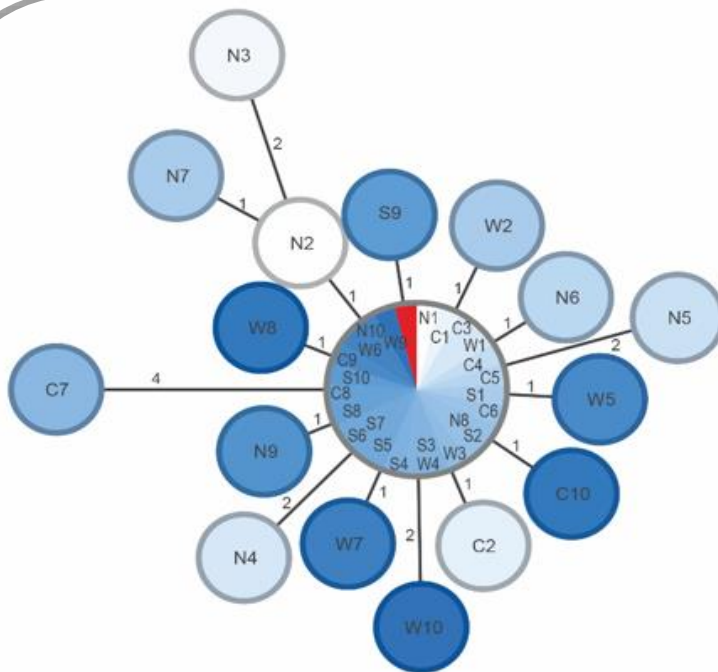


NorPREM



Nasjonalt kompetansenettverk for persontilpasset medisin
Nasjonal mikrobiologisk overvåkningsdatabase - fase 2
Fredrik Müller, Oslo universitetssykehus HF
Januar 2022 - juni 2022



Minimum spanning tree basert på genomsekvens av Pseudomonas aeruginosa ST3875 assosiert med det nasjonale utbruddet i 2021-2022. Figuren viser sammenligning av P. aeruginosa kliniske isolater fra alle helseregionene i Norge og isolat funnet i vaskekluter (rødt).¹

Illustrasjon¹

¹ <https://pubmed.ncbi.nlm.nih.gov/35514305/>



INNHold

Sammendrag	3
Innledning.....	4
1 Bakgrunn for prosjektet	5
1.1 Behov.....	5
1.2 Juridiske forhold og personvern.....	6
2 Om prosjektet.....	7
2.1 Resultatmål.....	7
2.2 Effektmål	7
2.3 Organisering av prosjektet	7
2.4 Fremdrifts- og milepælsplan	8
3 Løsningsforslag for nasjonal mikrobiologisk overvåkningsdatabase	9
3.1 Datahåndtering og lagring.....	9
3.2 Analyse	11
3.3 Visualisering.....	11
3.4 Administrasjon.....	11
3.5 Kontinuerlige overvåkninger	12
4 Oversikt og avgrensning av overvåkingsfunksjon	13
4.1 Metadata	14
5 Vedlegg	16
5.1 Vedlegg 1	16
5.2 Vedlegg 2	17



Sammendrag

Spredning av nye bakteriekloner som er tilpasset sykehusmiljøet, resistente mot våre vanlige antibiotika eller mer virulente, har betydelige helsemessige og økonomiske konsekvenser for helseforetakene, og det er viktig å oppdage dem så tidlig som mulig for å kunne iverksette effektive smitteverntiltak. Noen ganger kan det også handle om å oppdage spesifikke smittekilder, som for eksempel forurenset medisinsk utstyr eller forbruksmateriell.

Å oppdage og bidra til å oppklare smitteutbrudd er derfor en av sykehuslaboratoriernes viktige oppgaver og er en forutsetning for å kunne følge opp med smitteverntiltak lokalt. Dette gjøres ved å karakterisere bakteriene ved helgenomsekvensering, som utføres ved en rekke mikrobiologiske avdelinger i Norge. I dag eksisterer ingen nasjonal løsning for samlet lagring, sammenlikning og deling av slike sekvenser med tilhørende metadata. Dette gjør det vanskelig å håndtere infeksjonsutbrudd på tvers av helseforetak og ved andre helseinstitusjoner (for eksempel sykehjem).

Dette prosjektet har som formål å kartlegge forutsetninger for, og etablere en nasjonal infrastruktur for innsamling, lagring, sammenlikning og deling av mikrobielle helgenomdata. Helgenomdata må knyttes til relevante epidemiologiske og kliniske data (metadata) for å sikre god overvåkning og rask oppklaring av utbrudd. I tillegg til oppfølging av mistenkte utbrudd vil infrastrukturen også gi mulighet for kontinuerlig overvåkning av en rekke, fortrinnsvis resistente, bakterier i helseforetakene.

Prosjektgruppen har utarbeidet denne rapporten som kommer med en rekke anbefalinger som oppsummeres under.

Anbefalinger
1. Det må etableres nasjonal infrastruktur for å kunne etablere en felles mikrobiologisk overvåkingsdatabase
2. Det anbefales at helseforetakene innhenter nødvendighet IT – kompetanse som er en forutsetningen for god utvikling og drift av databasen
3. Det anbefales at overvåkingsdatabasen eies og driftes av RHF- ene* i fellesskap
4. Det anbefales at de mikrobiologiske miljøene fra hvert av RHF- ene inkluderes i utviklingsarbeidet av overvåkingsdatabasen
5. Det anbefales at juridiske forhold rundt personvern og databehandling avklares tidlig i etableringsfasen av overvåkingsdatabasen
6. Overvåkingsdatabasen vil være en del av helseforetakenes mikrobiologiske beredskap, og være et verktøy for å forebygge antibiotikaresistens. Dermed anbefales det at arbeidet med å etablere databasen starter i løpet av 2022 så snart som mulig

*Regionale helseforetakene



Innledning

Prosjektet om nasjonal overvåkningsdatabase for mikrobesequenser ble initiert i møte i regionalt fagnettverk for persontilpasset medisin i Helse Sør-Øst den 11. februar 2020. Det ble deretter utarbeidet et prosjektmandat som ble forankret og godkjent i det nasjonale kompetansenettverket i 2020 med Fredrik Müller, Avdeling for mikrobiologi, OUS som prosjektleder.

Den første fasen (AS – IS) av prosjektet bestod av utforming og utsendelse av et kartleggings skjema til miljøer innen infeksjonsmedisin, smittevern og mikrobiologi i Norge. Prosjektet mottok svar fra syv helseforetak:

- ✓ Oslo universitetssykehus HF
- ✓ Akershus universitetssykehus HF
- ✓ Sykehuset Innlandet HF
- ✓ Sykehuset Østfold HF
- ✓ Haukeland universitetssykehus HF
- ✓ Stavanger universitetssykehus HF
- ✓ St. Olavs hospital HF

I mars 2021 ble det gjort gjennomgang av kartleggingen som viste status med hensyn på organisering, personell, instrumentering, metoder og IKT-systemer. Miljøene støttet prosjektet og det ble opprettet et undernettverk for mikrobiologi innen NorPreM for å følge opp prosjektet videre. Det ble også avholdt et orienterende møte om våre planer med Folkehelseinstituttet 16.04.2021.



1 Bakgrunn for prosjektet

Å oppdage og bidra til å oppklare smitteutbrudd er en av sykehuslaboratoriernes viktige oppgaver. Spredning av nye bakteriekloner som er tilpasset sykehusmiljø, resistente mot våre vanlige antibiotika eller mer virulente har betydelige helsemessige og økonomiske konsekvenser, og det er viktig å oppdage dem så tidlig som mulig for å kunne iverksette effektive smitteverntiltak. Noen ganger kan det også handle om å oppdage spesifikke smittekilder, som for eksempel forurenset medisinsk utstyr eller forbruksmateriell.

Helgenomsekvensering (HGS) av bakterier åpner for rask og nøyaktig karakterisering og avklaring av smitteutbrudd. Dette åpner både for monitorering av utvalgte bakteriearter og utstrakt samarbeid mellom helseforetakene.

Bakgrunn for prosjektet «Nasjonal mikrobiologisk overvåkingsdatabase» er kartleggingsfasen som ble gjennomført for å kunne beskrive dagens situasjon.² Dette arbeidet viste med tydelighet at laboratoriene har behov for et felles system for lagring og deling av mikrobielle sekvensdata om de skal utnyttes effektivt. Det er også et behov for å kunne kople disse dataene med epidemiologiske og relevante kliniske data for å sikre god overvåkning og rask oppklaring av utbrudd.

I dag er det bare universitetssykehus og FHI som tilbyr helgenomsekvensering. HGS diagnostikk er sentralisert på mikrobiologiske eller genetiske avdelinger. Internasjonalt utfører større sykehuslaboratorier helgenomsekvensering ved mistanke om infeksjonsutbrudd. Dette er et viktig verktøy i utbruddsoppløring sammen med det lokale smittevernarbeidet. Erfaring fra tidligere utbrudd viser at overflytting av pasienter mellom sykehus kan gjøre det svært vanskelig å få oversikt over infeksjonsutbrudd og sporing av den enkelte pasient. Et aktuelt eksempel der det har vært behov for å få oversikt over sekvenser og analyser er utbruddet av *Pseudomonas aeruginosa* i flere helseregioner fra november 2021 og hvor utbruddet ble påvist ved hjelp helgenomsekvensering av *Pseudomonas aeruginosa* isolater og smittekilde senere identifisert i forurenset forbruksmateriell³.

1.1 Behov

De siste årene har flere norske helseforetak (HF) og helseinstitusjoner blitt berørt av større eller mindre infeksjonsutbrudd. Det gjelder særlig utbrudd med både antibiotikafølsomme og antibiotikaresistente bakterier som meticillin-resistente stafylokokker (MRSA), resistente gram-negative, ESBL/karbapenemase-produserende stavbakterier (ESBL; ekstendert spektrum betalaktamase) og vankomycin-resistente enterokokker (VRE).

Det er behov for å etablere en nasjonal database og infrastruktur som tillater innsamling av mikrobielle helgenomdata, kvalitetssikring av rådata, sammenlikning av sekvenser med tanke på mulig klonalitet, grafisk fremstilling av slektskap og overvåking. Samarbeid om utvikling, etablering og drift vil redusere kostnader og gi kvalitetsgevinster. I tillegg vil en nasjonal overvåkingsdatabase bidra til at man

² [PowerPoint-presentasjon \(spesialisthelsetjenesten.no\)](https://www.spesialisthelsetjenesten.no)

³ <https://pubmed.ncbi.nlm.nih.gov/35514305/>



umiddelbart vil kunne oppdage et nasjonalt eller regionalt utbrudd. Denne type infrastruktur har blitt etablert i andre land som for eksempel Sveits⁴ og Australia⁵.

Det er viktig å erkjenne at det ikke vil være egnet å planlegge dette som et tidsavhengig prosjekt der et IT-produkt skal ferdigstilles og distribueres ut. Dette handler om å etablere en felles plattform som skal inngå som en viktig del av laboratorienes virksomhet i mange år fremover. Plattformen må være fleksibel og kunne tilpasses nye behov og den rivende utviklingen innenfor feltet. Sekvenseringsteknologi og analyseverktøy er under konstant utvikling og endring. Dermed må plattformen ta høyde for forskjellige sekvenseringsteknologier. Den må ha en lagringsløsning for store datamengder med ulike eiere og mulighet for styrt utveksling av data mellom brukere, samt gode muligheter for visualisering av data og analyser. Ulike pågående overvåkinger må kunne få opprettet egne «hjemmesider» som fortløpende kan oppdateres. Det må være en avansert brukerstyring med ulike tilgangsnivåer. Siden løsningen skal benyttes innenfor diagnostikk må det være tilknyttet dedikert IT-personell og det må være support tilgjengelig på daglig basis i alle fall på hverdager slik at eventuelle tekniske problemer stort sett kan løses innen 24 timer. Av samme grunn må det vektlegges et enkelt og intuitivt brukergrensesnitt, slik at det blir enklere for laboratoriene å lære opp tilstrekkelig personell til å sikre kontinuerlig beredskap. Se tabell 5 for utdypende funksjonelle deler som bør inngå i en plattform og betraktninger av de enkelte delene.

1.2 Juridiske forhold og personvern

Forhold knyttet til personvern ved lagring av sekvenser med tilhørende metadata må gjennomgås grundig som en del av prosjektet. Under COVID – 19 - pandemien ble det en vel etablert praksis at man kan laste opp og dele sekvensdata fra mikrober, men det må avklares hvilke metadata som kan samles og hvordan disse kan deles. Gevinsten ved nasjonal overvåking basert på mikrobielle sekvenser er stor, men pseudonymiserte kliniske metadata er en forutsetning for at dette skal fungere. Det må legges opp differensiert tilgangsstyring og loggføring, og det må opprettes databehandleravtaler.

⁴ <https://www.spsp.ch/>

⁵ <https://www.cdgn.org.au/austrakka>



2 Om prosjektet

Prosjektet inngår i det nasjonale arbeidet i NorPreM. Kartleggingsfasen (AS- IS) har lagt grunnlaget for TO – BE – fasen som ble påbegynt i januar 2022. Prosjektgruppen har hatt seks arbeidsmøter hver tredje uke og har bestått av sju deltakere (vedlegg 2) fra mikrobiologiske miljøer der alle fire helseregioner har vært representert med hver sin koordinator. Prosjektgruppen har vært i kontakt med Swiss Institute of Bioinformatics i Sveits og har fått presentert og delt erfaringer fra Swiss Pathogen Surveillance Platform (SPSP) som tillater sammenlikning av helgenomsekvenser på nasjonalt nivå. Samt fått innsikt i hvilke utfordringer Sveits har hatt underveis, spesielt på det juridiske området.

Underveis i arbeidet valgte prosjektgruppen å dele seg i to arbeidsgrupper der den ene gruppen hadde hovedfokus på å definere metadata som skal importeres sammen med sekvensdata og den andre å implementere standardiserte verktøy for bioinformatisk analyse. Resultater av arbeidet er beskrevet i hvert sitt delkapittel i rapporten.

2.1 Resultatmål

- ✓ Likeverdig tilgang til mikrobielle helgenomsekvenser nasjonalt både innen klinisk rutine og forskning
- ✓ Mulighet til å påvise og følge infeksjonsutbrudd på tvers av helseforetak

2.2 Effektmål

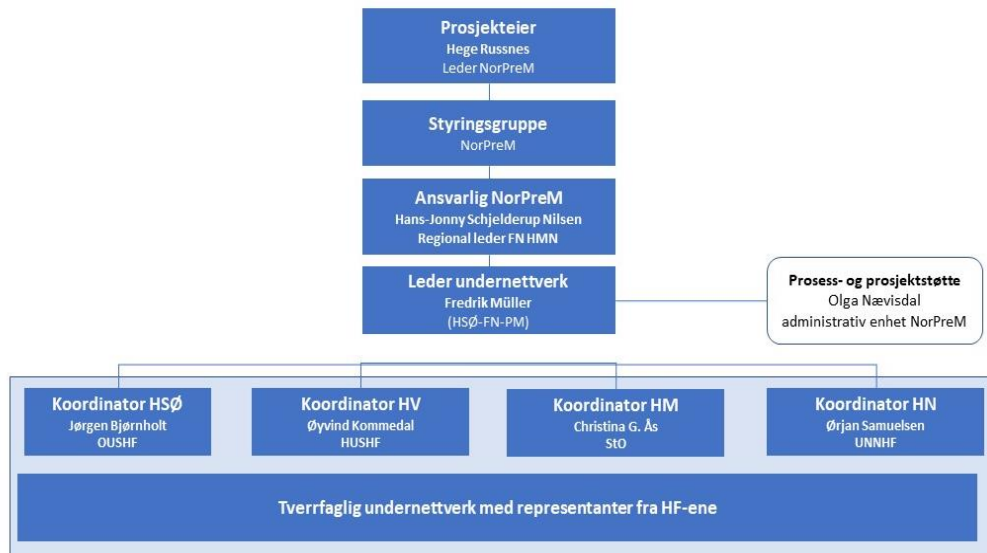
- ✓ Harmonisering av rutiner, protokoller og standarder for mikrobiell helgenomsekvensering nasjonalt
- ✓ Nasjonal plan for kompetansebehov med spesiell vekt på bioinformatikk
- ✓ Nasjonale rutiner for overvåking av infeksjonsutbrudd på sykehus
- ✓ Nasjonal mikrobiologisk overvåkningsdatabase

2.3 Organisering av prosjektet

Prosjektet inngår i NorPreM sin prosjektportefølje og er forankret hos Hans-Jonny Schjelderup Nilsen – regional fagnettverksleder i Helse Midt-Norge og eies av Hege Russnes – nasjonal leder i NorPreM. Prosjektbehovet ble initiert av Fredrik Müller – avdelingsleder på avdeling for mikrobiologi ved Oslo universitetssykehus. Han er medlem i det regionale fagnettverket for persontilpasset medisin i Helse Sør-Øst og er prosjektleder. Prosjektet får prosjektstøtte fra administrativ enhet for NorPreM.



Det har blitt oppnevnt representanter fra hver helseregion som har hatt koordinerende ansvar i hver sin region og opp mot de lokale helseforetakene.



Figur 1 illustrer organiseringen av prosjektgruppen

2.4 Fremdrifts- og milepælsplan

Tabell 1: Viser planlagt fremdrifts- og milepælsplan

Mikrobiologisk overvåkningsdatabase	2021		2022		
	Desember	Januar	Februar	Mars	April
Formål	Prosjektmandatet er forankret og prosjektgruppen er etablert	Behov for forbedringsområdet er forankret	Fremtidsløsning og anbefaling for en mikrobiologisk overvåkningsdatabase er utviklet og forankret		Anbefalinger er gitt til NorPreM
Aktivitet	Oppstartsmøte med forankring av mandat og fordeling av arbeidsoppgaver	Gjennomgang av nåsituasjon og identifisering av behov og forbedringstiltak	Utvikling av kravspesifikasjon til en mikrobiologisk overvåkningsdatabase og anbefaling for implementering		Rapportskriving og presentasjon i NorPreM
Leveranse	Prosjektmandat Fremdriftplan	Beskrivelse av nåsituasjon og behov	Rapportutkast med anbefaling		Sluttrapport
Milepæl					

Oppstart av prosjektet har blitt forsinket med en måneds tid grunnet en ny bølge med COVID - 19. Dermed ble også fremdriftsplanen forskjøvet ca. en måned. Fremdriftsplanene har ellers blir overholdt og det har vært et godt samarbeid i prosjektgruppen. Resultatet ble en rapport som gir en god beskrivelse av utfordringene i fagfeltet og behovet for nasjonal overvåkningsdatabase. Gruppen har vært enstemmig om gitte anbefalinger.



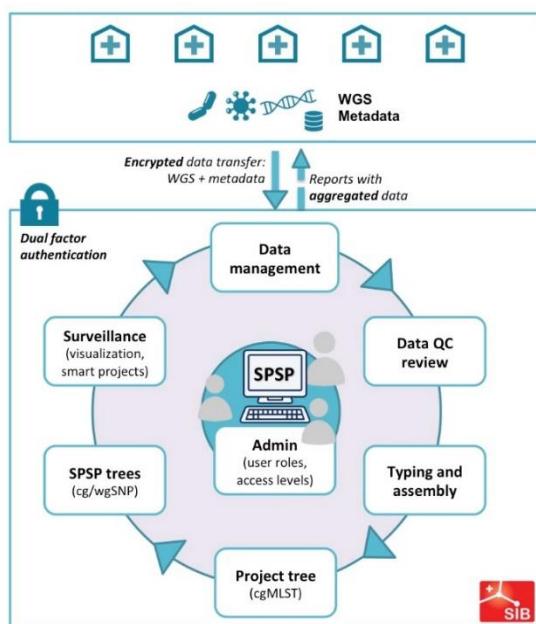
3 Løsningsforslag for nasjonal mikrobiologisk overvåkningsdatabase

En nasjonal mikrobiologisk overvåkningsdatabase bør dekke flere aspekter ved håndtering av helgenomdata. I dette kapitlet skisseres en løsning som består av fire hoveddeler: datahåndtering/lagring, analyse, visualisering og administrasjon.

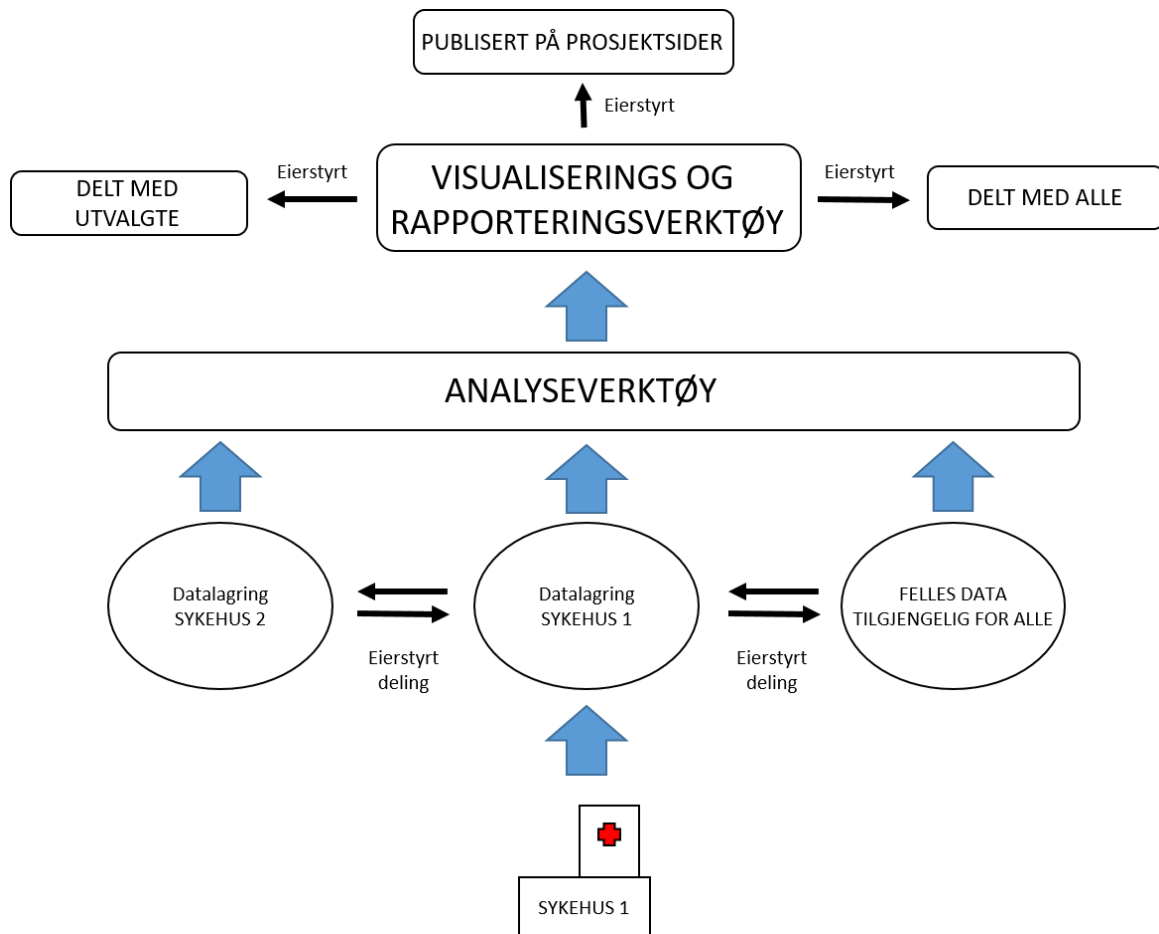
3.1 Datahåndtering og lagring

Det er behov for en portal hvor brukeren kan laste opp egne sekvensdata (rådata eller bearbejdede sekvenser) og tilhørende metadata. Tekniske metadata kan inkludere både obligatoriske (f. eks. prøvetype, sekvenseringsmetodikk, sekvenseringsdybde) og valgfrie datafelter, og bør defineres iht. en standardisert ontologi. Det er også viktig at det er mulig å endre og supplere metadata for hver enkelt bruker i ettetid. Det er videre behov for at datasett kan lastes opp gjennom et enkelt grafisk brukergrensesnitt (GUI), som gir en lavere terskel for bruk for personell uten høy bioinformatisk kompetanse. Samtidig bør det også tilrettelegges for automatisert opplasting av data via et programmeringsgrensesnitt (API), som kan være mer effektivt for laboratorier som produserer store mengder sekvensdata.

Sekvensdata, særlig rådata, krever stor lagringsplass. Ett enkelt bakterielt genom sekvensert med illumina-teknologi kan f.eks. produsere i underkant av en Gb med data. Hvert enkelt laboratorium kan forventes å laste opp hundrevis av nye genomer per år og dermed forventes det et stort og akkumulerende behov for lagring. Ny metodikk og nye bioinformatiske verktøy muliggjør nå også identifisering av mikrober fra metagenomdata hvor datamengden er enda større. Derfor bør det tilrettelegges for storskala skalerbar lagring. Et annet viktig moment er at data må forsikres mot teknisk svikt, og at rådata/analyserte data som ikke enkelt kan regenereres må lagres på områder med kontinuerlig back-up.



Figur 2: SWISS Pathogen Surveillance Platform (SPSP)



Figur 3 illustrer prinsippkisse av systemstruktur. De ulike elementene diskuteres i avsnittene nedenfor.



3.2 Analyse

Det bør utføres kvalitetskontroll av alle data som lastes opp til overvåkningsdatabasen. Dette kan inkludere ulike kvalitetsmål for både rådata og/eller bearbeidede sekvenser (assemblies). Data som ikke passerer ett eller flere kvalitetsmål bør flagges, slik at dette tydeliggjøres ved eventuell bruk i videre analyser. Videre skisseres behov for bioinformatiske analyser og verktøy innen ulike kategorier som er særlig relevante innen mikrobiologi, med ulike krav til prosesseringskapasitet som angitt i tabell 2.

Tabell 2 viser analyseverktøy og områder som anses som mest nødvendige i en nasjonal overvåkningsdatabase:

Analyseverktøy	Egenskaper/krav
Verktøy for kvalitetskontroll, de novo assembly og mapping av rådata til valgt referansegenom	Avhengig av prosesseringaktivitet
Typingsverktøy (e.g. MLST, cgMLST, SNP-typing, serotyping, spa-typing, kapselotyping) tilpasset de mest vanlige humane patogener	Mindre ressurskrevende analyser
Verktøy for fylogeni- og clusteranalyser både for kjernegenom og pangenom	Avhengig av prosesseringaktivitet
Prediksjon av gener og mutasjoner assosiert med antimikrobiell resistens.	Mindre ressurskrevende analyser
Prediksjon av gener assosiert med virulens	Mindre ressurskrevende analyser
Karakterisering av mobile genetiske elementer (e.g. plasmider, profager, patogenitetsøyer)	Mindre ressurskrevende analyser

Det understrekes at det er behov for et stort antall ulike verktøy og at det vil være et kontinuerlig behov for oppdateringer og oppgraderinger av både verktøy og databaser.

3.3 Visualisering

Visualisering og videre analyser er en svært viktig del av en nasjonal overvåkningsdatabase. Dette vil innebære visualisering av f.eks. fylogenetiske trær, epikurver og kart, hvor det er hensiktsmessig å kunne filtrere dynamisk, farge og reanalysere grupper av prøver basert på tekniske og kliniske metadata som f.eks. kvalitet agens, sekvenstype, tidsintervall (prøve tatt), helseregion, osv.

3.4 Administrasjon

Det er behov for en administrasjonsportal hvor hver bruker har tilgang til egne data, prosjekter og analyser og kan definere tilgang til definerte brukere, grupper eller gjøre data tilgjengelig for alle. Tilganger bør være dynamiske, slik at de enkelt kan endres ved behov. Det er viktig at all bruk av data



i plattformen er styrt av dataavtaler mellom institusjonene. Alle sekvensdata skal eies av den som laster dem opp og det anbefales at institusjoner (f.eks. en mikrobiologisk avdeling eller et referanselaboratorium) står som eiere - ikke enkeltpersoner. Bruk av andres sekvenser i publikasjoner må alltid avklares med eier, også for offentlige sekvenser.

3.5 Kontinuerlige overvåkninger

Sykehusene vil ha behov for kontinuerlig overvåkning av en rekke bakterier. Noen av disse overvåkningene vil være lokale, mens andre vil involvere andre sykehus eller hele landet. Dette krever at man fortløpende kan legge nye bakteriegenomer til eksisterende overvåkninger og deretter få oppdatert rapporter og visualiseringer tilsvarende. For fellesovervåkninger vil det også være behov for en kuratorrolle. En kurator skal kunne bestemme hvilke genomer som skal inngå i løpende overvåkninger og må både kunne legge til og fjerne genomer fra visualiseringer og rapporter.



4 Oversikt og avgrensning av overvåkingsfunksjon

Ved etablering av en overvåkingsfunksjon som den aktuelle databasen er tiltenkt å inneha, i tillegg til deling og analyse av mikrobielle sekvenser, foreligger flere anbefalinger og guidelines fra folkehelsefaglige hold (CDC, ECDC mfl.). Felles for disse er en anbefaling med beskrivelse av en rekke karakteristika ved systemet.

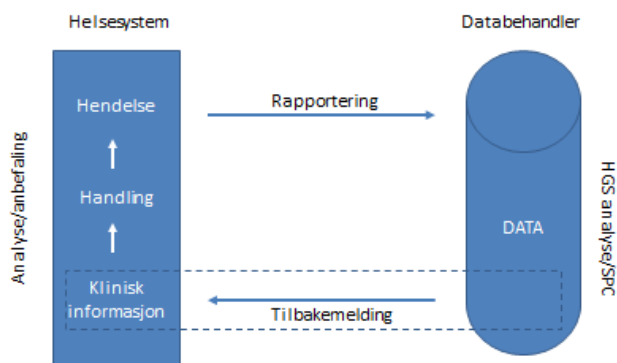
Tabell 3 presenterer karakteristika ved systemet med kommentarer knyttet til det aktuelle systemet:

Elementer *	Kommentar generelle/Nasjonal mikrobiologisk overvåkingsdatabase
Establish objectives	<ol style="list-style-type: none">1. Deling av sekvensdata nasjonalt2. Oppdaging og støtte til oppklaring av (nosokomiale) utbrudd av sykdomsfremkallende bakterier (og virus)
Develop case definitions	Kasusdefinisjon vil avhenge av agens, men ved aktiv innsending av sekvens skal det medfølge metadata- (tid, sted, person samt kliniske, laboratorie- og epidemiologiske opplysninger) av et omfang som tillater hensiktsmessig konstruksjon av kriterier for kasus definisjon.
Determine data sources, data collection mechanism (type of system)	Datakilde vil være norske mikrobiologiske laboratorier med sekvenserings-virksomhet. Innsamling vil være frivillig og systemet vil til enhver tid ha full oversikt over deltagende laboratorier. Detaljer vedrørende sikker oversendelse av sekvenser og metadata samt tilgang til data må beskrives i senere fase av prosjektet.
Determine data collection instruments	Tekniske og strukturelle detaljer er (se tabell 2).
Field test methods	Format av sekvenser og metadata skal defineres (se tabell 2).
Develop and test analytic approach	<ol style="list-style-type: none">1. Enkle, men kontinuerlig pågående temporale og spatiale analyser (statistisk prosess kontroll mm). Ved behov kan mere avanserte analyser utføres av bidragsytere.2. Se tabell 2 for bioinformatiske verktøy/analyse.3. Se også kasus definisjon.
Develop dissemination mechanism	Utkomme (resultater) av pågående overvåkingsanalyse meddeles f.eks. per epost ukentlig til relevant personell ved bidragsytende institusjoner (avdeling ev. regional kompetansetjeneste smittevern).
Assure use of analysis and interpretation	Årlige workshops for deltagende institusjoner skal sikre riktig bruk, analyse og tolkning av data. RHF'ene har overordnet tilsyn med analyseresultater samt drift og vedlikehold.

*modifisert fra [Epi Bul 1995](#).



Designutkast



Figur 4: Illustrer systemet, merk at figuren presenterer ikke sikkerhetstiltak (personvern, tilgangsstyring etc.)

4.1 Metadata

Etablering av en nasjonal overvåkningsdatabase for mikrobesequenser med hensikt å innsamle og kvalitetssikre mikrobielle helgenomdata for å sammenlikne sekvenser med tanke på slektskap og oppdage utbrudd eller vesentlige endringer i populasjonsstruktur fordrer et sett med metadata. Kliniske metadata defineres i denne sammenheng som opplysninger om tid, sted, person samt kliniske, laboratorie- og epidemiologiske opplysninger knyttet til hvert isolat (sekvens) se tabell 4.

En kasusdefinisjon vil bestå av ett bakteriefunn med tilhørende HGS-karakterisering. Det vil være enkelt å definere en rekke spesiesspesifikke kasusdefinisjoner knyttet opp mot resistensfenotyper (VRE, MRSA etc.), men det må også være hensiktsmessig å kunne definere nye – ikke resistens relaterte – kasusdefinisjoner, utelukkende på bakgrunn av bakteriefunn og HGS- karakterisering, jamf. det pågående *Pseudomonas aeruginosa* utbruddet⁶.

Tabell 4: Viser hvilke opplysninger systemet bør inneholde for å kunne oppfylle målsetningen om nødvendige metadata:

Variabel	Kommentar (format)
Person ID	Personnummer (D - nummer) Ev. bostedskordinater som kan kobles til et kart (anonymisering ved rapportering)
Tid for prøvetaking	tt:min:dd:mm:åå
Prøvemateriale	Blod, urin, luftveier etc
Sykehus/institusjon ev. avdeling og klinisk post	Standardisert navn
Infeksjonsstatus	Infeksjon/bærerskap

⁶ <https://pubmed.ncbi.nlm.nih.gov/35514305/>



Merk at anbefalte datasett er begrenset slik at det vil kunne rapporteres av innsendende mikrobiologiske avdeling uten henting av data fra pasientjournal, men vil kun oppdage utbrudd på bakgrunn av opphopning i tid og sted. For oppklaring av utbrudd vil en trenge vesentlig flere kliniske, laboratorie- og epidemiologiske opplysninger som bl.a. vil kunne supplere nevnerdata – dette bør foregå i regi av lokalt og/eller regionalt smittevern. Differensiering av metadata i noen obligatoriske og noen «frivillige» innsendte metadata, er hensiktsmessig for å sikre sammenligningsgrunnlag. De tekniske metadata bør være obligatoriske. Metadata må følge standardisert nomenklatur for å kunne benyttes i en søkefunksjon. Forhold knyttet til personvern ved lagring av sekvenser med tilhørende metadata må gjennomgås grundig som del av prosjektet.



5 Vedlegg

5.1 Vedlegg 1

Tabell 5: Oversikt over funksjonelle deler i en nasjonal plattform for overvåkning av bakterier og utbruddsoppløring

Opplastningsportal	Betraktninger
Opplasting av HGS rådata	Opplasting av data må kunne utføres både via GUI og API
Råsekvens	Mulighet for opplasting av råsekvens vil gjøre at flere med begrenset in-house kompetanse og bioinformatiske verktøy kan benytte plattformen
Assembly	Plattform må inneha verktøy for kvalitetskontroll av assemblies.
Opplasting av metadata	
Via online (GUI) skjema	Vil bidra til lavere terskel for bruk
Via strukturert tekst fil	Viktig ifht opplasting av data tilhørende flere prøver
Definert metadata vokabular, ontologi	Gyldige metadata må være klart definert for sammenstilling (f.eks. definerte "verdier" for prøvemateriale etc.) Noe epidemiologiske metadata (f.eks. Prøvetakningsdato, geografi, type prøvemateriale etc.) må være obligatorisk samt data angående sekvenseringsmetodikk, plattform/kit, sekvenseringslengde etc.
Mulighet for å oppdatere/endre meta- og sekvensdata tilhørende prøve i etterkant	
Analyseportal	
Assembly verktøy for råsekvens	Plattform bør ha verktøy for assembly av råsekvens. Det bør defineres om det skal være flere muligheter eller ett assembly verktøy
Kvalitetskontroll av råsekvens/assembly	Plattform må ha verktøy for kvalitetskontroll av sekvensdata (f.eks. Bifrost SSI - https://github.com/ssi-dk/bifrost). Det må defineres QC kriterier (dekning, dybde, antall contig, N50 ...etc) og hva som gjøres med evt. sekvensdata som ikke oppfyller QC og tilhørende metadata
Typningsverktøy (f.eks. MLST, cgMLST, SNP, spa, serotype,...)	Plattform må ha både mikrobepesifikke definerte verktøy og generiske typningsverktøy. Det må etableres rutiner for deling av data med internasjonale databaser for innhenting av data ved nye sekvenstyper (ST) etc
Resistengen/mutasjonsprediksjon	Database (f.eks. NCBI, CARD, ResFinder etc) og verktøy (f.eks. BLAST, k-mer) for resistengen/mutasjonprediksjon må defineres samt kriterier for prediksjon (% dekning, % identitet).
Virulensprediksjon, Plasmidreplikonprediksjon og evt. andre biomarkører	Må definere verktøy og databaser (finnes spesifikke verktøy/databaser avhengig av mikrobe)
Standardisert rapport	Det må vurderes om analyse skal generere standardisert rapport med definerte karakteristika for inkluderte prøver
Visualiseringsportal	
Fylogenetiske trær (basert på f.eks. SNP, cgMLST, K-mer etc.)	Må ha mulighet til å endre farge-koding av isolater etter metadata (f.eks. laboratorium, år, måned etc. /genetiske markører (f.eks. ST, resistensgener etc.). Legge inn definerte kriterier for klusterbestemmelse. Må ha zoom og søke funksjonalitet. Overordnede nasjonale trær for enkelt species etableres og subtrær for spesifikke analyser. Mulighet for å velge isolater til sub-analyse Mulighet for å opprette "levende" nasjonale trær som utvides kontinuerlig etterhvert som nye genomer kommer til. Disse trærne må ha en kurator som bestemmer hvilke nye genomer som skal inngå



Transmisjonsanalyse	permanent (f.eks. Referanselaboratoriet for aktuell mikrobe). Mulighet for “real-time” analyse. Muligheter for data inn i kart for geografisk analyse. Kriterier for epidemiologiske “advarsler” må defineres.
Rapport	Etablering av standardisert rapport må vurderes.
Administrasjonsportal	
Datakurerer – referansegenomer, databaser (MLST, spa, resistensgener/mutasjoner, virulens etc.), versjoner, endringslogg	
Datakurerer - fylogenetiske trær	Definering av ansvarlige for nasjonale mikrobe spesifikke trær
Tilgangsstyring - forskjellig nivå	Begrenset tilgang Utvidet tilgang Full tilgang Gruppetilgang (laboratorium/HF)/personlig tilgang, tilgang til persondata eller ikke
Loggføring av tilgang/bruk	

5.2 Vedlegg 2

Forfattere

RHF	HF	Navn	Fagområdet
Helse Sør - Øst	OUS	Fredrik Müller	Prosjektleder/ mikrobiolog
Helse Sør-Øst	OUS	Jørgen Vildersjø Bjørnholt	Mikrobiolog
Helse Sør-Øst	OUS	Andre Ingebretsen	Mikrobiolog
Helse Vest	HUS	Øyvind Kommedal	Mikrobiolog
Helse Nord	UNN	Ørjan Samuelsen	Mikrobiolog
Helse – Midt Norge	St. Olavs	Christina Gabrielsen Ås	Molekylærbiolog
Helse – Midt Norge	St. Olavs	Hans-Johnny S. Nilsen	Ansvarlig NorPreM/mikrobiolog
Helse Sør-Øst	OUS	Olga Nævisdal	Prosjektstøtte



Nasjonalt kompetansenettverk innen persontilpasset medisin (NorPreM)



NorPreM

Nasjonalt kompetansenettverk
for persontilpasset medisin