbal E, Mallah R, Jackson RG, Ball M, Ibrahim ZM, Broadbent M, et al. Identification of Adverse Drug Events from Free Text Electronic Patient Records and Information in a Large Mental Health Case Register. *PLoS One.* 2015;10(8):e0134208.
85. Iqbal E, Mallah R, Rhodes D, Wu H, Romero A, Chang N, et al. ADEPt, a semantically-enriched pipeline for extracting adverse drug events from free-text electronic health records. *PLoS One.* 2017;12(11):e0187121.
86. Irving J, Colling C, Shetty H, Pritchard M, Stewart R, Fusar-Poli P, et al. Gender differences in clinical presentation and illicit substance use during first episode psychosis: a natural language processing, electronic case register study. *BMJ Open.* 2021;11(4):e042949.
87. Irving J, Patel R, Oliver D, Colling C, Pritchard M, Broadbent M, et al. Using Natural Language Processing on Electronic Health Records to Enhance Detection and Prediction of Psychosis Risk. *Schizophr Bull.* 2021;47(2):405-14.
88. Jackson MRG, Ball M, Patel R, Hayes RD, Dobson RJ, Stewart R. TextHunter--A User Friendly Tool for Extracting Generic Concepts from Free Text in Clinical Research. *AMIA Annu Symp Proc.* 2014;2014:729-38.
89. Jackson R, Kartoglu I, Stringer C, Gorrell G, Roberts A, Song X, et al. CogStack - experiences of deploying integrated information retrieval and extraction services in a large National Health Service Foundation Trust hospital. *BMC Med Inform Decis Mak.* 2018;18(1):47.
90. Jackson R, Patel R, Velupillai S, Gkotsis G, Hoyle D, Stewart R. Knowledge discovery for Deep Phenotyping serious mental illness from Electronic Mental Health records. *F1000Res.* 2018;7:210.
91. Jackson RG, Patel R, Jayatilleke N, Kolliakou A, Ball M, Gorrell G, et al. Natural language processing to extract symptoms of severe mental illness from clinical text: the Clinical Record Interactive Search Comprehensive Data Extraction (CRIS-CODE) project. *BMJ Open.* 2017;7(1):e012012.
92. Jayasinghe L, Bittar A, Dutta R, Stewart R. Clinician-recalled quoted speech in electronic health records and risk of suicide attempt: a case-crossover study. *BMJ Open.* 2020;10(4):e036186.
93. Jian Z, Guo X, Liu S, Ma H, Zhang S, Zhang R, et al. A cascaded approach for Chinese clinical text de-identification with less annotation effort. *J Biomed Inform.* 2017;73:76-83.
94. Jonnagaddala J, Liaw ST, Ray P, Kumar M, Chang NW, Dai HJ. Coronary artery disease risk assessment from unstructured electronic health records using text mining. *J Biomed Inform.* 2015;58 Suppl(Suppl):S203-s10.
95. Kadra G, Stewart R, Shetty H, Jackson RG, Greenwood MA, Roberts A, et al. Extracting antipsychotic polypharmacy data from electronic health records: developing and evaluating a novel process. *BMC Psychiatry.* 2015;15:166.
96. Karystianis G, Dehghan A, Kovacevic A, Keane JA, Nenadic G. Using local lexicalized rules to identify heart disease risk factors in clinical notes. *J Biomed Inform.* 2015;58 Suppl(Suppl):S183-s8.

97. Kayaalp M, Browne AC, Callaghan FM, Dodd ZA, Divita G, Ozturk S, et al. The pattern of name tokens in narrative clinical text and a comparison of five systems for redacting them. *J Am Med Inform Assoc.* 2014;21(3):423-31.
98. Khalifa A, Meystre S. Adapting existing natural language processing resources for cardiovascular risk factors identification in clinical notes. *J Biomed Inform.* 2015;58 Suppl(Suppl):S128-s32.
99. Khan NZ, Javed MA. Use of Artificial Intelligence-Based Strategies for Assessing Suicidal Behavior and Mental Illness: A Literature Review. *Cureus.* 2022;14(7):e27225.
100. Khapre S, Stewart R, Taylor C. An evaluation of symptom domains in the 2 years before pregnancy as predictors of relapse in the perinatal period in women with severe mental illness. *Eur Psychiatry.* 2021;64(1):e26.
101. Kim HM, Smith EG, Ganoczy D, Walters H, Stano CM, Ilgen MA, et al. Predictors of suicide in patient charts among patients with depression in the Veterans Health Administration health system: importance of prescription drug and alcohol abuse. *J Clin Psychiatry.* 2012;73(10):e1269-75.
102. Kim J, Neumann L, Paul P, Day ME, Aratow M, Bell DS, et al. Privacy-protecting, reliable response data discovery using COVID-19 patient observations. *J Am Med Inform Assoc.* 2021;28(8):1765-76.
103. Kim Y, Heider PM, Meystre SM. Comparative Study of Various Approaches for Ensemble-based De-identification of Electronic Health Record Narratives. *AMIA Annu Symp Proc.* 2020;2020:648-57.
104. Kormilitzin A, Vaci N, Liu Q, Nevado-Holgado A. Med7: A transferable clinical natural language processing model for electronic health records. *Artif Intell Med.* 2021;118:102086.
105. Kraljevic Z, Searle T, Shek A, Roguski L, Noor K, Bean D, et al. Multi-domain clinical natural language processing with MedCAT: The Medical Concept Annotation Toolkit. *Artif Intell Med.* 2021;117:102083.
106. Kumar V, Stubbs A, Shaw S, Uzuner Ö. Creation of a new longitudinal corpus of clinical narratives. *J Biomed Inform.* 2015;58 Suppl(Suppl):S6-s10.
107. Laparra E, Mascio A, Velupillai S, Miller T. A Review of Recent Work in Transfer Learning and Domain Adaptation for Natural Language Processing of Electronic Health Records. *Yearb Med Inform.* 2021;30(1):239-44.
108. Lau IS, Kraljevic Z, Al-Agil M, Charing S, Quarterman A, Parkes H, et al. Natural language word embeddings as a glimpse into healthcare language and associated mortality surrounding end of life. *BMJ Health Care Inform.* 2021;28(1).
109. Le DV, Montgomery J, Kirkby KC, Scanlan J. Risk prediction using natural language processing of electronic mental health records in an inpatient forensic psychiatry setting. *J Biomed Inform.* 2018;86:49-58.
110. Lee DY, Kim C, Lee S, Son SJ, Cho SM, Cho YH, et al. Psychosis Relapse Prediction Leveraging Electronic Health Records Data and Natural Language Processing Enrichment Methods. *Front Psychiatry.* 2022;13:844442.
111. Lee HJ, Zhang Y, Roberts K, Xu H. Leveraging existing corpora for de-identification of psychiatric notes using domain adaptation. *AMIA Annu Symp Proc.* 2017;2017:1070-9.
112. Lee YQ, Wang BH, Su CH, Chen PT, Lin WQ, Wu CS, et al. Protected Health Information Recognition of Unstructured Code-Mixed Electronic Health Records in Taiwan. *Stud Health Technol Inform.* 2022;290:627-31.
113. Leightley D, Pernet D, Velupillai S, Stewart RJ, Mark KM, Opie E, et al. The Development of the Military Service Identification Tool: Identifying Military Veterans in a Clinical Research Database Using Natural Language Processing and Machine Learning. *JMIR Med Inform.* 2020;8(5):e15852.

114. Lenert LA, Zhu V, Jennings L, McCauley JL, Obeid JS, Ward R, et al. Enhancing research data infrastructure to address the opioid epidemic: the Opioid Overdose Network (O2-Net). *JAMIA Open*. 2022;5(2):ooac055.
115. Levis M, Levy J, Dufort V, Gobbel GT, Watts BV, Shiner B. Leveraging unstructured electronic medical record notes to derive population-specific suicide risk models. *Psychiatry Res*. 2022;315:114703.
116. Li M, Carrell D, Aberdeen J, Hirschman L, Kirby J, Li B, et al. Optimizing annotation resources for natural language de-identification via a game theoretic framework. *J Biomed Inform*. 2016;61:97-109.
117. Li Z, Kormilitzin A, Fernandes M, Vaci N, Liu Q, Newby D, et al. Validation of UK Biobank data for mental health outcomes: A pilot study using secondary care electronic health records. *Int J Med Inform*. 2022;160:104704.
118. Liu Q, Woo M, Zou X, Champaneria A, Lau C, Mubbashar MI, et al. Symptom-based patient stratification in mental illness using clinical notes. *J Biomed Inform*. 2019;98:103274.
119. Liu Z, Chen Y, Tang B, Wang X, Chen Q, Li H, et al. Automatic de-identification of electronic medical records using token-level and character-level conditional random fields. *J Biomed Inform*. 2015;58 Suppl(Suppl):S47-s52.
120. Lopez MH, Holve E, Sarkar IN, Segal C. Building the informatics infrastructure for comparative effectiveness research (CER): a review of the literature. *Med Care*. 2012;50 Suppl:S38-48.
121. Lucini FR, Krewulak KD, Fiest KM, Bagshaw SM, Zuege DJ, Lee J, et al. Natural language processing to measure the frequency and mode of communication between healthcare professionals and family members of critically ill patients. *J Am Med Inform Assoc*. 2021;28(3):541-8.
122. Madan S, Julius Zimmer F, Balabin H, Schaaf S, Fröhlich H, Fluck J, et al. Deep Learning-based detection of psychiatric attributes from German mental health records. *Int J Med Inform*. 2022;161:104724.
123. Magrangeas TT, Kolliakou A, Sanyal J, Patel R, Stewart R. Investigating the relationship between thought interference, somatic passivity and outcomes in patients with psychosis: a natural language processing approach using a clinical records search platform in south London. *BMJ Open*. 2022;12(8):e057433.
124. Mascio A, Stewart R, Botelle R, Williams M, Mirza L, Patel R, et al. Cognitive Impairments in Schizophrenia: A Study in a Large Clinical Sample Using Natural Language Processing. *Front Digit Health*. 2021;3:711941.
125. Mason A, Irving J, Pritchard M, Sanyal J, Colling C, Chandran D, et al. Association between depressive symptoms and cognitive-behavioural therapy receipt within a psychosis sample: a cross-sectional study. *BMJ Open*. 2022;12(5):e051873.
126. Mayer J, Shen S, South BR, Meystre S, Friedlin FJ, Ray WR, et al. Inductive creation of an annotation schema and a reference standard for de-identification of VA electronic clinical notes. *AMIA Annu Symp Proc*. 2009;2009:416-20.
127. McCoy TH, Castro VM, Cagan A, Roberson AM, Kohane IS, Perlis RH. Sentiment Measured in Hospital Discharge Notes Is Associated with Readmission and Mortality Risk: An Electronic Health Record Study. *PLoS One*. 2015;10(8):e0136341.
128. McCoy TH, Jr., Castro VM, Hart KL, Pellegrini AM, Yu S, Cai T, et al. Genome-wide Association Study of Dimensional Psychopathology Using Electronic Health Records. *Biol Psychiatry*. 2018;83(12):1005-11.
129. McCoy TH, Castro VM, Rosenfield HR, Cagan A, Kohane IS, Perlis RH. A clinical perspective on the relevance of research domain criteria in electronic health records. *Am J Psychiatry*. 2015;172(4):316-20.

130. McCoy TH, Jr., Hart KL, Perlis RH. Characterizing and predicting rates of delirium across general hospital settings. *Gen Hosp Psychiatry*. 2017;46:1-6.
131. McCoy TH, Jr., Pellegrini AM, Perlis RH. Research Domain Criteria scores estimated through natural language processing are associated with risk for suicide and accidental death. *Depress Anxiety*. 2019;36(5):392-9.
132. McCoy TH, Jr., Pellegrini AM, Perlis RH. Differences among Research Domain Criteria score trajectories by Diagnostic and Statistical Manual categorical diagnosis during inpatient hospitalization. *PLoS One*. 2020;15(8):e0237698.
133. McCoy TH, Jr., Wiste AK, Doyle AE, Pellegrini AM, Perlis RH. Association between child psychiatric emergency room outcomes and dimensions of psychopathology. *Gen Hosp Psychiatry*. 2019;59:1-6.
134. McCoy TH, Jr., Yu S, Hart KL, Castro VM, Brown HE, Rosenquist JN, et al. High Throughput Phenotyping for Dimensional Psychopathology in Electronic Health Records. *Biol Psychiatry*. 2018;83(12):997-1004.
135. McDonald K, Smith T, Broadbent M, Patel R, Geddes JR, Saunders KEA. Prevalence and incidence of clinical outcomes in patients presenting to secondary mental health care with mood instability and sleep disturbance. *Eur Psychiatry*. 2020;63(1):e59.
136. McMurry AJ, Fitch B, Savova G, Kohane IS, Reis BY. Improved de-identification of physician notes through integrative modeling of both public and private medical text. *BMC Med Inform Decis Mak*. 2013;13:112.
137. Meerwijk EL, Tamang SR, Finlay AK, Ilgen MA, Reeves RM, Harris AHS. Suicide theory-guided natural language processing of clinical progress notes to improve prediction of veteran suicide risk: protocol for a mixed-method study. *BMJ Open*. 2022;12(8):e065088.
138. Menger VJ, Spruit MR, Scheepers FE. [Knowledge discovery in clinical psychiatry: learning from electronic health records]. *Tijdschr Psychiatr*. 2021;63(4):294-300.
139. Metzger MH, Tvardik N, Gicquel Q, Bouvry C, Poulet E, Potinet-Pagliaroli V. Use of emergency department electronic medical records for automated epidemiological surveillance of suicide attempts: a French pilot study. *Int J Methods Psychiatr Res*. 2017;26(2).
140. Meystre S, Shen S, Hofmann D, Gundlapalli A. Can physicians recognize their own patients in de-identified notes? *Stud Health Technol Inform*. 2014;205:778-82.
141. Mitra A, Ahsan H, Li W, Liu W, Kerns RD, Tsai J, et al. Risk Factors Associated With Nonfatal Opioid Overdose Leading to Intensive Care Unit Admission: A Cross-sectional Study. *JMIR Med Inform*. 2021;9(11):e32851.
142. Morrison FP, Li L, Lai AM, Hripcsak G. Repurposing the clinical record: can an existing natural language processing system de-identify clinical notes? *J Am Med Inform Assoc*. 2009;16(1):37-9.
143. Morrison FP, Sengupta S, Hripcsak G. Using a pipeline to improve de-identification performance. *AMIA Annu Symp Proc*. 2009;2009:447-51.
144. Naismith H, Howard R, Stewart R, Pitman A, Mueller C. Suicidal ideation in dementia: associations with neuropsychiatric symptoms and subtype diagnosis. *Int Psychogeriatr*. 2022:1-8.
145. Noor K, Roguski L, Bai X, Handy A, Klapaukh R, Folarin A, et al. Deployment of a Free-Text Analytics Platform at a UK National Health Service Research Hospital: CogStack at University College London Hospitals. *JMIR Med Inform*. 2022;10(8):e38122.
146. Noori A, Magdamo C, Liu X, Tyagi T, Li Z, Kondepudi A, et al. Development and Evaluation of a Natural Language Processing Annotation Tool to Facilitate Phenotyping of Cognitive Status in Electronic Health Records: Diagnostic Study. *J Med Internet Res*. 2022;24(8):e40384.

147. O'Dushlaine C, Ripke S, Ruderfer DM, Hamilton SP, Fava M, Iosifescu DV, et al. Rare copy number variation in treatment-resistant major depressive disorder. *Biol Psychiatry*. 2014;76(7):536-41.
148. Obeid JS, Davis M, Turner M, Meystre SM, Heider PM, O'Bryan EC, et al. An artificial intelligence approach to COVID-19 infection risk assessment in virtual visits: A case report. *J Am Med Inform Assoc*. 2020;27(8):1321-5.
149. Obeid JS, Heider PM, Weeda ER, Matuskowitz AJ, Carr CM, Gagnon K, et al. Impact of De-Identification on Clinical Text Classification Using Traditional and Deep Learning Classifiers. *Stud Health Technol Inform*. 2019;264:283-7.
150. Panaite V, Devendorf AR, Finch D, Bouayad L, Luther SL, Schultz SK. The Value of Extracting Clinician-Recorded Affect for Advancing Clinical Research on Depression: Proof-of-Concept Study Applying Natural Language Processing to Electronic Health Records. *JMIR Form Res*. 2022;6(5):e34436.
151. Park J, Kotzias D, Kuo P, Logan Iv RL, Merced K, Singh S, et al. Detecting conversation topics in primary care office visits from transcripts of patient-provider interactions. *J Am Med Inform Assoc*. 2019;26(12):1493-504.
152. Parthipan A, Banerjee I, Humphreys K, Asch SM, Curtin C, Carroll I, et al. Predicting inadequate postoperative pain management in depressed patients: A machine learning approach. *PLoS One*. 2019;14(2):e0210575.
153. Patel R, Irving J, Brinn A, Taylor M, Shetty H, Pritchard M, et al. Associations of presenting symptoms and subsequent adverse clinical outcomes in people with unipolar depression: a prospective natural language processing (NLP), transdiagnostic, network analysis of electronic health record (EHR) data. *BMJ Open*. 2022;12(4):e056541.
154. Patel R, Jayatilleke N, Broadbent M, Chang CK, Foskett N, Gorrell G, et al. Negative symptoms in schizophrenia: a study in a large clinical sample of patients using a novel automated method. *BMJ Open*. 2015;5(9):e007619.
155. Patel R, Wee SN, Ramaswamy R, Thadani S, Tandi J, Garg R, et al. NeuroBlu, an electronic health record (EHR) trusted research environment (TRE) to support mental healthcare analytics with real-world data. *BMJ Open*. 2022;12(4):e057227.
156. Patel R, Wickersham M, Cardinal RN, Fusar-Poli P, Correll CU. Natural Language Processing: Unlocking the Potential of Electronic Health Record Data to Support Transdiagnostic Psychiatric Research. *Biol Psychiatry Cogn Neurosci Neuroimaging*. 2022.
157. Patel R, Wilson R, Jackson R, Ball M, Shetty H, Broadbent M, et al. Cannabis use and treatment resistance in first episode psychosis: a natural language processing study. *Lancet*. 2015;385 Suppl 1:S79.
158. Perera G, Broadbent M, Callard F, Chang CK, Downs J, Dutta R, et al. Cohort profile of the South London and Maudsley NHS Foundation Trust Biomedical Research Centre (SLaM BRC) Case Register: current status and recent enhancement of an Electronic Mental Health Record-derived data resource. *BMJ Open*. 2016;6(3):e008721.
159. Perlis RH, Iosifescu DV, Castro VM, Murphy SN, Gainer VS, Minnier J, et al. Using electronic medical records to enable large-scale studies in psychiatry: treatment resistant depression as a model. *Psychol Med*. 2012;42(1):41-50.
160. Potash JB. Electronic medical records: fast track to big data in bipolar disorder. *Am J Psychiatry*. 2015;172(4):310-1.
161. Ramsey LB, Aldrich SL, Poweleit E, Prows CA, Martin LJ, Strawn JR. Racial Differences in Escitalopram/Citalopram-Related Weight Gain in Children and Adolescents: A Natural Language Processing-Based Electronic Medical Record Study. *J Child Adolesc Psychopharmacol*. 2019;29(2):162-3.

162. Redd A, Pickard S, Meystre S, Scehnet J, Bolton D, Heavirland J, et al. Evaluation of PHI Hunter in Natural Language Processing Research. *Perspect Health Inf Manag.* 2015;12(Winter):1f.
163. Richard M, Aimé X, Krebs MO, Charlet J. Enrich classifications in psychiatry with textual data: an ontology for psychiatry including social concepts. *Stud Health Technol Inform.* 2015;210:221-3.
164. Rijcken E, Kaymak U, Scheepers F, Mosteiro P, Zervanou K, Spruit M. Topic Modeling for Interpretable Text Classification From EHRs. *Front Big Data.* 2022;5:846930.
165. Roberts K, Shooshan SE, Rodriguez L, Abhyankar S, Kilicoglu H, Demner-Fushman D. The role of fine-grained annotations in supervised recognition of risk factors for heart disease from EHRs. *J Biomed Inform.* 2015;58 Suppl(Suppl):S111-s9.
166. Rumshisky A, Ghassemi M, Naumann T, Szolovits P, Castro VM, McCoy TH, et al. Predicting early psychiatric readmission with natural language processing of narrative discharge summaries. *Transl Psychiatry.* 2016;6(10):e921.
167. Sánchez D, Batet M, Viejo A. Utility-preserving privacy protection of textual healthcare documents. *J Biomed Inform.* 2014;52:189-98.
168. Sanchez-Roige S, Palmer AA. Electronic Health Records Are the Next Frontier for the Genetics of Substance Use Disorders. *Trends Genet.* 2019;35(5):317-8.
169. Scaiano M, Middleton G, Arbuckle L, Kolhatkar V, Peyton L, Dowling M, et al. A unified framework for evaluating the risk of re-identification of text de-identification tools. *J Biomed Inform.* 2016;63:174-83.
170. Schirle L, Jeffery A, Yaqoob A, Sanchez-Roige S, Samuels DC. Two data-driven approaches to identifying the spectrum of problematic opioid use: A pilot study within a chronic pain cohort. *Int J Med Inform.* 2021;156:104621.
171. Searle T, Ibrahim Z, Teo J, Dobson R. Estimating redundancy in clinical text. *J Biomed Inform.* 2021;124:103938.
172. Senior M, Burghart M, Yu R, Kormilitzin A, Liu Q, Vaci N, et al. Identifying Predictors of Suicide in Severe Mental Illness: A Feasibility Study of a Clinical Prediction Rule (Oxford Mental Illness and Suicide Tool or OxMIS). *Front Psychiatry.* 2020;11:268.
173. Sezgin E, Huang Y, Lin D, Ramtekkar U, Pauline L, Lin S. Documented Reasons of Cancellation and Rescheduling of Telehealth Appointments During the Pandemic. *Telemed J E Health.* 2021;27(10):1143-50.
174. Sharma B, Dligach D, Swope K, Salisbury-Afshar E, Karnik NS, Joyce C, et al. Publicly available machine learning models for identifying opioid misuse from the clinical notes of hospitalized patients. *BMC Med Inform Decis Mak.* 2020;20(1):79.
175. Shek A, Jiang Z, Teo J, Au Yeung J, Bhalla A, Richardson MP, et al. Machine learning-enabled multitrust audit of stroke comorbidities using natural language processing. *Eur J Neurol.* 2021;28(12):4090-7.
176. Shivade C, Malewadkar P, Fosler-Lussier E, Lai AM. Comparison of UMLS terminologies to identify risk of heart disease using clinical notes. *J Biomed Inform.* 2015;58 Suppl(Suppl):S103-s10.
177. Sohn S, Kocher JP, Chute CG, Savova GK. Drug side effect extraction from clinical narratives of psychiatry and psychology patients. *J Am Med Inform Assoc.* 2011;18 Suppl 1(Suppl 1):i144-9.
178. Spalding WM, Bertoia ML, Bulik CM, Seeger JD. Treatment characteristics among patients with binge-eating disorder: an electronic health records analysis. *Postgrad Med.* 2023;135(3):254-64.
179. Stewart R, Velupillai S. Applied natural language processing in mental health big data. *Neuropsychopharmacology.* 2021;46(1):252-3.

180. Stubbs A, Kotfila C, Uzuner Ö. Automated systems for the de-identification of longitudinal clinical narratives: Overview of 2014 i2b2/UTHealth shared task Track 1. *J Biomed Inform.* 2015;58 Suppl(Suppl):S11-s9.
181. Stubbs A, Kotfila C, Xu H, Uzuner Ö. Identifying risk factors for heart disease over time: Overview of 2014 i2b2/UTHealth shared task Track 2. *J Biomed Inform.* 2015;58 Suppl(Suppl):S67-s77.
182. Stubbs A, Uzuner Ö. Annotating longitudinal clinical narratives for de-identification: The 2014 i2b2/UTHealth corpus. *J Biomed Inform.* 2015;58 Suppl(Suppl):S20-s9.
183. Stubbs A, Uzuner Ö. Annotating risk factors for heart disease in clinical narratives for diabetic patients. *J Biomed Inform.* 2015;58 Suppl(Suppl):S78-s91.
184. Thompson HM, Kronk CA, Feasley K, Pachwicewicz P, Karnik NS. Implementation of Gender Identity and Assigned Sex at Birth Data Collection in Electronic Health Records: Where Are We Now? *Int J Environ Res Public Health.* 2021;18(12).
185. Thompson HM, Sharma B, Bhalla S, Boley R, McCluskey C, Dligach D, et al. Bias and fairness assessment of a natural language processing opioid misuse classifier: detection and mitigation of electronic health record data disadvantages across racial subgroups. *J Am Med Inform Assoc.* 2021;28(11):2393-403.
186. To D, Sharma B, Karnik N, Joyce C, Dligach D, Afshar M. Validation of an alcohol misuse classifier in hospitalized patients. *Alcohol.* 2020;84:49-55.
187. Torii M, Fan JW, Yang WL, Lee T, Wiley MT, Zisook DS, et al. Risk factor detection for heart disease by applying text analytics in electronic medical records. *J Biomed Inform.* 2015;58 Suppl(Suppl):S164-s70.
188. Tsui FR, Shi L, Ruiz V, Ryan ND, Biernesser C, Iyengar S, et al. Natural language processing and machine learning of electronic health records for prediction of first-time suicide attempts. *JAMIA Open.* 2021;4(1):00ab011.
189. Turner RJ, Coenen F, Roelofs F, Hagoort K, Härmä A, Grünwald PD, et al. Information extraction from free text for aiding transdiagnostic psychiatry: constructing NLP pipelines tailored to clinicians' needs. *BMC Psychiatry.* 2022;22(1):407.
190. Urbain J. Mining heart disease risk factors in clinical text with named entity recognition and distributional semantic models. *J Biomed Inform.* 2015;58 Suppl(Suppl):S143-s9.
191. Uzuner Ö, Stubbs A, Filannino M. A natural language processing challenge for clinical records: Research Domains Criteria (RDoC) for psychiatry. *J Biomed Inform.* 2017;75s:S1-s3.
192. Vaci N, Liu Q, Kormilitzin A, De Crescenzo F, Kurtulmus A, Harvey J, et al. Natural language processing for structuring clinical text data on depression using UK-CRIS. *Evid Based Ment Health.* 2020;23(1):21-6.
193. Velupillai S, Dalianis H, Hassel M, Nilsson GH. Developing a standard for de-identifying electronic patient records written in Swedish: precision, recall and F-measure in a manual and computerized annotation trial. *Int J Med Inform.* 2009;78(12):e19-26.
194. Velupillai S, Epstein S, Bittar A, Stephenson T, Dutta R, Downs J. Identifying Suicidal Adolescents from Mental Health Records Using Natural Language Processing. *Stud Health Technol Inform.* 2019;264:413-7.
195. Velupillai S, Suominen H, Liakata M, Roberts A, Shah AD, Morley K, et al. Using clinical Natural Language Processing for health outcomes research: Overview and actionable suggestions for future advances. *J Biomed Inform.* 2018;88:11-9.
196. Viani N, Kam J, Yin L, Bittar A, Dutta R, Patel R, et al. Temporal information extraction from mental health records to identify duration of untreated psychosis. *J Biomed Semantics.* 2020;11(1):2.

197. Viani N, Kam J, Yin L, Verma S, Stewart R, Patel R, et al. Annotating Temporal Relations to Determine the Onset of Psychosis Symptoms. *Stud Health Technol Inform.* 2019;264:418-22.
198. Warren JR, Posey B, Thornton T, Parang P. Can computer autoacquisition of medical information meet the needs of the future? A feasibility study in direct computation of the fine grained electronic medical record. *Proc AMIA Symp.* 1999:445-9.
199. Wu H, Hodgson K, Dyson S, Morley KI, Ibrahim ZM, Iqbal E, et al. Efficient Reuse of Natural Language Processing Models for Phenotype-Mention Identification in Free-text Electronic Medical Records: A Phenotype Embedding Approach. *JMIR Med Inform.* 2019;7(4):e14782.
200. Wu H, Toti G, Morley KI, Ibrahim ZM, Folarin A, Jackson R, et al. SemEHR: A general-purpose semantic search system to surface semantic data from clinical notes for tailored care, trial recruitment, and clinical research. *J Am Med Inform Assoc.* 2018;25(5):530-7.
201. Xie J, McPherson T, Powell A, Fong P, Hogan A, Ip W, et al. Ensuring Adolescent Patient Portal Confidentiality in the Age of the Cures Act Final Rule. *J Adolesc Health.* 2021;69(6):933-9.
202. Xu Y, Lee S, Martin E, D'Souza A G, Doktorchik CTA, Jiang J, et al. Enhancing ICD-Code-Based Case Definition for Heart Failure Using Electronic Medical Record Data. *J Card Fail.* 2020;26(7):610-7.
203. Yang H, Garibaldi JM. A hybrid model for automatic identification of risk factors for heart disease. *J Biomed Inform.* 2015;58 Suppl(Suppl):S171-s82.
204. Yang H, Garibaldi JM. Automatic detection of protected health information from clinic narratives. *J Biomed Inform.* 2015;58 Suppl(Suppl):S30-s8.
205. Yang JC, Thygesen JH, Werbeloff N, Hayes JF, Osborn DPJ. Antipsychotic polypharmacy and adverse drug reactions among adults in a London mental health service, 2008-2018. *Psychol Med.* 2022:1-8.
206. Yang X, Lyu T, Li Q, Lee CY, Bian J, Hogan WR, et al. A study of deep learning methods for de-identification of clinical notes in cross-institute settings. *BMC Med Inform Decis Mak.* 2019;19(Suppl 5):232.
207. Zhang R, Pakhomov SVS, Arsoniadis EG, Lee JT, Wang Y, Melton GB. Detecting clinically relevant new information in clinical notes across specialties and settings. *BMC Med Inform Decis Mak.* 2017;17(Suppl 2):68.
208. Zhao M, Havrilla J, Peng J, Drye M, Fecher M, Guthrie W, et al. Development of a phenotype ontology for autism spectrum disorder by natural language processing on electronic health records. *J Neurodev Disord.* 2022;14(1):32.
209. Zhong QY, Mittal LP, Nathan MD, Brown KM, Knudson González D, Cai T, et al. Use of natural language processing in electronic medical records to identify pregnant women with suicidal behavior: towards a solution to the complex classification problem. *Eur J Epidemiol.* 2019;34(2):153-62.
210. Zhu VJ, Lenert LA, Barth KS, Simpson KN, Li H, Kopschik M, et al. Automatically identifying opioid use disorder in non-cancer patients on chronic opioid therapy. *Health Informatics J.* 2022;28(2):14604582221107808.
211. Zhu VJ, Lenert LA, Bunnell BE, Obeid JS, Jefferson M, Halbert CH. Automatically identifying social isolation from clinical narratives for patients with prostate Cancer. *BMC Med Inform Decis Mak.* 2019;19(1):43.

