

## Diagnostisk nøyaktighet av ulike molekylære tester for påvisning av SARS-CoV-2 arvestoff – omtale av en oppsummering fra det europeiske nettverket for metodevurdering, EUnetHTA

I en ny rapport fra det europeiske nettverket for metodevurdering (EUnetHTA) er diagnostisk nøyaktighet av 12 ulike klasser av molekylære tester for å avdekke SARS-CoV-2 arvestoff vurdert. Forfatterne fant at diagnostisk nøyaktighet på tvers av testklassene er sammenliknbare når testene ble brukt til å påvise SARS-CoV-2 hos pasienter med covid-19 symptomer. Funnene gir grunnlag til å anta at påvisning av SARS-CoV-2 arvestoff med alternative tester til revers transkriptase sanntids polymerasekjedereaksjon (RT-PCR) kan bidra til å løse problemer knyttet til gjeldende diagnostiske protokoller og øke testkapasiteten. Forfatterne fant imidlertid ikke forskning som gir sikre svar knyttet til diagnostisk nøyaktighet ved testing av personer uten symptomer, pasienter som følges opp etter sykdom eller ved kontaktsporing. Forfatterne fremhever at begrensningene i kunnskapsgrunnlaget må tas hensyn til når beslutninger knyttet til valg av teststrategi skal tas.

Omtale av oppsummert forskning

Vigdis Lauvrak, seniorrådgiver, Folkehelseinstituttet

Lene Juvet, fagdirektør, Folkehelseinstituttet

Folkehelseinstituttet, januar 2021

### Bakgrunn

Tilgang til diagnostiske verktøy for å begrense spredningen av SARS-CoV-2-viruset og dempe den globale effekten av covid-19-pandemien er avgjørende (3). Testkapasiteten kan begrenses av en rekke faktorer. Disse inkluderer en global mangel på utstyr og reagenser, behandlingstid for resultater og andre ressurskrav. I mangel av en gullstandard for påvisning av covid-19, anbefaler WHO at bekreftelse av mistenkte tilfeller av covid-19 bør være basert på påvisning viralt RNA (arvestoff) ved nukleinsyre-amplifikasjonstester slik som RT-PCR (1,3). RT-PCR-testing krever spesialutdannet personale og utstyr. På grunn av tid og ressursavhengigheten, kan RT-PCR være uegnet når rask respons for å unngå økt smitteoverføring er viktig. Balansen mellom ulike variabler, inkludert formålet med teststrategien og kapasiteten, kan være drivere for utforming av testpolitikk, og kontaktsporingsprogrammer. En rekke ulike molekylære tester med forskjellige metoder og kommersielt tilgjengelige test systemer er utviklet og markedsført for å øke testkapasiteten og gjennomføringshastigheten. Hver av variasjonene i metodene gjenspeiler forskjellige diagnostiske testprestasjoner. Likevel vil hver variasjon kunne ha sine egne fordeler og kan brukes avhengig av konteksten til teststrategien og tilgjengelige ressurser. Det Europeiske nettverket for metodevurdering, EUnetHTA har derfor gjennomført en oppsummering av den diagnostiske nøyaktigheten for molekylære tester for påvisning av SARS-CoV2 virus arvestoff.

Målet med rapporten fra EUnetHTA (1) var å gi beslutningstakere i ulike europeiske land grunnlag for å besvare to spørsmål:

- Hvordan best teste pasienter med kliniske manifestasjoner av SARS-CoV-2 for å bekrefte en diagnose av COVID-19?
- Hvordan best identifisere bærere uten symptomer (asymptomatiske) og overvåke nærkontakter?

Rapporten er ment å kunne formidles helt eller delvis til relevante beslutningstakere i Europa.

## Hva bygger konklusjonene på?

Rapporten har systematisk gjennomgått og vurdert forskning om diagnostisk nøyaktighet av molekylære tester for påvisning av SARS-CoV2 arvestoff. Rapporten gir ingen anbefalinger knyttet til analyse, utvinningsplattform eller produsent. For en oversikt over kommersielle diagnostiske tester og deres regulatoriske status i Europa, henviser forfatterne til Joint Research Center (JRC) opprettet av Europakommisjonen (2), og [JRC COVID-19 test databasen](#).

Forskningsspørsmålet i rapporten var:

Hva er den diagnostiske nøyaktigheten av molekylære metoder for påvisning av SARS-CoV-2 arvestoff hos personer mistenkt for å ha Covid-19?

Diagnostisk nøyaktighet ble sammenliknet mot RT-PCR (på prøver fra øvre luftveier) eller RT-PCR kombinert med kliniske funn. Det var planlagt å gjøre analyser basert på testegenskaper og om mulig subgruppeanalyser basert på hvordan testene var tatt og behandlet. Tabell 1 viser detaljer for spørsmålet og inklusjonskriteriene for studier. For å besvare spørsmålet utførte forfatterne uttømmende systematiske søk etter forskning. Søket ble sist oppdatert mellom 29 og 31 juli 2020. Forfatterne gir også en oversikt over pågående forskning og har sjekket for publikasjoner fra denne forskningen i november 2020.

**Tabell 1** Forskningsspørsmål og inklusjonskriterier

<b>Populasjon</b>	Mulige eller mistenkte tilfeller (uansett alder, men der det er mulig kategorisert som barn hvis alder <18 og voksen hvis alder ≥18) av covid-19 testet for diagnose på grunnlag av kliniske symptomer, kontaktsporing eller som en del av massescreening
<b>Helsetilstand (target)</b>	Aktiv infeksjon med SARS-CoV-2
<b>Indeks test</b>	Enhver molekylær test basert på amplifikasjon av nukleinsyre, slik som RT-PCR eller isotermske RNA-amplifikasjonsmetoder, som er designet for å oppdage tilstedeværelsen av SARS-CoV-2-infeksjon hos personer med mistanke om covid-19.
<b>Referansestandard</b>	RT-PCR ved bruk av en validert analyse, alene eller i kombinasjon med kliniske funn
<b>Utfall</b>	<ul style="list-style-type: none"> <li>• Andel sanne og falske positive testresultater samt sanne og falske negative testresultater</li> <li>• Sensitivitet</li> <li>• Spesifisitet</li> <li>• Areal under kurven fra ROC-kurven</li> <li>• Positive og negative prediktive verdier</li> <li>• Andel av inkonklusive testresultater</li> </ul>
<b>Studiedesign</b>	Retrospektive og prospektive kohorter, case serier / case control studier (med minimum 10 deltakere) og tverrsnittstudier som evaluerer diagnostisk nøyaktighet og testegenskaper av molekylære tester for påvisning av SARS-CoV-2

Vurderingen omfattet ikke påvisning av virus antigener, antistoff respons eller metoder som febermåling eller andre kliniske symptomer. Studier hvor en 2x2 tabell over positive og negative testresultater ikke kunne settes opp ble ekskludert. Kvaliteten av inkluderte studier ble vurdert ved hjelp av The Quality Assessment of Diagnostic Accuracy Studies 2 (QUADAS-2) og risiko for skjevhet ble vurdert på studie-nivå for hver enkelt 2x2 tabell. Statistiske analyser ble i hovedsak utført i programmet Stata/IC 16.1 (StataCorp LLC, USA). Vi har på bakgrunn av funnene i metodevurderingen

gjennomført vår egen vurdering av tillit til resultatene med redskapet Grading of Recommendations Assessment, Development and Evaluation (GRADE, <https://www.gradeworkinggroup.org/>).

## Hovedfunn i metodevurderingen

Metodevurderingen identifiserte 3 systematiske oversikter, 103 relevante primærstudier, 14 hurtigvurderinger og 10 pågående kliniske studier. Totalt 168 2x2 tabeller ble laget basert på data fra primærstudiene. Resultatene ble fordelt til 12 relevante diagnostiske testklasser basert på fellestrekk. Det ble ikke funnet tilstrekkelig litteratur til å vurdere diagnostisk nøyaktighet av molekylære tester i screening av asymptomatiske, rekonvalesenspopulasjoner eller som en del av kontaktsporing. Alle resultatene som presenteres er derfor basert på testing av pasienter med symptomer på covid-19 uten hensyn til hvordan prøvene ble tatt. Testklassene og deres fellestrekk er gjengitt i tabell 2.

**Tabell 2** Identifiserte klasser av molekylære tester for påvisning av SARS-Cov-2 arvestoff

Indeks test klasse	Beskrivelse av fellestrekk
Automatiserte RT-PCR systemer	Integrert, høy gjennomstrømming, fullt automatisert arbeidsflyt
Kommersielle RT-PCR kit	Manuelle kommersielle kit basert på RT-PCR teknologi
POCT Systemer	Automatiserte nær pasienten tester (Point of care test, POCT) basert på reaksjoner i lukkede beholdere på stedet (ikke PCR)
Ulike RT-PCR metoder	Egenproduserte (In house) metoder basert på RT-PCR
RT-qPCR	Manuelle systemer basert på kvantitativ RT-PCR
RT-RAA	Manuelle systemer basert på RT rekombinase amplifikasjons teknologien
RT-nPCR	Manuelle systemer basert på RT nested PCR teknologien
dRT-PCR	Manuelle systemer basert på digital RT-PCR teknologi
LAMP	Manuelle systemer basert på LAMP (loop-mediated isothermal amplification) med kolorimetrisk automatisk eller øye basert måling.
RT-LAMP	Manuelle systemer basert på RT- LAMP (reverse transcriptase loop-mediated isothermal amplification) med kolorimetrisk automatisk eller øye basert måling.
TMA	Manuelle systemer basert på TMA (transcription-mediated amplification) teknologien
CRISPR	Manuelle systemer basert på CRISPR (clustered regularly interspaced short palindromic repeats) teknologien.

Forfatterne rapporterer en rekke modeller og analyser. Diagnostisk nøyaktighet uttrykt som sensitivitet og spesifisitet ble gjengitt for hver enkelt inkludert studie og analysert i meta-analyser. Sensitivitet er sannsynligheten for at en person med en gitt tilstand (syk/smittet) får tilstanden påvist med indeks testen sammenliknet med en referanse testen, dvs positiv test. Spesifisitet er sannsynligheten for at en person uten tilstanden (frisk/ikke smittet) får riktig svar med indeks testen sammenliknet med referanse testen, dvs negativ test. Positive prediktive verdier ble beregnet i modeller der det ble benyttet en jevn distribusjon av forekomst (prevalens) på 0-30%. Sannsynligheten for at en test gir falske positive eller falske negative resultater kalles testens positive prediktive verdi (PPV) og testens negative prediktive verdi (NPV) og påvirkes av prevalens (forekomsten av sykdommen i befolkningen som testes). Tabell 3 oppsummerer hovedfunn for samlede estimer for hver testklasse og vår vurdering av tillit til estimatene. Alle resultatene bygger på RT-PCR ved bruk av en validert analyse, alene eller i kombinasjon med kliniske funn som referansestandard. GRADE vurderingene reflekterer vår tillit til at effektestimater er nært en sann verdi. Jo lavere tillit vi har jo mer sannsynlig er det at nye studier vil kunne endre estimatet.

Tabell 3 Hovedfunn og vår vurdering av tillit til resultatene

Indeks test klasse	Kunnskapsgrunnlag	Sensitivitet (95 % KI)	Spesifisitet (95 % KI)	PPV	NPV	GRADE tillit til estimerer**
Automatiserte RT-PCR systemer	10 studier, 10 analyser, 2983 parede tester	0,95 (0,94-0,99)	0,99 (0,99-1,00)	0,93 (0,90-0,96)	1,00 (0,97-1,00)	⊕⊕⊕○ MODERAT <sup>a,b</sup>
Kommersielle RT-PCR kit	13 studier, 25 analyser, 3005 parede tester	0,94 (0,89-0,97)	0,94 (0,89-0,97)	0,99 (0,99-1,00)	0,99 (0,98-0,99)	⊕⊕⊕○ MODERAT <sup>a,b</sup>
POCT Systemer	20 studier, 29 analyser, 4576 parede tester	0,95 (0,91-0,98)	1,00 (0,99-1,00)	0,93 (0,90-0,97)	0,99 (0,96-1,00)	⊕⊕⊕○ MODERAT <sup>a,b</sup>
Ulike RT-PCR metoder	16 studier, 23 analyser, 3676 parede tester	0,97 (0,93-0,98)	1,00 (0,98-1,00)	0,96 (0,93-0,99)	0,99 (0,97-1,00)	⊕⊕⊕○ MODERAT <sup>a,b</sup>
RT-qPCR	6 studier, 9 analyser, 728 parede tester	0,98 (0,96-0,99)	0,99 (0,90-1,00)	0,86 (0,78-0,94)	1,00 (0,92-1,00)	⊕⊕○○ LAV <sup>a,c</sup>
RT-RAA	4 studier, 4 analyser, 1118 parede tester	0,99 (0,73-1,00)	0,99 (0,79-1,00)	0,91 (0,85-0,97)	1,00 (0,94-1,00)	⊕⊕○○ LAV <sup>a,d</sup>
RT-nPCR	3 studier, 4 analyser, 1810 parede tester	0,95 (0,84-0,98)	1,00 (0,50-1,00)	1,00 (0,99-1,00)	0,99 (0,99-0,99)	⊕⊕○○ LAV <sup>a,c,d</sup>
dRT-PCR	3 studier, 3 analyser, 379 parede tester	0,99 (0,86-1,00)	0,83 (0,03-1,00)	Ikke beregnbart	Ikke beregnbart	⊕○○○ SVÆRT LAV <sup>a,c,e</sup>
LAMP	8 studier, 8 analyser, 1059 parede tester	0,87 (0,67-0,96)	1,00 (0,81-1,00)	0,95 (0,91-0,99)	0,98 (0,94-1,00)	⊕⊕○○ LAV <sup>a,d</sup>
RT-LAMP	20 studier, 24 analyser, 3343 parede tester	0,92 (0,82-0,97)	0,99 (0,97-1,00)	0,89 (0,84-0,94)	0,99 (0,94-1,00)	⊕⊕⊕○ MODERAT <sup>a</sup>
TMA	4 studier, 4 analyser, 604 parede tester	0,97 (0,94-0,98)	0,99 (0,98-1,00)	0,93 (0,88-0,97)	0,99 (0,95-1,00)	⊕⊕○○ LAV <sup>a,d</sup>
CRISPR	6 studier, 7 analyser, 517 parede tester	0,97 (0,90-0,99)	0,99 (0,92-1,00)	0,93 (0,87-0,98)	0,99 (0,94-1,00)	⊕⊕○○ LAV <sup>a,d</sup>

KI= konfidens intervall. PPV = Positive prediktive verdier, NPV = Negativt prediktive verdier. PPV og NPV er presentert som summen av sensitivitets analyser med en antatt jevn distribusjon av forekomst (prevalens) mellom 0 og 30 %.

\*GRADE Vurderingen gjenspeiler tillit til resultatene og er gjort av forfatteren av omtalen basert på informasjon i metodevurderingen. <sup>a</sup>Det er betydelig heterogenitet i analysene ( $I^2 > \text{enn } 70\%$  både for sensitivitet og spesifisitet og statistisk variasjon også vurdert ved  $p < 0,05$ ),

<sup>b</sup>Publikasjonsskjevhet (større sjanse for at positive resultater blir publisert) kan ikke utelukkes, <sup>c</sup>Sterk mistanke om publikasjonsskjevhet, <sup>d</sup>Brede konfidensintervall sensitivitet og spesifisitet, <sup>e</sup>Stor risiko for skjevhet i oppsett av studiene.

## Andre resultater

Diagnostisk nøyaktighet beregnet som areal under ROC kurver ble gjengitt og resultatene bidrar til hovedkonklusjonen. Variabler som påvirker heterogeniteten er undersøkt, men det er vanskelig å si hvor sikre vurderingene er gitt svakhetene ved de inkluderte studiene.

Forfatterne fant 10 relevante pågående studier: [NCT04405492](#), [NCT04403672](#), [NCT04245631](#), [NCT04326387](#), [NCT04311398](#), [NCT04408170](#), [NCT04447495](#), [NCT04468217](#), [DRKS00021578](#) og [ISRCTN14966673](#). Vi har sjekket disse og finner at ingen er satt opp for å spesifikt å undersøke diagnostisk nøyaktighet hos asymptomatiske personer. Forfatterne fremhever at det per november 2020 ikke forelå resultater for disse studiene.

## Forfatternes hovedkonklusjon

Forfatterne konkluderer at testene gjennomgående, med unntak av LAMP og dRT-PCR, er sammenliknbare med hensyn til diagnostisk nøyaktighet og derfor egnet både til å bekrefte og avkrefte SARS-CoV-2 infeksjon hos symptomatiske pasienter. LAMP er brukbar for påvisning av virus, men ikke for å utelukke infeksjon. For dRT-PCR fant forfatterne ikke grunnlag for å gjennomføre analyser for prediktive verdier.

## Forfatternes kritiske kommentarer til resultatene

Forfatterne trekker frem at dataene bygger på analyser av symptomatiske pasienter og at utvalget av pasienter i de inkluderte studiene kan gi risiko for skjevheter og overestimering av testnøyaktighetene dersom resultatene overføres til asymptomatiske, rekonvalesentpasienter, og kontaktsporingsprogrammer. Forfatterne trekker også frem at det kan være svakheter med grupperingen i 12 klasser av tester og at det er en del uforklart heterogenitet i analysene. Det vil si en annen gruppering eller eksklusjon av studier fra analyser vil påvirke resultatet. Noe heterogenitet kan forklares med ulikheter i måten prøvene var tatt på og videre behandlet på, men studiene ga ikke grunnlag til å trekke sikre konklusjoner. Forfatterne fremhever at det er behov for studier av høyere kvalitet for å gi sikrere svar.

## Vår vurdering av rapporten og relevans for Norge

Forfatternes konklusjoner er basert på en grundig og systematisk gjennomgang av forskningslitteratur, det er brukt adekvate metoder og rapporten har vært igjennom en grundig fagfellevurdering. Det foreligger flere rapporter om diagnostisk nøyaktighet av alternativer til RT-PCR på øvre luftveisprøver, både de som er inkludert av forfatterne og de som er publisert etter at forfatterne gjennomførte sine søk. Litt ulik tid for søk og kriterier for inklusjon vil føre til at det blir noe ulikhet i primærstudier som blir inkludert. Vi stoler imidlertid på resultatene og støtter hovedkonklusjonene. Vår GRADE vurdering av tillit til resultatene gjenspeiler større usikkerhet knyttet til noen av testklassene.

Vi vurderer at resultatene er relevante for norske forhold, men i Norge har hurtigtester basert på påvisning av SARS-CoV-2 antigener i øvre luftveisprøver vært vurdert som mer aktuelle å innføre enn andre hurtigtester og det har vært diskutert om det i noen tilfeller kan være aktuelt å bruke spytt som prøvemateriale. Antigentester er ikke omfattet av vurderingen fra EUnetHTA. Det irske metodevurderingsinstituttet HIQA publiserte i oktober en metodevurdering (4) som spesifikt så på hurtigtester inkludert antigentester. Rapporten var basert på oppsummert forskning og forfatterne fant svært variable resultater avhengig av test-type. Forfatterne fremhever at det er pågående og nye studier som kan gi sikrere svar. Vi har gjort et forenklet litteratursøk og fant ikke mer oppdaterte (med søk etter september 2020) oppsummeringer om diagnostisk nøyaktighet av antigen tester, men kan ikke utelukke at det foreligger. Tester som hadde brukt spytt som prøvemateriale ble ikke ekskludert av EUnetHTA rapporten, men forfatterne oppsummerte ikke disse som en gruppe. Folkehelseinstituttet har nylig gjennomgått studier som ser på diagnostisk nøyaktighet ved bruk av

spytt som prøvemateriale (5). Søket i vår rapport er mer oppdatert enn søket i EUnetHTA rapporten og vi inkluderte flere studier med spesielt fokus på spytt og RT-PCR, noe som ga oss større grunnlag til å konkludere med hensyn til bruk av spytt som prøvemateriale.

## Kilder

1. EUnetHTA RCRC02 Authoring Team, The diagnostic accuracy of molecular methods that detect the presence of the SARS-CoV-2 virus in people with suspected COVID-19, Collaborative Assessment, Diemen (The Netherlands): EUnetHTA; 2020 4 th of December, N°187 pages, Report No.:RCR02, Tilgjengelig fra: <https://eunetha.eu/rcrot02/> Contact the EUnetHTA Secretariat [EUnetHTA@zinl.nl](mailto:EUnetHTA@zinl.nl) with inquiries about this assessment,
2. European Commission, Current performance of COVID-19 test methods and devices and proposed performance criteria - working document of Commission services 2020. Tilgjengelig fra: <https://ec.europa.eu/docsroom/documents/40805>,
3. WHO, Diagnostic testing for SARS-CoV-2: interim guidance, 11 September 2020, World Health Organization; 2020. Tilgjengelig fra: <https://www.who.int/publications/i/item/diagnostic-testingfor-SARS-cov-2>.
4. Rapid health technology assessment (HTA) of alternatives to laboratorybased real-time RT-PCR to diagnose current infection with severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) Submitted to NPHT: 7 October. Health Information and Quality Authority (HIQA), Ireland. Tilgjengelig fra: <https://www.hiqa.ie/reports-and-publications/health-technology-assessment/rapid-hta-alternative-tests-detect-current>
5. Lauvrak V, Juvet LK. Saliva sample for testing SARS-CoV-2 infection, 1st update on diagnostic accuracy. Systematic review, COVID-19 rapid response 2020. Oslo: Norwegian Institute of Public Health, 2020. Tilgjengelig fra: <https://www.fhi.no/publ/2020/spyttprover-for-testing-av-SARS-cov-2-infeksjon-forste-oppdatering-om-diagn/>