

# Diagnostikk av temporomandibulære tilstander

Rapport fra Kunnskapssenteret nr 10–2013

Systematisk oversikt



 kunnskapssenteret

**Bakgrunn:** Temporomandibulær dysfunksjon (TMD) er en fellesbetegnelse på en gruppe tilstander i kjeveleddet, inkludert omkringliggende tyggemuskulatur og andre strukturer. Det finnes ulike tester for å påvise TMD-tilstander. I denne oversikten over systematiske oversikter oppsummeres vitenskapelig dokumentasjon om diagnostisk nøyaktighet av ulike tester for å identifisere og påvise TMD. **Hovedfunn:** •Undersøkelse med magnetisk resonanstomografi (MR) ser ut til å ha akseptabel nøyaktighet for å påvise eller utelukke diskusforskyvning, diskusdeformitet og ossøse kondylendringer i kjeveleddet, sammenliknet med frysesnitt (post mortem) eller biopsier. •Ultral lyd ser ut til å ha akseptabel nøyaktighet for å påvise, men ikke utelukke, diskusforskyvning, sammenliknet med MR-undersøkelse. •Kliniske tester for undersøkelse av kjeveleddslyder, smerte og kjeveleddsbevegelser, ser ikke ut til å ha tilfredsstillende nøyaktighet for å påvise eller utelukke TMD-tilstander, sammenliknet med MR-undersøkelse. •Elektromyografi ser ikke ut til å ha tilfredsstillende nøyaktighet for å påvise eller

|                         |  |
|-------------------------|--|
| <b>Tittel</b>           | Diagnostikk av temporomandibulære tilstander   |
| <b>English title</b>    | Diagnostics of temporomandibular disorders   |
| <b>Institusjon</b>      | Nasjonalt kunnskapssenter for helsetjenesten   |
| <b>Ansvarlig</b>        | Nylenna, Magne, <i>direktør</i>  |
| <b>Forfattere</b>       | Strøm, Vegard, <i>prosjektleder, Forsker</i><br>Brurberg, Kjetil G, <i>Forsker</i><br>Dahm, Kristin Thuve, <i>Forsker</i><br>Kirkehei, Ingvild, <i>Forskningsbibliotekar</i>                   |
| <b>ISBN</b>             | 978-82-8121-535-1  |
| <b>ISSN</b>             | 1890-1298  |
| <b>Rapport</b>          | Nr 10 – 2013   |
| <b>Prosjektnummer</b>   | 716  |
| <b>Publikasjonstype</b> | Systematisk oversikt   |
| <b>Antall sider</b>     | 57 (75 inklusiv vedlegg)   |
| <b>Oppdragsgiver</b>    | Helsedirektoratet  |
| <b>Emneord(MeSH)</b>    | Kjeveleddsdysfunksjon, temporomandibulær dysfunksjon, TMD, diagnostikk, systematisk oversikt.  |
| <b>Sitering</b>         | Strøm, V, Brurberg KG, Dahm KT, Kirkehei I. Diagnostikk av temporomandibulære tilstander. Rapport fra Kunnskapssenteret nr. 10–2013. Oslo: Nasjonalt kunnskapssenter for helsetjenesten, 2013. |

Nasjonalt kunnskapssenter for helsetjenesten fremskaffer og formidler kunnskap om effekt av metoder, virkemidler og tiltak og om kvalitet innen alle deler av helsetjenesten. Målet er å bidra til gode beslutninger slik at brukerne får best mulig helsetjenester. Kunnskapssenteret er formelt et forvaltningsorgan under Helse- direktoratet, men har ingen myndighetsfunksjoner og kan ikke instrueres i faglige spørsmål.

Kunnskapssenteret vil takke Heming Olsen-Bergem, Brynjar Fure og Inger Natvig Norderhaug for å ha bidratt med sin ekspertise i dette prosjektet. Kunnskapssenteret tar det fulle ansvaret for synspunktene som er uttrykt i rapporten.

Nasjonalt kunnskapssenter for helsetjenesten  
Oslo, juni 2013

# Hovedfunn

Temporomandibulær dysfunksjon (TMD) er en fellesbetegnelse på en gruppe tilstander i kjeveleddet, inkludert omkringliggende tyggemuskulatur og andre strukturer. Det finnes ulike tester for å påvise TMD-tilstander.

I denne oversikten over systematiske oversikter oppsummeres vitenskapelig dokumentasjon om diagnostisk nøyaktighet av ulike tester for å identifisere og påvise TMD. I mangel av en metodisk gullstandard for diagnostisering av TMD-tilstander gir funnene i denne oversikten ikke grunnlag for å avgjøre hvilke tester som best påviser eller utelukker TMD. Derimot kan funnene si noe om og i hvor stor grad ulike diagnostiske tester fører til samme konklusjon.

Oppsummerte resultater fra seks systematiske oversikter med i alt 67 primærstudier og over 5400 deltakere viser at:

- Undersøkelse med magnetisk resonanstomografi (MR) ser ut til å ha akseptabel nøyaktighet for å påvise eller utelukke diskusforskyvning, diskusdeformitet og ossøse kondylendringer i kjeveleddet, sammenliknet med frysesnitt (post mortem) eller biopsier.
- Ultralyd ser ut til å ha akseptabel nøyaktighet for å påvise, men ikke utelukke, diskusforskyvning, sammenliknet med MR-undersøkelse.
- Kliniske tester for undersøkelse av kjeveleddslyder, smerte og kjeveleddsbevegelser, ser ikke ut til å ha tilfredsstillende nøyaktighet for å påvise eller utelukke TMD-tilstander, sammenliknet med MR-undersøkelse.
- Elektromyografi ser ikke ut til å ha tilfredsstillende nøyaktighet for å påvise eller utelukke myofascial TMD sammenliknet med kliniske tester.

Resultatene må tolkes med forsiktighet, fordi primærstudiene i de inkluderte systematiske oversiktene kan ha metodiske svakheter, det er stor variasjon i resultatene, og feilmarginene i nøyaktighetsmålene er gjennomgående brede eller ikke oppgitt.

## Tittel:

Diagnostikk av temporomandibulære tilstander

## Publikasjonstype:

Systematisk oversikt

## Svarer ikke på alt:

- Ingen helseøkonomisk evaluering
- Ingen anbefalinger

## Hvem står bak denne publikasjonen?

Kunnskapssenteret har gjennomført oppdraget etter forespørsel fra Helsedirektoratet, Allmennhelse/ primærhelsesdivisjonen v/ avd.dir. Jon-Torgeir Lunke

## Når ble litteratursøket utført?

Søk etter studier ble avsluttet september 2012.

## Fagfeller:

Fure, Brynjar, forskningsleder, Kunnskapssenteret

Norderhaug, Inger Natvig, forskningsleder, Kunnskapssenteret

Olsen-Bergem, Heming, spesialist i oral kirurgi og oral medisin, universitetslektor, Universitetet i Oslo

---

# Sammendrag

---

---

## Bakgrunn

---

I denne oversikten over systematiske oversikter oppsummerer vi vitenskapelig dokumentasjon om diagnostisk nøyaktighet av ulike tester for å påvise temporomandibulær dysfunksjon (temporomandibular disorder; TMD). TMD er en fellesbetegnelse på en gruppe tilstander i kjeveleddet, inkludert omkringliggende tyggemuskulatur og andre strukturer. Tilstandene karakteriseres av kliniske tegn og symptomer som muskel- og/eller kjeveleddssmerter, ansiktssmerter, kjeveleddslyder (klikke-/skrapelyder; krepitasjon) og begrenset eller avvikende gapeevne.

Symptomer og tegn på TMD forekommer hyppig, men den reelle forekomsten av TMD i Norge er ikke kartlagt. Man skiller mellom ekstraartikulære og intraartikulære TMD-tilstander. I den første gruppen inngår myofasciale smertetilstander, det vil si smerter i muskulatur og sener. De intraartikulære tilstandene omfatter interne kjeveleddsforandringer (som diskusforskyvning med eller uten tilbakegang), degenerative kjeveleddsforandringer (som osteoartritt), og inflammatoriske tilstander (som ulike artritt-/synovitt-sykdommer).

Tester eller prosedyrer for å diagnostisere en gitt tilstand bør være reliabilitetstestet og validert. Studier som reliabilitetstester og validerer prosedyrer, tester eller diagnostiske kriterier, kjennetegnes ved at en indekstest, dvs. den testen man skal evaluere, benyttes på pasienter for så å sammenliknes med resultater fra en gullstandard (eller referansetest) som er benyttet på de samme pasientene. Nøyaktigheten av en diagnostisk test kan karakteriseres ved dens sensitivitet (evne til å oppdage sykdom), spesifisitet (evne til å utelukke sykdom) og sannsynlighetsratioer (sannsynlighet for å ha eller ikke ha sykdom).

---

## Metode

---

Vi søkte systematisk etter vitenskapelig litteratur, dvs. systematiske oversikter som evaluerte diagnostiske tester for påvisning av TMD, t.o.m. september 2012, i følgende databaser: MEDLINE, EMBASE, AMED, PsycINFO, Cochrane Database of Systematic Reviews, Database of Abstracts of Reviews of Effects (DARE), ISI Web of knowledge, PubMed, CINAHL, Prospero, Clinical Evidence og UptoDate. Vi vurderte alle artikler der sammendrag var tilgjengelig på engelsk eller skandinavisk. Vi inkluderte systematiske oversikter av høy eller moderat metodisk kvalitet, med sammenstillinger av primærstudier som reliabilitetstester

eller validerer diagnostiske tester, avbildningsprosedyrer eller kriteriesett for påvisning av kjeveleddsdysfunksjon. Vi hadde også krav om at indekstest og referansetest var benyttet på de samme personene/pasientmaterialet. Tre prosjektmedarbeidere gikk gjennom resultatene fra søkene uavhengig av hverandre. Alle systematiske oversikter som møtte inklusjonskriteriene ble kritisk vurdert ved hjelp av sjekklister, og kvaliteten ble kategorisert som ”høy”, ”moderat” eller ”mangelfull”.

---

## Resultater

---

Litteratursøket ledet til 406 unike referanser, hvorav 29 artikler ble innhentet i fulltekst. Av disse inkluderte vi til slutt seks systematiske oversikter. Disse var publisert i tidsrommet 2006-2012. Totalt omfattet oversiktene 67 unike primærstudier med til sammen ca. 5400 deltakere. Tre av oversiktene vurderte vi til å ha høy- og tre til å ha moderat metodisk kvalitet. Indekstestene som ble evaluert omfattet kliniske tester, magnetisk resonanstomografi (MR), ultralyd og elektromyografi (EMG). Det var benyttet ulike referansetester, men oftest MR.

Tre systematiske oversikter med til sammen 30 primærstudier og ca. 3900 deltakere, omhandlet diagnostisk nøyaktighet av kliniske tester sammenliknet med MR for påvisning av ulike TMD-tilstander. For *kjeveleddslyder* varierte punkttestimatene for sensitivitet fra 0,02 til 0,95, og for spesifisitet fra 0,14 til 1,0. Tilsvarende resultater var for *smerte* fra 0,05 til 1,0, og fra 0,02 til 0,96, for *kjeveleddsbevegelse* fra 0,05 til 0,92, og fra 0,08 til 0,97, og for *testkluster* fra 0,05 til 0,97 og fra 0,26 til 1,0, for hhv. sensitivitet og spesifisitet. De positive sannsynlighetsratioene (LR+) var med få unntak lavere enn 2,5, og de negative (LR-) høyere enn 0,2.

Én systematisk oversikt med 22 primærstudier og ca. 430 deltakere omhandlet diagnostisk nøyaktighet av MR sammenliknet med biopsi eller frysesnitt (post mortem) for påvisning av diskusforskyvning, degenerative eller inflammatoriske lidelser i kjeveleddet. For påvisning av diskusforskyvning varierte sensitivitet fra 0,75 til 0,90 og spesifisitet fra 0,63 til 1,0, og for diskusdeformitet var sensitivitet rundt 0,8, men med svært varierende spesifisitet. For påvisning av ossøse endringer i kjeveleddet varierte sensitivitet fra 0,50 til 0,87 og spesifisitet fra 0,71 til 1,0.

Én systematisk oversikt med 15 enkeltstudier og til sammen 924 deltakere omhandlet diagnostisk nøyaktighet av ultralyd sammenliknet med MR for påvisning av diskusforskyvning. Sensitivitet (ved lukket munn) varierte fra 0,50 til 0,93, og spesifisitet fra 0,47 til 0,92. Statisk ultralyd hadde noe høyere sensitivitet enn dynamisk. Ved ultralyd med munnen åpen var sensitivitet, 0,81 (95 % KI 0,77 -0,84) og spesifisitet 0,78 (95 % KI 0,74 -0,82). Dynamisk avbildning hadde noe bedre diagnostiske egenskaper enn statisk avbildning. Omtrent halvparten av primærstudiene hadde høy eller uklar risiko for metodiske skjevheter.

Én systematisk oversikt omhandlet diagnostisk nøyaktighet av elektromyografi (EMG) sammenliknet med kliniske tester for påvisning av myofascial TMD. Sensitiviteten varierte fra

0,15 til 0,69, og spesifisiteten fra 0,67 til 0,98. Disse resultatene var basert på kun to enkeltstudier som begge hadde høy risiko for metodiske skjevheter.

---

## Diskusjon

---

I mangel av en metodisk gullstandard for diagnostisering av TMD-undergrupper er det viktig å presisere at funnene i denne oversikten ikke gir grunnlag for å avgjøre hvilke tester som best påviser eller utelukker TMD, men derimot om og i hvor stor grad ulike diagnostiske tester fører til samme konklusjon. Mange av primærstudiene har metodiske svakheter som øker risikoen for at resultatene er utsatt for metodiske feilkilder. I tillegg er det gjennomgående et problem at punkttestimatene for diagnostisk nøyaktighet i studier som ser på tilnærmet samme problemstilling varierer mye. En tests sensitivitet og spesifisitet er avhengig av nivået man har satt som grense for om man er syk eller ikke. Forsøk på å øke sensitiviteten til en test vil redusere spesifisiteten og omvendt. Sensitivitet  $\geq 0,7$  og spesifisitet  $> 0,95$  kan anses som akseptable terskler ved diagnostikk av TMD. Dette betyr, basert på det utvalget man har studert, at man vil oppdage minst 70 % av de som faktisk har sykdommen, og at testen vil være negativ hos minst 95 % av de som faktisk er friske. Dette betyr også at 30 % av de som er syke feilaktig vil bli diagnostisert som friske (falsk negative), og 5 % av de friske vil feilaktig bli diagnostisert som syke (falsk positive). Det er imidlertid få enkeltstudier som dokumenterer akseptable verdier. Populasjonen i de inkluderte studiene omfattet gjennomgående selekterte pasientgrupper i spesialisthelsetjenesten. Dette kan ha betydning for generaliserbarheten av resultatene.

---

## Konklusjon

---

- MR-undersøkelse, særlig bilder kombinert i sagittal- og koronalplan, ser ut til å ha akseptabel nøyaktighet for å påvise eller utelukke diskusforskyvning, diskusdeformitet og ossøse kondylendringer i kjeveleddet, sammenliknet med fryseseint (post mortem) eller biopsier.
- Ultralyd ser ut til å ha akseptabel nøyaktighet for å påvise, men ikke utelukke, diskusforskyvning, sammenliknet med MR-undersøkelse.
- Kliniske tester for undersøkelse av kjeveleddslyder, smerte og kjeveleddsbevegelser, ser ikke ut til å ha tilfredsstillende nøyaktighet for å påvise eller utelukke TMD-tilstander, sammenliknet med MR-undersøkelse.
- Elektromyografi ser ikke ut til å ha tilfredsstillende nøyaktighet for å påvise eller utelukke myofascial TMD sammenliknet med kliniske tester.
- Resultatene må tolkes med forsiktighet, fordi primærstudiene i de inkluderte systematiske oversiktene kan ha metodiske svakheter, det er stor variasjon i resultatene, og feilmarginerne (konfidensintervallene) i nøyaktighetsmålene er gjennomgående brede eller ikke oppgitt.

# Key messages (English)

Temporomandibular disorder (TMD) is a term used to describe a number of related disorders affecting the temporomandibular joints, masticatory muscles, and associated structures. To diagnose subgroups of TMD, numerous tests are available. The aim of this overview of systematic reviews is to summarize scientific research on accuracy of different diagnostic tests for the presence of TMD. In the absence of a methodological gold standard for the diagnosis of TMD, it is to be emphasized that the findings of this review does not provide a basis for deciding which tests best detect or rule out TMD, but rather whether and to what extent various diagnostic tests lead to the same conclusion.

## Main findings

Results from six systematic reviews, comprising 67 primary studies and about 5400 participants demonstrate that:

- MRI, especially sagittal + coronal images, appears to have acceptable accuracy to detect or exclude disc displacement, disc configuration and osseous changes of the temporomandibular joint compared with cryosections (postmortem) or biopsies.
- Ultrasonography seems to have acceptable accuracy to detect, but not exclude, disc displacement compared with MRI.
- Clinical tests of joint sounds, pain and movements do not seem to have satisfactory accuracy to detect or exclude TMD when compared with MRI.
- Electromyography does not seem to have sufficient accuracy to detect or rule out myofascial TMD compared with clinical tests.
- The results must be interpreted with caution, because the primary studies included in the systematic reviews may have methodological weaknesses, the results show great variation, and confidence intervals in accuracy measures are wide or not reported.

Title:  
Diagnostics of  
temporomandibular disorders

Type of publication:  
**Systematic review**

A review of a clearly formulated question that uses systematic and explicit methods to identify, select, and critically appraise relevant research, and to collect and analyse data from the studies that are included in the review. Statistical methods (meta-analysis) may or may not be used to analyse and summarise the results of the included studies.

**Doesn't answer everything:**

- Excludes studies that fall outside of the inclusion criteria
- No health economic evaluation
- No recommendations

**Publisher:**

Norwegian Knowledge Centre  
for the Health Services

**Updated:**

Last search for studies:  
September 2012

---

# Executive summary (English)

---

## Background

---

In this overview of systematic reviews we summarize the accuracy of different tests to identify and diagnose temporomandibular disorders (TMD). TMD is a collective term used to describe a number of related disorders affecting the temporomandibular joints, masticatory muscles, and associated structures. The most common signs and symptoms include facial and jaw pain, which can be aggravated by jaw movements, TMD joint noises (clicking or crepitus), and restriction of mandibular movements. The prevalence of TMD signs and symptoms is relatively high, but in most cases resolved spontaneously. In Norway, it is assumed that about 80 persons suffer severe TMD.

Temporomandibular joint disorders may be categorized into extra articular (extra capsular), as myofascial pain disorders, and intra articular (intra capsular) disorders. Intra articular disorders comprise internal derangements (with the sub categories of disc displacement with or without reduction), degenerative joint disorders, and inflammatory TMD. The diagnostic process is seldom straight forward as TMD might comprise different disorders with similar sign and symptoms. Tests used in diagnostics should be accurate, reliable and valid for the purpose. Studies validating diagnostic tests are characterised by the use of an index test, i.e. the test to be evaluated, compared to the results of a reference standard (or gold standard) applied to the same patients. Accuracy of a diagnostic test can be expressed as its sensitivity, specificity and likelihood ratios. Sensitivity is a measure of a tests ability to identify persons that actually has the condition, while specificity is a measure of the tests ability to rule out the presence of a condition.

---

## Methods

---

We searched systematically for research literature until September 2012, in the following databases: MEDLINE, EMBASE, AMED, PsycINFO, Cochrane Database of Systematic Reviews, Database of Abstracts of Reviews of Effects (DARE), ISI Web of knowledge, PubMed, CINAHL, Prospero, Clinical Evidence og UptoDate. We appraised all articles with an abstract in English or a Scandinavian language. Systematic reviews judged to be of high or moderate methodological quality were included, that assessed the diagnostic efficacy of different tests or imaging techniques in detecting TMD, and where the index test was compared with a reference standard and applied on the same patients. Titles and abstracts of identified studies



were assessed independently by three reviewers to judge if the studies matched the inclusion criteria. Included studies were read in full-text and reassessed according to the inclusion criteria. The same three reviewers made independent assessments of the methodological quality of the included studies, categorized as high, medium or unsatisfactory. Check lists were used for this purpose. Information about kappa statistics, from reliability studies, and sensitivity, specificity and likelihood ratios were summarized.

---

## Results

---

The literature search identified 406 unique references of where 29 articles were read in full-text. Finally, we included six systematic reviews, published in the time period between 2006 and 2012. The reviews comprised totally 67 distinctive studies comprising about 5400 participants. We evaluated three of the reviews to be of high methodological quality, and three to be of moderate quality. The index tests evaluated were clinical tests, magnetic resonance imaging (MRI), ultrasonography and electromyography (EMG). As reference standard, MRI was most often used.

Three systematic reviews, comprising 30 primary studies and about 3900 participants, assessed the diagnostic accuracy of clinical tests compared with MRI in detecting TMD. For *joint sounds*, the estimates for sensitivity ranged from 0.02 to 0.95, and for specificity from 0.14 to 1.0. The corresponding result for *pain* were from 0.05 to 1.0, and from 0.02 to 0.96, for *joint movements* from 0.05 to 0.92, and from 0.08 to 0.97, and for *test clusters* from 0.05 to 0.97 and from 0.26 to 1.0, for sensitivity and specificity, respectively. The positive likelihood ratios (LR+) were in general below 2.5, and the negative ratios (LR-) larger than 0.2.

One systematic review, comprising 22 primary studies and about 430 participants, assessed the diagnostic accuracy of MRI compared with biopsies or cryosections in detecting disk displacements, degenerative or inflammatory TMD. In detecting disk displacements, sensitivity ranged from 0.75 to 0.90 and specificity from 0.63 to 1.0. For disk configuration, the sensitivity was around 0.8, however, with large variance in specificity. For detection of osseous changes the sensitivity ranges from 0.50 to 0.87 and specificity from 0.71 to 1.0. These results were based on 12 primary studies assessed, by the review authors, as having a moderate level of evidence.

One systematic review, comprising 15 primary studies and 924 participants, assessed the diagnostic accuracy of ultrasonography compared with MRI in detecting disk displacements with or without reduction. For detection of disk displacements, with the mouth in closed position, the sensitivity ranged from 0.50 to 0.93, and specificity from 0.47 to 0.92. Static ultrasonography had higher sensitivity than dynamic ultrasonography. For ultrasonography with the mouth in open position, the sensitivity was 0.81 (95 % CI 0.77 -0.84) and the specificity 0.78 (95 % CI 0.74 -0.82). Dynamic ultrasonography had higher sensitivity than static ultrasonography.

One systematic review, comprising two primary studies and 148 participants, assessed the diagnostic accuracy of electromyography compared with clinical test in detecting myofascial TMD. The sensitivity ranged from 0.15 to 0.69, and specificity from 0.67 to 0.98. The two studies had high risk of bias.

---

## Discussion

---

In the absence of a methodological gold standard for the diagnosis of TMD, it is important to emphasize that the findings of this review does not provide a basis for deciding which tests best detect or rule out TMD, but rather whether and to what extent various diagnostic tests lead to the same conclusion. Many of primary studies have methodological weaknesses that increase the risk that the results are subject to methodological errors. In addition, there is generally a problem that the estimates of diagnostic accuracy studies looking at approximately the same topics vary widely. This may be due to differences in protocols, different use of thresholds and differences in patient populations. A test's sensitivity and specificity depends on the limits set to be ill or not. Attempts to increase the sensitivity of a test will reduce the specificity and vice versa. A sensitivity  $\geq 0.7$  and specificity  $> 0.95$  can be considered as acceptable thresholds for the diagnosis of TMD. Thus, based on the selection studied, at least 70% of those who actually have the disease will be diagnosed, and the test will be negative in at least 95% of those who actually are healthy. This also means that 30% of those who are sick will be falsely diagnosed as healthy (false negative), and 5% of the healthy will erroneously be diagnosed as diseased (false positive). However, there are few individual studies that measure acceptable values. The lack of precision can be attributed to small primary studies, and considering the wide confidence intervals in the context of the inconsistencies across the included studies, firm conclusions are difficult to draw based on the available data. The populations in the included studies consisted mainly of selected patients from the special health care system, in which purpose was to distinguish between different forms of TMD.

---

## Conclusions

---

- MRI, especially sagittal + coronal images, appears to have acceptable accuracy to detect or exclude disc displacement, disc configuration and osseous changes of the temporomandibular joint compared with cryosections (postmortem) or biopsies.
- Ultrasonography seems to have acceptable accuracy to detect, but not exclude, disc displacement compared with MRI.
- Clinical tests of joint sounds, pain and movements do not seem to have satisfactory accuracy to detect or exclude TMD when compared with MRI.
- Electromyography does not seem to have sufficient accuracy to detect or rule out myofascial TMD compared with clinical tests.
- The results must be interpreted with caution, because the primary studies included in the systematic reviews may have methodological weaknesses, the results show great variation, and confidence intervals in accuracy measures are wide or not reported.

---

# Innhold

|   |           |
|---|-----------|
| <b>HOVEDFUNN</b>                                  | <b>2</b>  |
| <b>SAMMENDRAG</b>                                 | <b>3</b>  |
| Bakgrunn  | 3         |
| Metode  | 3         |
| Resultater  | 4         |
| Diskusjon   | 5         |
| Konklusjon  | 5         |
| <b>KEY MESSAGES (ENGLISH)</b>                     | <b>6</b>  |
| <b>EXECUTIVE SUMMARY (ENGLISH)</b>                | <b>7</b>  |
| Background  | 7         |
| Methods   | 7         |
| Results   | 8         |
| Discussion  | 9         |
| Conclusions                                       | 9         |
| <b>INNHold</b>                                    | <b>10</b> |
| <b>FORORD</b>                                     | <b>12</b> |
| <b>PROBLEMSTILLING</b>                            | <b>13</b> |
| <b>BAKGRUNN</b>                                   | <b>14</b> |
| Temporomandibulær dysfunksjon                     | 14        |
| Utredning og diagnostikk                          | 15        |
| <b>METODE</b>                                     | <b>19</b> |
| Litteratursøk                                     | 19        |
| Inklusjonskriterier                               | 19        |
| Eksklusjonskriterier                              | 20        |
| Utvelgelse av studier for inklusjon               | 20        |
| Vurdering av metodisk kvalitet                    | 21        |
| Dataekstraksjon og analyser                       | 21        |
| <b>RESULTATER</b>                                 | <b>23</b> |
| Beskrivelse av inkluderte systematiske oversikter | 24        |
| Vurdering av metodologisk kvalitet                | 25        |

|   |           |
|---|-----------|
| Diagnostisk nøyaktighet av kliniske tester                                | 25        |
| Diagnostisk nøyaktighet av MR   | 34        |
| Diagnostisk nøyaktighet av CT   | 37        |
| Diagnostisk nøyaktighet av ultralyd                                       | 37        |
| Diagnostisk nøyaktighet av elektromyografi                                | 39        |
| <b>DISKUSJON</b>  | <b>41</b> |
| Hva viser resultatene?  | 41        |
| Utfordringer knyttet til manglende referansestandard                      | 43        |
| Oversikt over oversikter – fordeler og ulemper                            | 44        |
| Dokumentasjonens kvalitet   | 45        |
| <b>KONKLUSJON</b>   | <b>48</b> |
| <b>REFERANSER</b>   | <b>49</b> |
| <b>VEDLEGG</b>  | <b>58</b> |
| Vedlegg 1 - Søkestrategier  | 58        |
| Vedlegg 2 - Sjekkliste for vurdering av metodisk kvalitet                 | 64        |
| Vedlegg 3 - Ekskluderte studier   | 66        |
| Vedlegg 4 - Primærstudier som inngår i inkluderte systematiske oversikter | 67        |
| Vedlegg 5 – Tabeller, diagnostisk nøyaktighet                             | 69        |
| Vedlegg 6 - Begrep og forkortelser  | 74        |

---

# Forord

Helsedirektoratet har fra Helse- og omsorgsdepartementet fått i oppdrag å iverksette tiltak for å styrke tilbudet til personer med temporomandibulær dysfunksjon (TMD), herunder å iverksette strakstiltak for utredning og behandling av personer med alvorlig TMD. Videre skal direktoratet utarbeide en veileder for behandlingsforløp for alle personer med TMD. Nasjonalt kunnskapssenter for helsetjenesten fikk i den forbindelse i oppdrag fra Helsedirektoratet å oppsummere tilgjengelig forskning om utredning, diagnostikk og behandling av TMD. Rapportene som omhandler effekt av behandling er publisert (se [www.kunnskapssenteret.no](http://www.kunnskapssenteret.no)). Mandatet for denne rapporten er å sammenfatte det vitenskapelige materiale som er publisert omkring utredning og diagnostikk av TMD.

Prosjektgruppen har bestått av:

- Prosjektleder: Forsker Vegard Strøm, Kunnskapssenteret
- Forsker Kjetil G. Brurberg, Kunnskapssenteret
- Forsker Kristin Thuve Dahm, Kunnskapssenteret
- Forskningsbibliotekar Ingvild Kirkehei, Kunnskapssenteret

Denne oversikten er i hovedsak ment å bistå helsepersonell, gjennom kunnskap om reliabilitet og validitet av ulike undersøkelsesmetoder, til å kunne bedre kvaliteten i utredning av TMD-pasienter. I møtet med den enkelte pasient må forskningsbasert dokumentasjon ses i sammenheng med andre relevante forhold, pasientenes behov og egne kliniske erfaringer.

Gro Jamtvedt  
*Avdelingsdirektør*

Liv Merete Reinar  
*Seksjonsleder*

Vegard Strøm  
*Prosjektleder*

---

# Problemstilling

I denne oversikten over systematiske oversikter oppsummerer vi tilgjengelig kunnskap om reliabilitetstestede og validerte tester, prosedyrer eller kriteriesett for å utrede og diagnostisere temporomandibulær dysfunksjon.

Vi søker å gi svar på følgende problemstilling:

Hva er nøyaktigheten av ulike diagnostiske tester sammenliknet med referansestandard for å identifisere og påvise temporomandibulær dysfunksjon?

---

# Bakgrunn

---

## Temporomandibulær dysfunksjon

---

Temporomandibulær dysfunksjon (temporomandibular disorder; TMD), eller kjeveleddsdysfunksjon, er en fellesbetegnelse på en gruppe tilstander i kjeveleddet, inkluderer omkringliggende tyggemuskulatur og andre strukturer (1). Kjeveleddet er et av de mest kompliserte leddstrukturene i kroppen. Leddet kan beveges i flere dimensjoner som inkluderer hengsel- eller rotasjonsbevegelse og glide- eller translasjonsbevegelse (fremoverglidning) (2). Flere muskler tar samtidig del i, og påvirker disse bevegelsene, som først og fremst styres av tyggemusklene.

Temporomandibulær dysfunksjon er klassifisert som en muskelskjelettlidelse, og karakteriseres gjerne av symptomer og kliniske tegn som muskel- og/eller kjeveleddssmerte, ømhet i tyggemuskulatur, ansiktssmerter, kjeveleddslyder (klikke-/skrapelyder; krepitasjon) og begrenset eller avvikende gapeevne (3). Smertene stråler ofte ut mot øre og nakke og kan utløses eller forsterkes av bevegelser i kjeven. Studier viser at enkeltstående symptomer på TMD forekommer hyppig i den generelle befolkningen. Opp mot 75 % kan ha minst ett objektivt tegn på TMD, og 5 til 33 % rapporterer subjektive symptomer (3;4). For de aller fleste går imidlertid symptomene oftest over av seg selv etter en tid, men 5-10 % av de som er plaget vil ha behov for behandling (2;3). Forekomsten av TMD i Norge ikke er kartlagt, men det er anslått at om lag 60-80 personer er i Norge i dag er plaget med invalidiserende kjeveleddsdysfunksjon (Helsedirektoratet, personlig meddelelse).

Man skiller gjerne mellom ekstraartikulære (ekstrakapsulære) og intraartikulære (intrakapsulære) TMD-tilstander (3). I den første gruppen inngår myofasciale smertetilstander. De intraartikulære tilstandene kan deles i tre undergrupper; interne kjeveleddsforandringer (som diskusforskyvning med eller uten tilbakegang), eller degenerative kjeveleddsforandringer (som osteoartritt), og inflammatoriske tilstander (som ulike artritt-/synovittsykdommer). Kombinasjoner av disse tilstandene er hyppig forekommende (3).

*Myofasciale smertetilstander* antas å være den mest utbredte av disse tilstandene, og kjennetegnes gjerne ved ikke-spesifikke ansiktssmerter og muskulær ømhet, begrenset gapeevne og smerter ved palpasjon og kjeveleddsbevegelser. Myofasciale smerter er ofte også sammenfallende med spenningshodepine og øre- eller tannverk (5).

*Intern kjeveleddsforandring* kan defineres som et unormalt forhold mellom den mandibulære kondylen og den intraartikulære diskusen ved sammenbitt av tennene (6), og innebærer vanligvis en forskyvning av kjeveleddsdiskus fra dens normale posisjon. *Diskusforskyvning med tilbakegang* (el. reduksjon) er definert som fremoverforskyvning av diskusen ved lukket munn som går tilbake (med el. uten klikk) til normal posisjon i løpet av en gapebevegelse. Hos noen personer med denne tilstanden kan man høre et klikk også når munnen lukkes. Dette kalles resiprokalt klikk. Når diskusforskyvningen ikke går tilbake til normal posisjon i løpet av gapebevegelsen kalles tilstanden gjerne *diskusforskyvningen uten tilbakegang* (reduksjon), eller ”closed lock” (6).

*Degenerative kjeveleddsforandringer* omfatter nedbryting av leddbrusk, mens *inflammatoriske artritter* først og fremst omfatter de bløte delene av leddet (synovitt). Artritter kan skyldes ytre faktorer som traumer og infeksjoner, men kommer oftere av indre årsaker som for eksempel revmatisme, artrose med mer. Artrose/degenerative forandringer kan blant annet komme av diskusforandring (som også kan gi synovitt), lavgradige traumer over tid og tidligere inngrep i leddet (3).

For personer som har symptomer og kliniske tegn på TMD er det viktig å få korrekt diagnose slik at de kan henvises videre i behandlingsapparatet, og når Helsedirektoratet nå skal utarbeide en ny behandlingslinje for TMD omfatter dette også den diagnostiske prosessen. Det er derfor viktig å undersøke om det finnes reliabilitetstestede og validerte diagnostiske tester ved TMD som kan egne seg for implementering i en ny norsk behandlingslinje.

---

## **Utredning og diagnostikk**

---

Siden TMD omfatter ulike tilstander som ofte har liknende kliniske symptomer, kan det være vanskelig å standardisere den diagnostiske prosessen (7). Kliniske tester er ofte et av de første stegene i en diagnostiseringsprosess (7). *Kliniske tester* omfatter kartlegging/registrering/måling av individuelle symptomer som smerte og kjeveleddslyder og begrensninger i kjeveleddsbevegelser, enten enkeltvis eller samlet sett (test-kluster). Et slik test-kluster er etablert gjennom ”Research Diagnostic Criteria for TMD” (RDC/TMD) (8). RDC/TMD er et diagnostisk system bestående av to akser; Akse I er en protokoll for klinisk vurdering, dvs. diagnosesetting av ekstra- eller intraartikulære TMD-tilstander, og Akse II er et screeningsinstrument for å kartlegge psykologisk status og smerterelatert funksjon. Totalt utgjør disse to aksene en omfattende biopsykososial evaluering (9). RDC/TMD er et mye brukt system internasjonalt, og er også tatt i bruk i Norge, særlig ved universitetene i Oslo og Bergen. En forenklet utgave av dette diagnostikksystemet er forventet å komme i løpet av 2013, da under betegnelsen DC/TMD (Diagnostic Criteria/TMD). Det vil bli et forenklet system som skal kunne brukes både til forskning og klinisk bruk i diagnostikk av voksne personer.

Kliniske tester vil imidlertid i utilstrekkelig grad kunne påvise abnormaliteter i kjeveleddet (1;10), og i tilfeller hvor det er behov for ytterligere diagnostisk informasjon gjøres i tillegg ofte radiologiske undersøkelser eller annen bildediagnostikk for å påvise interne og degene-



relative kjeveleddsforandringer (se (7)). For påvisning av myofasciale tilstander har muskelfunksjonsanalyser også blitt benyttet (11).

*Radiologiske undersøkelser* inkluderer hovedsakelig ordinært røntgenbilde av kjeven, computertomografi (CT), lavdose CT (cone-beam computertomografi; CBCT) og magnetisk resonanstomografi (MR) (7). MR anses gjerne som et førstevalg, hvor diskusposisjoner, morfologi, benstrukturer og eventuelle inflammasjoner kan visualiseres for påvisning og diagnostisering av intraartikulære TMD-tilstander (7;10). Tidligere har man vist at andre diagnostiske metoder som *artroskopi/artrografi* har vært nyttig for å påvise intern kjeveleddsforandring (3). Dette er imidlertid en invasiv metode og har etter hvert blitt erstattet av MR. *Ultralyd* er en bildediagnostisk metode som i de senere år har blitt introdusert i TMD-diagnostikk, og som ser ut til å kunne egne seg for påvisning av diskusposisjoner, både gjennom statiske og dynamiske målinger (12). I tillegg er ultralyd ofte mindre kostnadskrevenende og lettere tilgjengelig enn MR. *Muskelfunksjonsanalyser* baserer seg primært på elektromyografiske (EMG) data (11). Bruken av EMG er basert på antakelsen om at patologiske og dysfunksjonelle tilstander kan identifiseres gjennom unormal aktivitet i kjeveleddsmuskulaturen (11).

Tester eller prosedyrer for å identifisere og diagnostisere en gitt tilstand, bør være reliabilitetstestet og validert. Studier som reliabilitetstester og validerer prosedyrer, tester, eller diagnostiske kriteriesett kjennetegnes ved at en indekstest, dvs. den testen/prosedyren man skal evaluere, benyttes på pasienter for så å sammenliknes med resultater fra en referansestandard, eller gullstandard, som er benyttet på de samme pasientene (13). En referansestandard bør være den best tilgjengelige metoden for å avgjøre om en person har den gitte tilstanden eller ikke. Per i dag eksisterer det, så vidt vi vet, ingen omforent referansestandard i diagnostisering av de ulike TMD tilstandene, selv om bruk av MR ofte anses som en slik standard, særlig for de intraartikulære tilstandene (14;15).

### **Diagnostisk nøyaktighet**

Å diagnostisere innebærer bruk av tester/undersøkelser for å avklare tilstedeværelse eller fravær av en sykdom, tilstand eller syndrom (16). Forutsetningen for at vi skal ha nytte av en test er at den er nøyaktig, og at den er til å stole på og måler det den er ment å måle, altså at testen er reliabel og valid. En ideell diagnostisk test er så nøyaktig at den uten feil skiller de personene som er syke og de som er friske (17). Nøyaktigheten av en diagnostisk test kan karakteriseres ved dens sensitivitet og spesifisitet (18).

#### *Sensitivitet og spesifisitet*

Sensitivitet og spesifisitet er to måter en diagnostisk test kan kvantifiseres på. Sensitivitet er et mål på i hvilken grad testen klarer å identifisere personer som har en sykdom eller tilstand. Den er altså testens evne til å klassifisere syke personer (som her; personer med TMD) som syke. Mer presist angir sensitivitet andelen av sanne positive (de som faktisk har en sykdom/tilstand; Figur 1) som korrekt identifiseres av testen (18). Spesifisitet angir testens presisjon for å utelukke sykdom eller tilstand, det vil si testens evne til å klassifisere friske per-

soner som friske. Mer presist angir spesifisitet andelen av sanne negative (de som faktisk ikke har en sykdom/tilstand; Figur 1) som korrekt identifiseres av testen (18).

For å beregne sensitivitet og spesifisitet bruker man gjerne en 2 x 2 tabell (Figur 1) hvor man setter inn resultater fra testingen. Hvis resultatet for hhv. sensitivitet og spesifisitet blir f.eks. 0,90 og 0,75, kan man si, basert på det utvalget man har studert, at man vil oppdage 90 % av de som faktisk har sykdommen, og at testen vil være negativ hos 75 % av de som faktisk er friske. Dette betyr også at 10 % av de som er syke feilaktig vil bli diagnostisert som friske (falsk negative), og 25 % av de friske vil feilaktig bli diagnostisert som syke (falsk positive). Det ideelle er at både sensitivitet og spesifisitet er på 100 %. Det er imidlertid verd å merke seg at en tests sensitivitet og spesifisitet er avhengig av terskelverdien man har satt som grense for om man er syk eller ikke. Det er også viktig å være klar over at forsøk på å øke sensitiviteten til en test ofte skjer på bekostning av redusert spesifisitet og omvendt.

|   |              | <b>Patologi</b><br>(Resultat av referansetest) |                              | Total   | <b>Prediktiv verdi</b>  |
|---|--------------|--|------------------------------|---------|-------------------------|
|   |              | Syk  | Ikke syk                     |         |                         |
| <b>Test</b><br>(Resultat av indekstest) | Positiv test | <b>Sann positiv a</b>                          | <b>Falsk positiv b</b>       | a+b     | <b>Positiv</b><br>a/a+b |
|   | Negativ test | <b>Falsk negativ c</b>                         | <b>Sann negativ d</b>        | c+d     | <b>Negativ</b><br>d/c+d |
| Total                                   |              | a+c  | b+d                          | a+b+c+d |                         |
|   |              | <b>Sensitivitet</b><br>a/a+c                   | <b>Spesifisitet</b><br>d/b+d |         |                         |

**Figur 1.** Beregning av diagnostisk nøyaktighet.

### *Positiv og negativ prediktiv verdi*

I klinisk praksis vet man ofte bare testresultatet, f.eks. at testen er positiv, derfor vil man gjerne vite hvor god en test er til å kunne predikere dette funnet (19). Hvor stor er sannsynlighet for at testen har gitt riktig resultat? Når vi møter en pasient som vi er usikre på om har TMD, har vi derfor ofte mer nytte av å vite testens prediktive verdi. En tests prediktive verdi kombinerer sann og falsk positiv (eller negativ) i én verdi (Figur 1).

Hvis en person tester positivt, hva er da sjansen for at han eller hun har tilstanden? *Positiv prediktiv verdi* (PPV) er altså sannsynligheten for at en person med et positivt testresultat

faktisk har sykdommen eller tilstanden (19). Hvis testen har høy positiv prediktiv verdi for eksempel 98 %, så betyr det at det er 98 % sjanse for at personen har sykdommen dersom resultatet av testen er positiv. Vi vil da være nesten helt sikre på at pasienten tester positivt fordi vedkommende faktisk er syk, og ikke fordi testen er falsk positiv. Hvis en person tester negativt, hva er da sjansen for han eller hun ikke har tilstanden? *Negativ prediktiv verdi* (NPV) er sannsynligheten for at en person med et negativt testresultat ikke har sykdommen eller tilstanden (19). Dersom testens negative prediktive verdi er lav, for eksempel 48 %, vet vi at selv om en pasient har en negativ test, har ca. halvparten av dem som får negativt utslag likevel sykdommen.

Mens sensitivitet og spesifisitet er knyttet til selve testen er de prediktive verdiene ikke bare avhengig av sensitiviteten og spesifisiteten, men også av forekomsten (prevalensen) av sykdommen/tilstanden (19). Prevalensen i utvalget kan beregnes ved  $(a+c)/(a+b+c+d)$  (Figur 1). Dette betyr også at de prediktive verdiene ikke er overførbare fra en pasient til en annen, eller fra en setting til en annen (Attia 2003).

### *Sannsynlighetsratio*

For et gitt testresultat kan man sammenlikne sannsynligheten for at pasienten virkelig har tilstanden/sykdommen, med tilsvarende sannsynlighet for at personen ikke har tilstanden. Dette kalles sannsynlighetsratio ("likelihood ratio"; LR) (20). Sannsynlighetsratio er uavhengig av sykdomsprevalensen.

Hvis testen bare kan ha positivt eller negativt svar, blir sannsynlighetsratio for positivt svar ( $LR^+$ ), lik sensitivitet/(1-spesifisitet), og sannsynlighetsratio for negativt svar ( $LR^-$ ) blir (1-sensitivitet)/spesifisitet.  $LR^+$  indikerer således hvor mye mer sannsynlig det er å få et positivt testresultat hos personer som virkelig har tilstanden, enn hos dem som ikke har tilstanden. Den beste testen for å diagnostisere en person med en tilstand er den testen med høyest  $LR^+$ . En  $LR^+$  større enn 10 indikerer en stor og ofte konklusiv sannsynlighet for at sykdommen er tilstede (21). En *negativ LR* ( $LR^-$ ) indikerer hvor mye mer sannsynlig det er å få et negativt testresultat hos personer som er friske enn hos personer som er syke. Den beste testen for å utelukke en tilstand hos en person er den testen med den laveste  $LR^-$ . En  $LR^-$  mindre enn 0,1 indikerer gjerne en konklusiv sannsynlighet for at tilstanden ikke er tilstede. En  $LR^+$  eller  $LR^-$  på 1,0 betyr at et henholdsvis positivt eller negativt testresultat er like sannsynlig hos en person som har tilstanden som hos en person som ikke har tilstanden (21).

Sannsynlighetsratioer kan tolkes på følgende måte (21):

| $LR^-$  | $LR^+$ |                                       |
|---------|--------|---------------------------------------|
| 1       | 1      | Ingen klinisk verdi                   |
| 0,2–0,5 | 2–5    | Liten forskjell som kan være relevant |
| 0,1–0,2 | 5–10   | Beskjeden, men vesentlig forskjell    |
| <0,1>   | >10    | Klinisk viktig forskjell              |

---

# Metode

---

## Litteratursøk

---

### Søkebaser

Vi søkte systematisk etter litteratur i følgende databaser: MEDLINE, EMBASE, AMED, PsycINFO, Cochrane Database of Systematic Reviews, Database of Abstracts of Reviews of Effects (DARE), ISI Web of knowledge, PubMed, CINAHL, Prospero, Clinical Evidence og UptoDate.

Søkene var sammensatt av emneord og tekstord for temporomandibular dysfunksjon og avgrenset videre med søkeord for diagnostikk og valideringsstudier. Søket ble til sist avgrenset med et presist søkefilter for systematiske oversikter.

Alle søk ble gjennomført av forskningsbibliotekar (IK) og avsluttet i september 2012.

Søkestrategier er rapportert i Vedlegg 1a.

For å sikre oss at alle relevante systematiske oversikter var identifisert, gjennomgikk vi også referansene fra et søk som var utført i april 2012 i forbindelse med utarbeidelse av rapport om behandling av TMD. Dette søket var bredere og var bygd opp for å finne alle systematiske oversikter om temporomandibular dysfunksjon. Søkestrategiene for dette søket er rapportert i Vedlegg 1b.

---

## Inklusjonskriterier

---

### Studiedesign

Vi ønsket primært å inkludere systematiske oversikter med sammenstillinger av primærstudier som reliabilitetstester eller validerer diagnostiske prosedyrer, tester, eller kriteriesett for TMD. Vi åpnet for å inkludere primærstudier dersom vi ikke fant systematiske oversikter som dekket feltet på en tilfredsstillende måte.

### Populasjon

Pasienter med tegn eller symptomer på kjeveleddsfunksjon, også kalt temporomandibular dysfunksjon eller TMD.

## **Indekstest**

Ulike diagnostiske tester, avbildningsprosedyrer eller kriteriesett som er benyttet for å utrede og diagnostisere kjeveledds-dysfunksjon.

## **Referansestandard eller sammenligningstest**

Siden det ikke finnes en omforent referansestandard, eller gullstandard, for diagnostisering av kjeveledds-dysfunksjoner, har vi valgt å inkludere alle studier som sammenlignet ulike diagnostiske tester uavhengig av hvilken test som var definert som referansestandard i hver enkelt oversikt eller primærstudie. Sammenligningstesten må imidlertid være brukt på de samme personene/pasientmaterialet som indekstesten.

## **Utfall**

Fra reliabilitetstester har vi trukket ut informasjon om inter-rater og intra-rater reliabilitet (kappastatistikk). Fra valideringsstudier hvor komparator eller referansestandard er benyttet har vi trukket ut informasjon om sensitivitet, spesifisitet, positiv og negativ prediktiv verdi og sannsynlighetsratioer. I mangel av enhetlig referansestandard vil vi ikke kunne avgjøre hvilke tester som er best, men vi kan sammenligne ulike diagnostiske metoder med henblikk på å avgjøre om de fører til ulike resultater.

## **Språk**

Vi vurderte alle artikler der sammendrag var tilgjengelig på engelsk eller et av de skandinaviske språk.

---

## **Eksklusjonskriterier**

---

Systematiske oversikter som vi vurderte til å være av lav metodisk kvalitet.

---

## **Utvelgelse av studier for inklusjon**

---

Første del av seleksjonsprosessen ble gjort på grunnlag av tittel og sammendrag av de artiklene som ble identifisert i søkene, vurdert opp mot seleksjonskriteriene spesifisert ovenfor. Tre prosjektmedarbeidere (VS, KGB og KTD) gikk, i par, gjennom resultatene fra søkene uavhengig av hverandre. Prosjektleder (VS) leste alle titler og sammendrag, mens de to prosjektmedarbeiderne delte søketreffene mellom seg. Dersom lesepåret var enig om inklusjon, eller ved uenighet om inklusjon av en artikkel, ble artikkelen innhentet i fulltekst og vurdert opp mot inklusjons- og eksklusjonskriteriene av to prosjektmedarbeidere som gjorde sine vurderinger uavhengig av hverandre. Ved uenighet om inklusjon ble spørsmålet løst ved diskusjon, eller ved å søke råd hos en tredje prosjektmedarbeider.

---

## Vurdering av metodisk kvalitet

---

Alle systematiske oversikter som møtte inklusjonskriteriene ble kritisk vurdert med tanke på metodisk kvalitet, av tre personer (VS, KGB og KTD) uavhengig av hverandre. Ved uenighet eller uklarheter diskuterte vi oss frem til enighet.

Metodisk kvalitet av inkluderte systematiske oversikter ble vurdert ved hjelp av sjekkliste (modifisert utgave av Kunnskapssenterets sjekkliste for systematiske oversikter). Sjekklisten inneholder åtte hovedspørsmål, som alle besvares ”ja”, ”uklart” eller ”nei”;

1. Er kriteriene for inklusjon av studier klare?
2. Ble det utført et tilfredsstillende litteratursøk?
3. Ble det sikret mot systematiske skjevheter (bias) ved seleksjon av studier?
4. Ble risiko for systematiske feil (bias) på de inkluderte studiene vurdert ved bruk av relevante kriterier?
5. Er resultatene fra studiene sammenfattet på en tilfredsstillende måte?
6. Er metodene som ble brukt da resultatene ble sammenfattet, klart beskrevet?
7. Er graden av heterogenitet beskrevet?
8. Er forfatterens konklusjoner støttet av data og/eller analysen som er rapportert i oversikten?

Hvert av punktene inneholder i tillegg flere underpunkter (Vedlegg 2)

For hver systematiske oversikt foretok vi en samlet vurdering og kategoriserte kvaliteten som ”høy”, ”moderat” eller ”lav/mangelfull”, med følgende til grunn:

*Høy kvalitet:* Brukes hvis alle eller de fleste kriteriene fra sjekklisten er oppfylt. Dersom noen av kriteriene ikke er oppfylt, må det være veldig lite sannsynlig at studiens konklusjon blir påvirket.

*Moderat kvalitet:* Brukes hvis noen av kriteriene fra sjekklisten ikke er oppfylt og/eller der kriteriene ikke er tilfredsstillende beskrevet. Samlet vurdering tilsier at det er lite sannsynlig at studiens konklusjon påvirkes.

*Lav/mangelfull:* Brukes hvis få eller ingen kriterier i sjekklisten er oppfylt og/eller ikke er tilfredsstillende beskrevet. Samlet vurdering tilsier at det er sannsynlig at studiens konklusjon kan forandres.

---

## Dataekstraksjon og analyser

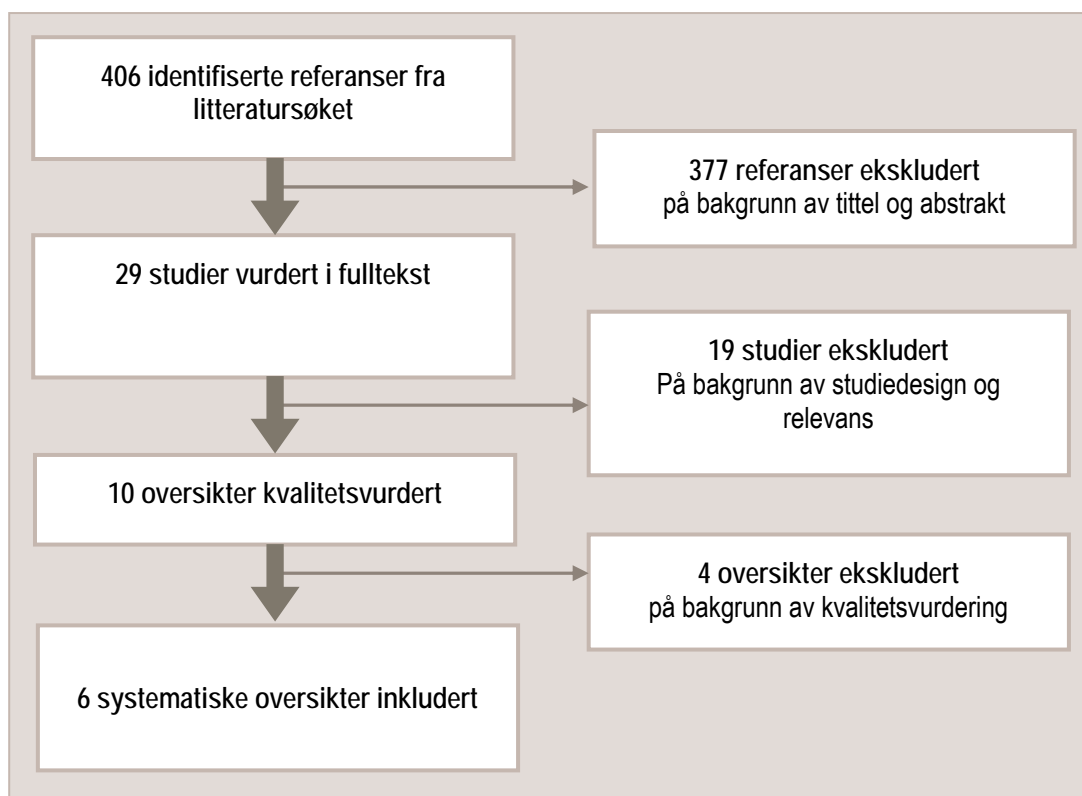
---

Prosjektleder (VS) og prosjektmedarbeidere (KGB og KTD) hentet ut relevante data fra hver av studiene som inngikk i de systematiske oversiktene, og beskrev dette i tabeller. Vi ekstraherte følgende data fra de inkluderte studiene: Tittel, forfattere, detaljer om publikasjonen, land, populasjon, type intervensjon, sammenlikninger, utfallsmål og hovedresultater. En prosjektmedarbeider gikk gjennom beskrivelsene for å sikre at all relevant informasjon var kommet med. Resultatene om diagnostisk nøyaktighet er oppsummert etter hvilken type test

(indekstest) som evalueres, og fremstilt som beskrivende synteser, i tabeller og deskriptive analyser. Forest-plott er benyttet for parvis visning av estimatene, samt 95 % konfidensintervall der disse er oppgitt, for sensitivitet og spesifisitet for hver enkeltstudie. Forest-plottene er ment å gi et visuelt inntrykk av variasjonen i resultatene mellom de ulike studiene, og er en indikasjon på presisjonen i de diagnostiske målene, uten at dataene er slått sammen i en meta-analyse.

# Resultater

Vi fant 406 unike referanser i søket etter systematiske oversikter. Vi gjennomgikk tittel og sammendrag til alle referansene og innhentet 29 artikler i fulltekst (Figur 2). Av disse ekskluderte vi 19 artikler som følge av at studiene ikke var systematiske oversikter, eller ikke omhandlet diagnostikk av TMD (Vedlegg 3). Ti oversikter ble kvalitetsvurdert, hvorav fire ble ekskludert som følge av lav metodisk kvalitet (Vedlegg 3), slik at vi til slutt inkluderte seks systematiske oversikter.



**Figur 2:** Flytskjema over utvelgelsen av studier for inklusjon.



## Beskrivelse av inkluderte systematiske oversikter

Tabell 1 gir en oversikt over de seks inkluderte systematiske oversiktene. Disse var publisert i tidsrommet 2006-2012 i internasjonale vitenskapelige tidsskrifter. Oversiktene kommer fra Sverige (n=2) (22;23), USA (n=2)(24;25), Canada (n=1) (14) og Kina (n=1) (26). Totalt omfatter oversiktene 67 enkeltstudier (Vedlegg 4) med til sammen ca. 5400 personer som har gjennomgått både en indekstest og en referansetest. Indekstester som evalueres i disse oversiktene er EMG, kliniske tester, MR og ultralyd (Tabell 1). Som referansestandard var MR oftest benyttet.

**Tabell 1.** Oversikt over inkluderte systematiske oversikter.

| Forfatter/<br>År          | Antall<br>primærstu-<br>dier<br>(deltakere) | Populasjon<br>(diagnostisk kategori)  | Test som<br>evalueres<br>(indekstest)            | Komparator<br>(referansetest)                             |
|---------------------------|---|---|--|---|
| Al-Saleh<br>2012 (24)     | 2<br>(148)                                  | - Myofascial TMD  | Elektromyografi<br>(EMG)                         | Kliniske tester<br>(sykehistorie/subjektive<br>symptomer) |
| Chaput<br>2012 (14)       | 8<br>(952)                                  | - Diskusforskyvning med eller<br>uten tilbakegang   | Kliniske tester<br>(smerte/ funk-<br>sjon/lyder) | MR  |
| Koh<br>2009 (22)          | 23<br>(ca 2015)                             | - Myofascial TMD<br>- Diskusforskyvning med eller<br>uten tilbakegang<br>- Degenerativ kjeveleddsforand-<br>ring (artrose)                              | Kliniske tester<br>(smerte/ funk-<br>sjon/lyder) | MR  |
| Li<br>2012 (26)           | 15<br>(924)                                 | Diskusforskyvning med eller<br>uten tilbakegang   | Ultralyd   | MR  |
| Limchaichana<br>2006 (23) | 22<br>(ca 430)                              | - Diskusforskyvning med eller<br>uten tilbakegang<br>- Degenerativ kjeveleddsforand-<br>ring (artrose)<br>- Inflammatorisk TMD (syno-<br>vitt/effusjon) | MR   | Biopsier/<br>Frysesnitt<br>(post mortem)                  |
| Reneker<br>2011 (25)      | 7<br>(937)                                  | - Diskusforskyvning med eller<br>uten tilbakegang<br>- Degenerativ kjeveleddsforand-<br>ring<br>- Inflammatoriske TMD-<br>tilstander                    | Klinisk tester<br>(smerte/ funk-<br>sjon/lyder)  | MR<br>Artroskopi  |

MR; Magnetisk Resonanstomografi. TMD; Temporomandibular Disorder (temporomandibulær dysfunksjon).

---

## Vurdering av metodologisk kvalitet

---

Tre av de inkluderte systematiske oversiktene har vi vurdert til å ha høy metodisk kvalitet og tre til å være av moderat kvalitet (Tabell 2). Årsakene til nedgradering av kvalitet skyldes hovedsakelig manglende rapportering av resultater fra vurdering av risiko for systematiske skjevheter i de inkluderte primærstudiene (22), at litteratursøket etter vår mening ikke var tilfredsstillende utført grunnet utilstrekkelig bruk av MeSH-termer (22;23), og at det i en av oversiktene kun var benyttet én database (Pubmed) i litteratursøket (23).

**Tabell 2.** Vurdering av metodologisk kvalitet av inkluderte systematiske oversikter.

| Sjekkliste*  | Al-Saleh<br>(24) | Chaput<br>(14) | Koh<br>(22) | Li<br>(26) | Limchaichana<br>(23) | Reneker<br>(25) |
|--|------------------|----------------|-------------|------------|----------------------|-----------------|
| Er kriteriene for inklusjon av studier klare?  | Ja               | Ja             | Ja          | Ja         | Ja                   | Ja              |
| Ble det utført et tilfredsstillende litteratursøk?   | Ja               | Uklart         | Nei         | Ja         | Nei                  | Uklart          |
| Ble det sikret mot systematiske skjevheter (bias) ved seleksjon av studier?                                | Ja               | Ja             | Ja          | Ja         | Ja                   | Ja              |
| Ble risiko for systematiske feil (bias) på de inkluderte studiene vurdert ved bruk av relevante kriterier? | Ja               | Ja             | Uklart      | Ja         | Ja                   | Ja              |
| Er resultatene fra studiene sammenfattet på en tilfredsstillende måte?                                     | Ja               | Ja             | Ja          | Ja         | Ja                   | Uklart          |
| Er metodene som ble brukt da resultatene ble sammenfattet, klart beskrevet?                                | Uklart           | Ja             | Ja          | Uklart     | Ja                   | Ja              |
| Er graden av heterogenitet beskrevet?  | Ja               | Ja             | Ja          | Ja         | Ja                   | Nei             |
| Er forfatternes konklusjoner støttet av data og/eller analysen som er rapportert i oversikten?             | Ja               | Ja             | Ja          | Ja         | Ja                   | Ja              |
| Samlet kvalitetsvurdering  | Høy              | Høy            | Moderat     | Høy        | Moderat              | Moderat         |

\*Sjekklisten er utformet ved hjelp av: 1. Sjekkliste for systematiske oversikter fra Håndboka 2009, 2. SURE checklist (adapted from the checklist used to prepare SUP-PORT summaries (<http://www.support-collaboration.org/summaries.htm>) and revised to incorporate additional considerations from the AMSTAR checklist), og 3. Sjekkliste for vurdering av en oversiktsartikkel (basert på Guyatt G, Rennie D. The Evidence-Based Medicine Working Group. Users' Guides to the Medical literature. JAMA Archives Journals, AMA Press, 2002. og Critical Appraisal Skills Programme (<http://www.phru.nhs.uk/Pages/PHD/CASP.htm>)).

---

## Diagnostisk nøyaktighet av kliniske tester

---

Tre av de seks inkluderte systematiske oversiktene (14;22;25), med til sammen 30 primærstudier og ca. 3900 deltakere, har undersøkt diagnostisk nøyaktighet av kliniske tester sammenliknet med MR, eller artroskopi, for påvisning av diskusforskyvning, degenerativ kjeveleddsforandring, eller inflammatorisk kjeveleddstilstander. Kjønn- og alderfordeling er oppgitt kun i oversikten til Chaput (14), hvor andelen kvinner var 80 %, og alder varierte fra 14-79 år.

## Litteratursøk

Til tross for at de tre oversiktene (14;22;25) alle undersøkte diagnostisk nøyaktighet av kliniske tester for påvisning av TMD var det flere ulikheter i måten litteratursøkene var foretatt på, og i antall søketreff (Tabell 3). Det var også relativt lite overlapp i inkluderte primærstudier mellom de tre systematiske oversiktene (Vedlegg 4). Vi har derfor valgt å oppsummere resultater fra alle tre oversiktene.

**Tabell 3.** Oversikt over litteratursøk og søkestrategi i tre inkluderte systematiske oversikter som har undersøkt diagnostisk nøyaktighet av kliniske tester for påvisning av TMD.

| Systematisk oversikt     | Chaput 2012 (14)  | Reneker 2011 (25)   | Koh 2009 (22)   |
|--------------------------|---|---|---|
| Benyttede databaser      | Medline<br>Embase   | Medline<br>Cinahl<br>SportDiscus<br>+ Manuelt søk i referanselister   | PubMed<br>Cochrane Library<br>+Manuelt søk i referanselister  |
| Tidsspenn<br>Søk/søkeord | Januar 1994 til oktober 2009<br>Medline: (((tmj)) AND ((“diagnosis”[subheading] OR “diagnosis”[All Fields] OR diagnos*)) AND ((mri)) AND ((sensitivity specificity OR validity OR predictive value of testes OR validation)).   | Ikke oppgitt<br>MeSH-termer: “temporomandibular joint disorders” OR temporomandibular joint dysfunction syndrome” OR temporomandibular joint” AND “diagnosis, differential” OR “diagnosis” AND “sensitivity and specificity” AND “physical examination”                   | 1988 til og med 2007<br>MeSH-termer: “temporomandibular joint” AND “magnetic resonance imaging”   |
| Avgrensning av søk       | Språk: Engelsk<br>Humanstudier  | Språk: Engelsk  | Språk: Engelsk<br>Humanstudier, Alder ≥19 år, Tilgjengelig sammendrag, Klinisk studie, meta-analyse, praktiske retningslinjer, randomisert kontrollert studie, oversiktsartikkel, eller komparativ studie                   |
| Antall søketreff         | 4229  | 131   | 219   |
| Vurdert i fulltekst      | 36  | 24  | 82  |
| Antall inkl artikler     | 8   | 7   | 23  |
| Inklusjons-kriterier     | 1) Benyttet kliniske tester, enten individuelt eller på klusternivå.<br>2) Benyttet MRI som referansestandard.<br>3) Kasus-kontrollstudie eller prospektiv kohortstudie.<br>4) Personer med TMD, intern kjeveledds-forandring med/uten tilbakegang.<br>5) Alder ≥ 14 år | 1) Vurdering av individuelt klinisk mål for TMD, spesielt leddlyder, begrensning i leddbevegelse, muskel- og leddsmerte.<br>2) Rapportert diagnostisk nøyaktighet av disse målene, og<br>3) En akseptabel referanse-standard for sammenlikning (MRI eller led-dartoskopi) | 1) Beskrivelse av klinisk diagnose eller kliniske funn og en MRI-diagnose/funn ved vurdering av degenerativ eller inflammatorisk TMD.<br>2) Beskrivelse av forhold mellom klinisk diagnose/ funn og MRI diagnose/funn.      |
| Eksklusjons-kriterier    | Ikke oppgitt  | Studier:<br>- som bare målte resultater av kliniske funn fra test-klustere,<br>- hvor de kliniske testene ikke var tilstrekkelig forklart.  | Studier som:<br>- rapporterte kliniske funn om muskeltilstander, ikke TMD,- rapporterte laboratoriefunn og ikke kliniske funn el. MRI-funn,<br>- undersøkte tumorer, trauma, og synovial kondromatose, - var kasusrapporter |

## Vurdering av risiko for metodiske skjevheter

Vurdering av intern validitet/risiko for metodiske skjevheter i de inkluderte primærstudiene ble i to av oversiktene (14;25) foretatt ved hjelp av QUADAS-sjekkliste (13) (14 kriterier, skala 0-14, hvor 14 er best), mens det i én oversikt (22) ble benyttet en modifisert QUADAS-sjekkliste. *Chaput* (14) inkluderte åtte primærstudier (27-34), alle prospektive kohortstudier av høy kvalitet, dvs. at studiene var lite utsatt (hadde lav risiko) for metodiske skjevheter, definert som QUADAS-skåre ≥ 9. *Reneker* (25) inkluderte syv studier (30;34-39) og vurderte artikler med QUADAS-skåre ≥ 10 til å ha høy kvalitet, og < 10 til å ha lav kvalitet. Tre av de

syv inkluderte primærstudiene i denne oversikten ble vurdert til å være av høy metodologisk kvalitet (30;36;37). Koh (22) inkluderte 23 studier (27;28;31;32;34;37;40-56), men i denne oversikten ble QUADAS-skårene ikke rapportert. Til tross for høy kvalitet i studiene inkludert i oversikten til Chaput (14), var det i følge oversiktsforfatterne to faktorer som likevel trakk kvaliteten noe ned, i det fem studier ikke oppga tidsspennet mellom testene, og at blinding av testere ikke var enhetlig rapportert. I oversikten til Reneker (25) var lav kvalitet i studiene hovedsakelig grunnet uklar beskrivelse av indekstesten og/eller referansestandard, og mangel på beskrivelse av blinding.

## **Resultater om diagnostisk nøyaktighet**

Diagnostisk nøyaktighet av kliniske tester er i oversiktene rapportert for enkelttester (hhv. kjeveleddslyder; klikk og krepitasjon, kjeveleddsbevegelser og smerte), og for grupper av tester (test-kluster; kumulativ effekt av enkelttester), for påvisning av hhv. diskusforskyvning med tilbakegang, diskusforskyvning uten tilbakegang, degenerativ kjeveleddsforandring og inflammatoriske tilstander. Oversiktsforfatterne har imidlertid benyttet noe ulik tilnærming ved utregning av diagnostisk nøyaktighet, idet Chaput (14) har lagt antall inkluderte kjeveledd til grunn, mens Koh (22;57) og Reneker (25) har benyttet antall personer. De kliniske testene er sammenliknet med resultater fra MR-bilder i sagittalplan og/eller frontalplan (koronalplan), med unntak av én studie (35) i oversikten til Reneker (25), som benyttet artroskopi som referansetest.

Kjeveleddsbevegelse ble definert som begrenset bevegelighet (både aktiv og passiv), begrensnig i kondylar translasjon, og ufrivillig bevegelser i forbindelse med åpning og lukking av munnen. Begrensning i aktiv åpning av munnen (maksimal gapeevne) var definert som mindre enn 40 mm. Passiv åpning av munnen ble oppnådd ved å påføre et lett trykk i 15 sekunder, og deretter måle interincisal avstand korrigert for vertikalt overlapp. Begrenset protrusjon var definert som mindre enn 5 mm. Krepitasjon ble definert som en vedvarende lyd gjennom gapebevegelsen ("like grating or grinding"), mens klikk var definert som en distinkt, kort lyd med tydelig start og slutt.

## ***Bruk av kliniske tester for påvisning av diskusforskyvning med tilbakegang***

### ***Kjeveleddslyder***

Klikk versus MR for deteksjon av diskusforskyvning med tilbakegang ble rapportert i totalt fem enkeltstudier (27;30;34;37;38) i de tre oversiktene, med til sammen omlag 500 deltakere. Resultatene viser at punkttestimatene for sensitivitet varierte fra 0,46 til 0,89, og for spesifisitet fra 0,20 til 0,95 (Figur 3). LR<sup>+</sup> varierte fra lavest 1,12 til høyest 15,20 (resiprokalt klikk), mens LR<sup>-</sup> varierte fra høyest 0,92 til lavest 0,25. For krepitasjon var sensitiviteten svært lav (0,02-0,11) og spesifisiteten varierende fra 0,64 til 0,91, rapportert fra to enkeltstudier med til sammen 313 deltakere (Figur 3). LR<sup>+</sup> varierte fra lavest 0,21 til høyest 0,88, mens LR<sup>-</sup> varierte fra høyest 1,39 til lavest 1,02 (Vedlegg 5; Tabell 6).

### *Smerte*

Leddsmarter: Smerter angitt ved palpasjon, kjeveposisjoner, og provokasjonstester.

Syv enkeltstudier med til sammen nesten 800 deltakere sammenlignet den diagnostiske nøyaktigheten av smerte versus MR for deteksjon av diskusforskyvning med tilbakegang. Sensitivitet varierte fra 0,05 til 1,0, og spesifisitet fra 0,12 til 0,93 (Figur 3). LR<sup>+</sup> varierte fra lavest 0,56 til høyest 1,99, og LR<sup>-</sup> varierte fra høyest 1,83 til lavest 0,00 (Vedlegg 5; Tabell 6).

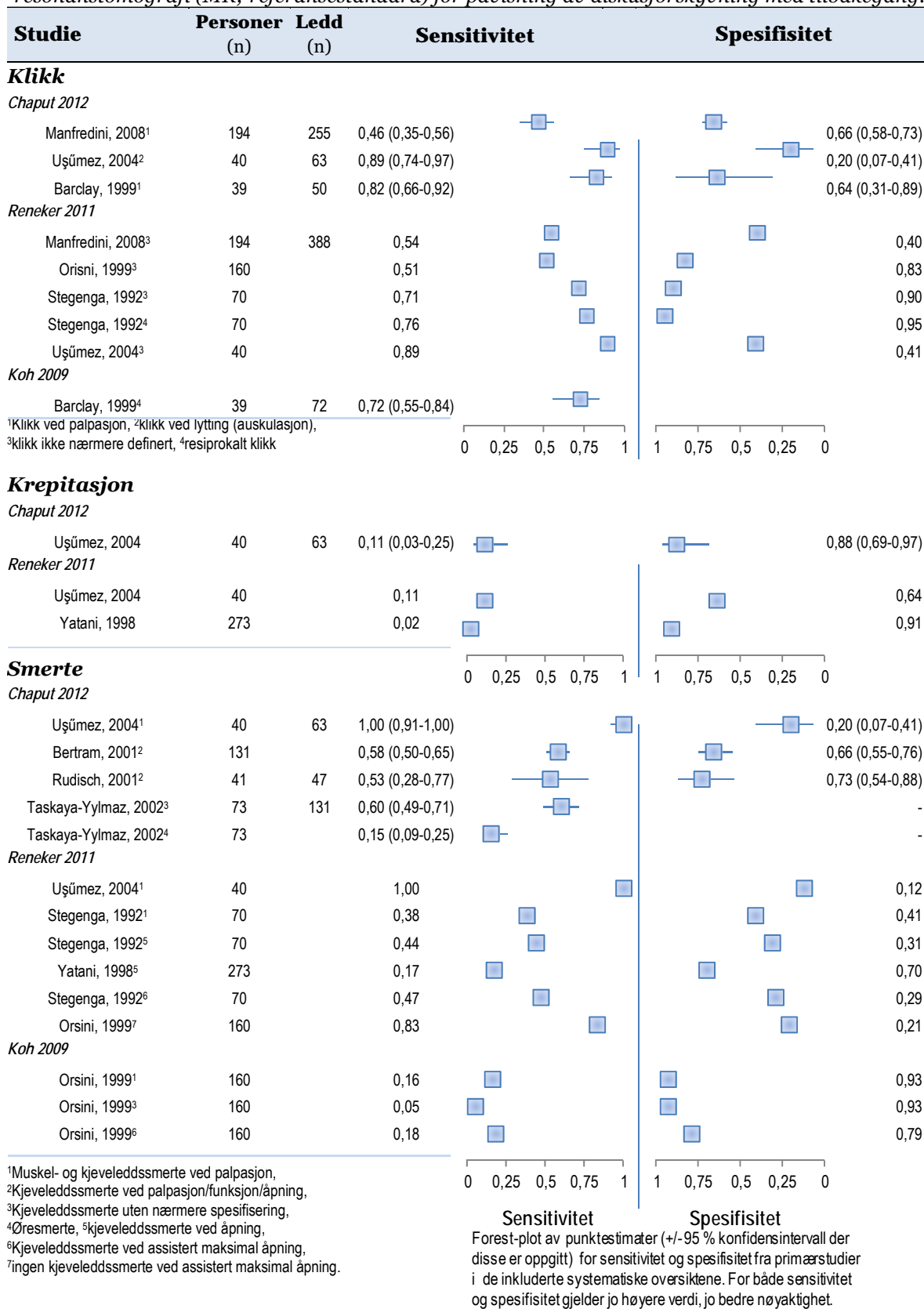
### *Kjeveleddsbevegelser*

Fire enkeltstudier sammenlignet den diagnostiske nøyaktigheten av ulike kjeveleddsbevegelser versus MR for deteksjon av diskusforskyvning med tilbakegang. Disse studiene hadde til sammen over 500 deltakere. Sensitivitet varierte fra 0,04 til 0,92, og spesifisitet fra 0,08 til 0,90. LR<sup>+</sup> varierte fra lavest 0,17 til høyest 2,59, og LR<sup>-</sup> varierte fra høyest 2,95 til lavest 0,26.

### *Test-kluster*

Chaput (14) rapporterer resultater fra tre enkeltstudier som hadde benyttet test-kluster for diagnostisering av diskusforskyvning med tilbakegang, sammenliknet med MR. Alle testene i klusteret måtte være positive for å kunne påvise diagnosen. Det var stor variasjon både i sensitivitet (fra 0,05 til 0,97) og spesifisitet (fra 0,26 til 0,88). LR<sup>+</sup> varierte fra lavest 0,42 til høyest 3,05, og LR<sup>-</sup> varierte fra høyest 1,08 til lavest 0,12.

**Figur 3.** Diagnostisk nøyaktighet (sensitivitet og spesifisitet) med forest-plot av ulike kliniske enkelttester (indekstest: hhv. klikk, krepitasjon og smerte) sammenliknet med magnetisk resonanstomografi (MR; referansestandard) for påvisning av diskusforskyvning med tilbakegang.



## ***Bruk av kliniske tester for påvisning av diskusforskyvning uten tilbakegang***

### *Kjeveleddslyder*

Resultater fra seks enkeltstudier med til sammen 776 deltakere som sammenliknet kjeveleddslyder med MR for påvisning av diskusforskyvning uten tilbakegang, viste for klikk en variasjon i sensitivitet fra 0,04 til 0,72, og for spesifisitet fra 0,14 til 0,76 (Figur 4). LR<sup>+</sup> varierte fra lavest 0,08 til høyest 1,53, og LR<sup>-</sup> varierte fra høyest 4,93 til lavest 0,70 (Vedlegg 5; Tabell 7). For krepitasjon varierte sensitivitet fra 0,08 til 0,71, mens verdiene for spesifisitet var mer heterogene (0,88-1,0; Figur 4). LR<sup>+</sup> var relativt høy, varierende fra lavest 5,30 til høyest 6,36. LR<sup>-</sup> varierte fra høyest 0,86 til lavest 0,33 (Vedlegg 5; Tabell 7).

### *Smerte*

Syv enkeltstudier med til sammen 788 deltakere har sammenliknet smerte med MR for påvisning av diskusforskyvning uten tilbakegang. Sensitivitet varierte fra 0,15 til 1,0, og spesifisitet fra 0,08 til 0,96 (Figur 4). LR<sup>+</sup> varierte fra lavest 1,09 til høyest 6,02 (kjeveleddsmerte ved assistert maksimal åpning), og LR<sup>-</sup> varierte fra høyest 0,51 til lavest 0,00 (Vedlegg 5; Tabell 7).

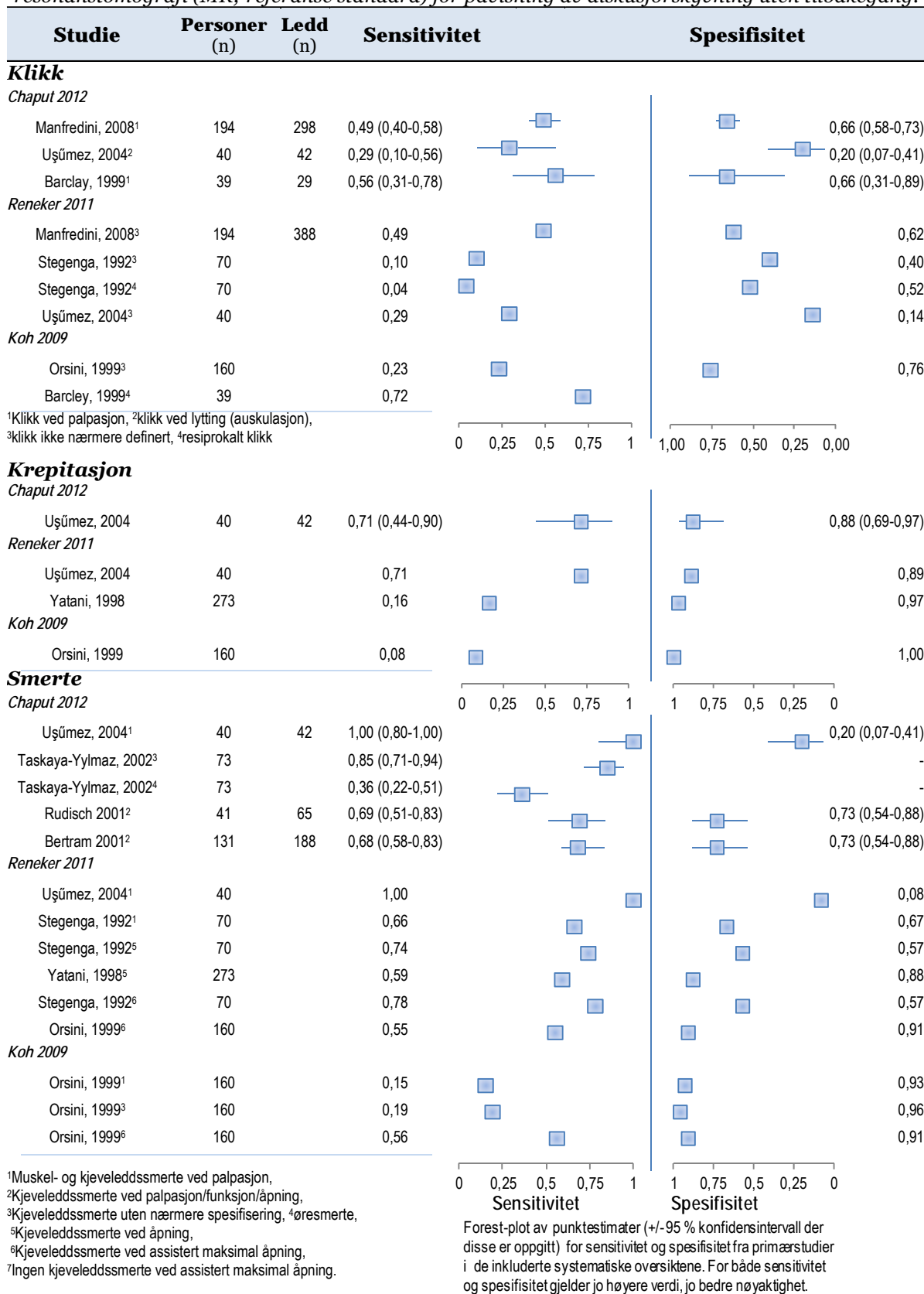
### *Kjeveleddsbevegelser*

To av oversiktene rapporterer resultater fra fire enkeltstudier med til sammen 543 deltakere som har sammenliknet ulike kjeveleddsbevegelser med MR for påvisning av diskusforskyvning uten tilbakegang. Sensitivitet varierer fra 0,14 til 0,78 og spesifisitet fra 0,08 til 0,97. LR<sup>+</sup> varierte fra lavest 0,33 til høyest 9,00 (defleksjon), og LR<sup>-</sup> varierte fra høyest 8,13 til lavest 0,26.

### *Test-kluster*

Sensitivitet varierte fra 0,06 til 0,78 og spesifisitet fra 0,28 til 1,0. LR<sup>+</sup> varierte fra lavest 0,08 til høyest 6,33 (krepitasjon/defleksjon/smerte/begrenset gapeevne), og LR<sup>-</sup> varierte fra høyest 3,36 til lavest 0,27.

**Figur 4.** Diagnostisk nøyaktighet (sensitivitet og spesifisitet) med forest-plot av ulike kliniske enkelttester (indekstest; hhv. klikk, krepitasjon og smerte) sammenliknet med magnetisk resonanstomografi (MR; referanse standard) for påvisning av diskusforskyvning uten tilbakegang.

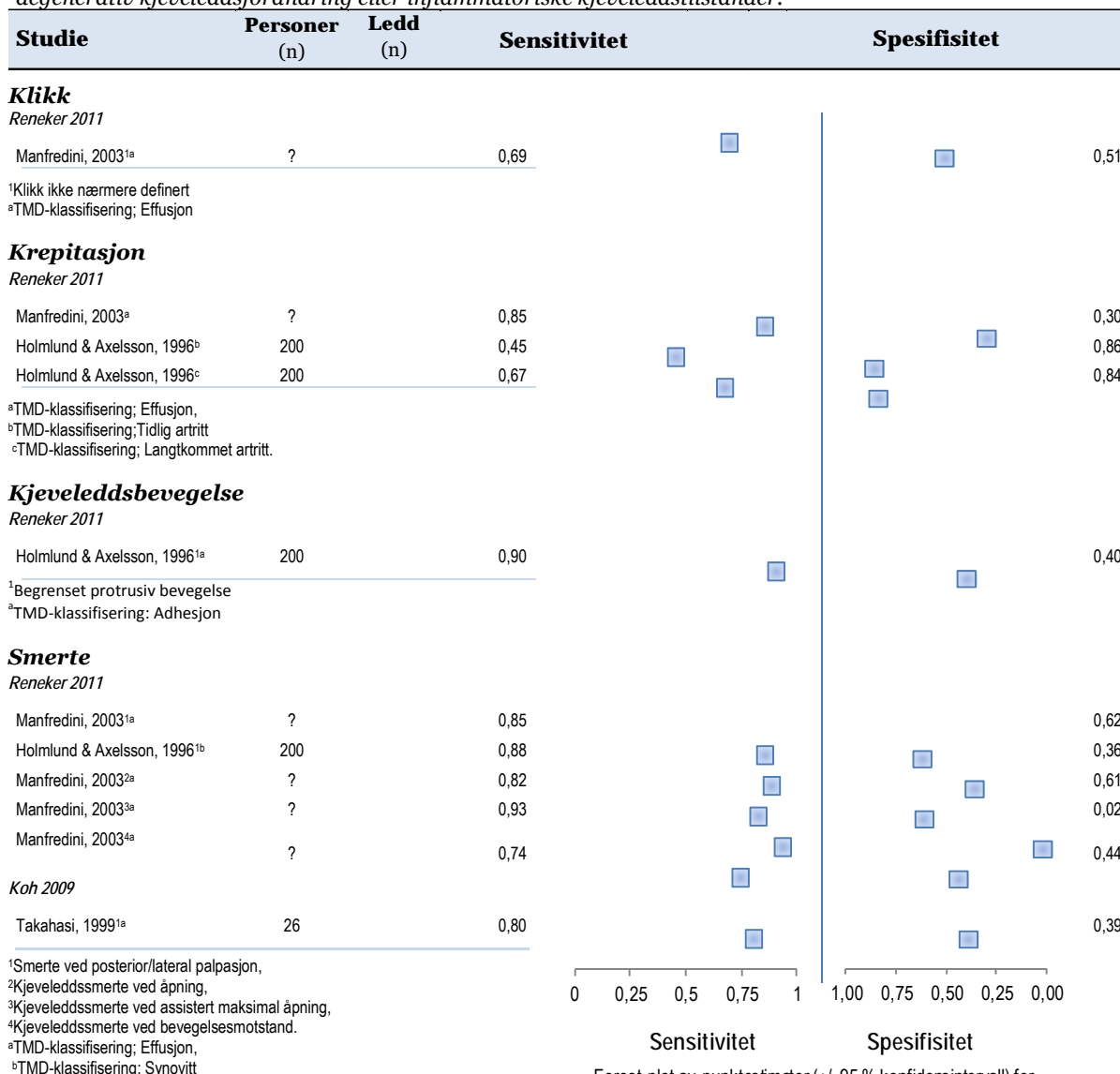




## Bruk av kliniske tester for påvisning av degenerativ kjeveleddsforandring eller inflammatorisk TMD

Oversiktene til Koh (22) og Reneker (25) rapporterer resultater fra tre enkeltstudier (35;36;56) med til sammen over 200 deltakere som har sammenlignet diagnostisk nøyaktighet av enkelttester versus MR eller artroskopi (35) for deteksjon av enten artritt, synovitt, leddeffusjon, eller adhesjon. Resultater for sensitivitet og spesifisitet er oppsummert i Figur 5 og Vedlegg 5; Tabell 8)

**Figur 5.** Diagnostisk nøyaktighet (sensitivitet og spesifisitet) med forest-plot av ulike kliniske enkelttester (indekstest) sammenliknet med magnetisk resonanstomografi (MR) eller artroskopi for påvisning av degenerativ kjeveleddsforandring eller inflammatoriske kjeveleddstilstander.



Forest-plot av punktestimater (+/- 95 % konfidensintervall) for sensitivitet og spesifisitet fra primærstudier i de inkluderte systematiske oversiktene. Forbåde sensitivitet og spesifisitet gjelder jo høyere verdi, jo bedre nøyaktighet.

### **Oppsummering av resultater om kliniske tester**

Tre systematiske oversikter (14;22;25), med til sammen 30 primærstudier og ca. 3900 deltagere, har undersøkt diagnostisk nøyaktighet av kliniske tester sammenliknet med MR, eller artroskopi, for påvisning av diskusforskyvning, degenerativ kjeveleddsforandring, eller inflammatoriske TMD-tilstander.

De oppsummerte resultatene om diagnostisk nøyaktighet viser en gjennomgående tendens til at det er en sammenheng mellom høy sensitivitet og lav spesifisitet, og omvendt.

For *kjeveledslyder* (klikk og krepitasjon) varierer punkttestimatene for sensitivitet fra 0,02 til 0,95, og spesifisitet fra 0,14 til 1,0. Denne variasjonen var uavhengig av type klikk (klikk ved palpasjon, auskulasjon, resoprokalt klikk, eller der klikk ikke er nærmere definert), og av type TMD-tilstand. Sannsynlighetsratioene for å bekrefte eller utelukke TMD ved hjelp av kjeveleddslyder varierte, med få unntak, mellom 0,5 og 2,5 for LR<sup>+</sup> og fra 3,1 til 0,5 for LR<sup>-</sup>. I én studie var LR<sup>+</sup> høyere enn 10 (bruk av resiprokalt klikk for påvisning av diskusforskyvning med tilbakegang).

For *smerte* varierer punkttestimatene for sensitivitet fra 0,05 til 1,0, og spesifisitet fra 0,02 til 0,96. Minst variasjon i sensitiviten var det for påvisning av degenerative eller inflammatoriske TMD-tilstander. Synlighetsratioene ved benyttelse av smerte til å bekrefte eller utelukke TMD var, med få unntak, hhv. lavere enn 2,5 (LR<sup>+</sup>) og høyere enn 0,2 (LR<sup>-</sup>). Høyest positiv sannsynlighetsratio (LR<sup>+</sup>) ble rapportert for kjeveleddssmerte ved assistert maksimal åpning for påvisning av diskusforskyvning uten tilbakegang (LR<sup>+</sup> >6).

For *kjeveleddsbevegelse* varierer punkttestimatene for sensitivitet fra 0,05 til 0,92, og spesifisitet fra 0,08 til 0,97. Synlighetsratioene ved benyttelse av kjeveleddsbevegelser for å bekrefte eller utelukke TMD, var stort sett lavere enn 2,5 (LR<sup>+</sup>) og høyere enn 0,2 (LR<sup>-</sup>). Høyest positiv sannsynlighetsratio ble rapportert for defleksjon ved påvisning av diskusforskyvning uten tilbakegang (LR<sup>+</sup> fra 6,3 til 9,0).

For *testkluster* varierer punkttestimatene for sensitivitet fra 0,05 til 0,97 og spesifisitet fra 0,26 til 1,0. Synlighetsratioene ved benyttelse av testklusters for å bekrefte eller utelukke TMD, var stort sett lavere enn 2,5 (LR<sup>+</sup>) og høyere enn 0,2 (LR<sup>-</sup>). Høyest positiv sannsynlighetsratio ble rapportert for testkluster bestående av krepitasjon, defleksjon, smerte og begrenset gapeevne for påvisning av diskusforskyvning uten tilbakegang (LR<sup>+</sup> 6,33).

---

## Diagnostisk nøyaktighet av MR

---

Én inkludert systematisk oversikt (23), med til sammen 22 primærstudier og ca. 430 deltakere, har undersøkt diagnostisk nøyaktighet av magnetisk resonanstomografi (MR) sammenlignet med biopsi eller frysesenitt for påvisning av diskusforskyvning, degenerative eller inflammatoriske lidelser i temporomandibulærleddet. Kjønn- og alderfordeling var ikke oppgitt.

### Litteratursøk og vurdering av risiko for metodiske skjevheter

Limchaichana og medarbeidere (23) gjennomførte et enkelt søk etter engelskspråklig litteratur i PubMed (1988 til juli 2005) avgrenset av MeSH-termene "Temporomandibular joint" og "Magnetic resonance imaging". Forfatterne gjennomgikk i tillegg referanselistene til 36 utvalgte publikasjoner. Søketreff ble gjennomgått og vurdert opp mot seleksjonskriteriene av tre uavhengige forfattere. Kvaliteten til inkluderte studier ble vurdert ved hjelp av en sjekkliste som inneholdt elementer som:

- Er spekteret av pasientene som inngår i studien representativt for de pasienter som vil gjennomgå undersøkelsen i klinisk praksis?
- Ble MR-undersøkelsen foretatt blindet og uavhengig av referansetesten?
- Ble referansetesten foretatt blindet og uavhengig av MR-undersøkelsen?
- Ble metodene beskrevet detaljert nok til å tillate replikasjon?

Inkluderte studier ble gruppert i tre evidensnivåer – lav, moderat og høy. Resultatene ble beskrevet for hver enkelt studie i tabeller, og forfatterne gjorde ingen egne meta-analyser. Det elektroniske litteratursøket ga 462 unike treff som ledet til inklusjon av 22 studier. Ti av de 22 studiene var ifølge oversiktsforfatterne assosiert med så store metodiske feil og mangler at resultatene ikke ble presentert. De resterende tolv studiene ble vurdert å ha moderat evidensgrad, og resultatene av disse ble gruppert etter type sammenligning og presentert i tabeller. Sju studier presenterte reliabilitetsdata, mens fem undersøkte diagnostisk nøyaktighet av MR sammenlignet med biopsi eller frysesenitt (Tabell 4).

**Tabell 4.** Inkluderte studier i oversikten til Limchaichana og medarbeidere (23) om diagnostisk nøyaktighet av MR versus biopsi eller frysesenitt (post mortem).

| Førsteforfatter/år | Påvisning av                                      | MR-teknikk og sammenligning                             |
|--------------------|---|---|
| Westesson 1987     | Diskusposisjon, diskusmorfologi, ossøse endringer | T1 sagittal versus sagittale frysesenitt                |
| Katzberg 1988      | Diskusposisjon, ossøse endringer                  | T1 koronal versus koronale frysesenitt                  |
| Schwaighofer 1990  | Diskusposisjon                                    | T1 koronal + sagittal versus koronale frysesenitt       |
| Tasaki 1993        | Diskusposisjon, diskusmorfologi, ossøse endringer | PD + T2, koronal + sagittal versus koronale frysesenitt |
| Larheim 1999       | Benmargsforandinger                               | PD, T1 eller T2, koronal + sagittal versus biopsi       |

T1 (T1-vektet), T2 (T2-vektet), PD (Protontetthet eller protonvektet). MR utført med 1,5-Tesla-apparatur.

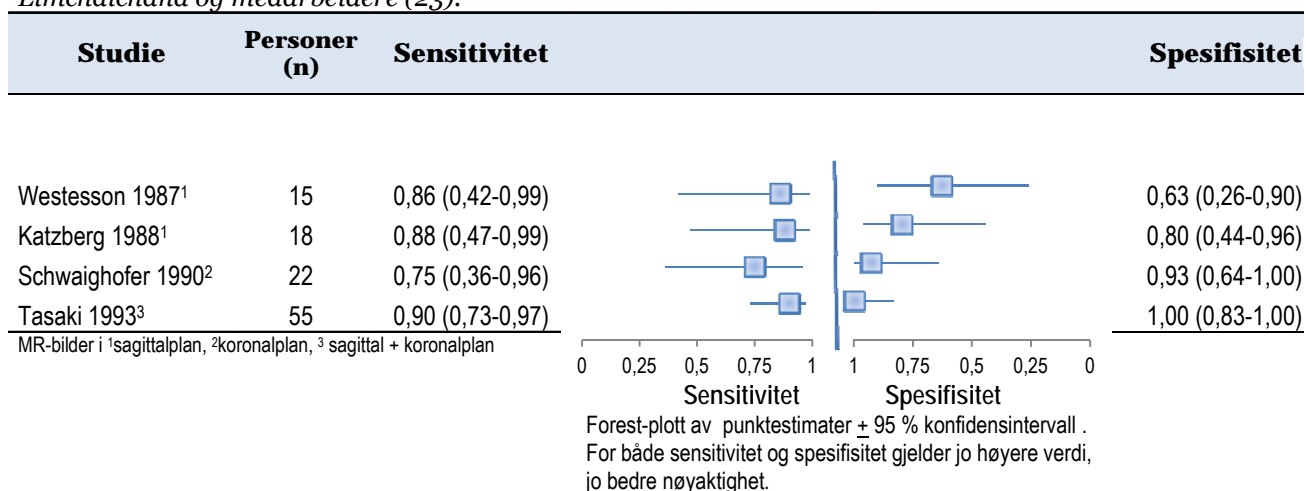
## Resultater om diagnostisk nøyaktighet

### Bruk av MR for evaluering av diskusposisjon i temporomandibulærleddet

Oversiktsforfatterne inkluderte fire studier (58-61) som undersøkte reliabilitet (kappa-verdier). Én studie viste godt samsvar mellom undersøkelser foretatt av samme observatør (Kappa ~0,85). Målt på tvers av ulike observatører var variasjonen større, men kappa økte etter kalibrering av observatører og ved bruk av bilder med bedre kvalitet. Enigheten var best for påvisning av diskusforskyvning uten tilbakegang (Kappa=0,91), og dårligst for påvisning av lett fremoverforskyvning av diskusen (Kappa=0,19).

Oversiktsforfatterne inkluderte fire studier (62-65), av moderat evidensnivå, som sammenholdt resultater av MR-avbildning med postmortem frysesnitt for vurdering av diagnostisk nøyaktighet (Tabell 8). Studiene var heterogene med henblikk på bruk av diagnostiske kriterier, bakenforliggende prevalens (spekter av pasienter) og type avbildning, og oversiktsforfatterne valgte å trekke konklusjoner på tvers av de inkluderte studiene. For påvisning av diskusforskyvning varierte punkttestimatene for sensitivitet fra 0,75 til 0,90 og for spesifisitet fra 0,63 til 1,0 (Figur 6). Punkttestimatene for LR<sup>+</sup> var > 2,3, og for LR<sup>-</sup> < 0,2 (Vedlegg 5; Tabell 9).

**Figur 6.** Diagnostisk nøyaktighet av magnetisk resonanstomografi (MR) sammenliknet med frysesnitt foretatt post mortem for påvisning av diskusforskyvning. Enkeltstudier hentet fra oversikten til Limchaichana og medarbeidere (23).



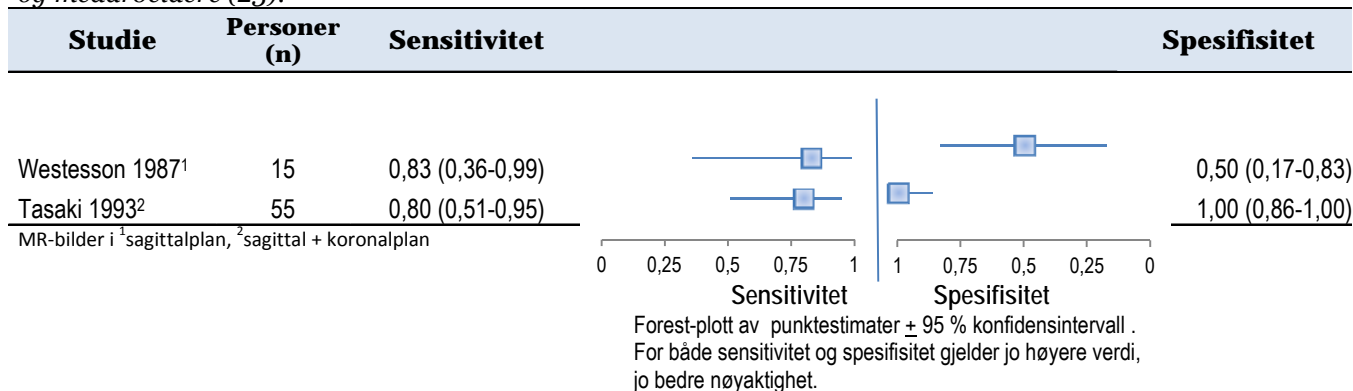
### Bruk av MR for evaluering av diskusmorfologi

Oversiktsforfatterne inkluderte to studier (59;61) som rapporterte kappa-verdier. Én studie viste begrenset samsvar mellom undersøkelser foretatt av samme observatør (Kappa ~0,65). Begge studiene viste begrenset samsvar mellom bedømmingene til ulike observatører med kappa-verdier rundt 0,4.

Oversiktsforfatterne inkluderte to studier (64;66), av moderat evidensnivå, som sammenliknet MR med post mortem frysesnitt for påvisning av diskusdeformitet. Den ene studien sammenliknet sagittale T1-vektede MR-bilder som kunne påvise deformert diskus med like stor nøyaktighet som frysesnitt, mens den andre studien sammenliknet protonvektede og T2-

vektede MR-bilder tatt i koronal- og sagittalplanet versus frysesnitt i koronal- og sagittalplanet. Begge studiene viste sensitivitet rundt 0,8, men med svært varierende spesifisitet (Figur 7). Punkttestimatene for LR<sup>+</sup> var > 1,7, og for LR<sup>-</sup> < 0,3 (Vedlegg 5; Tabell 10).

**Figur 7.** Diagnostisk nøyaktighet av magnetisk resonanstomografi (MR) sammenliknet med frysesnitt foretatt post mortem for påvisning av diskusdeformitet. Enkeltstudier hentet fra oversikten til Limchaichana og medarbeidere (23).

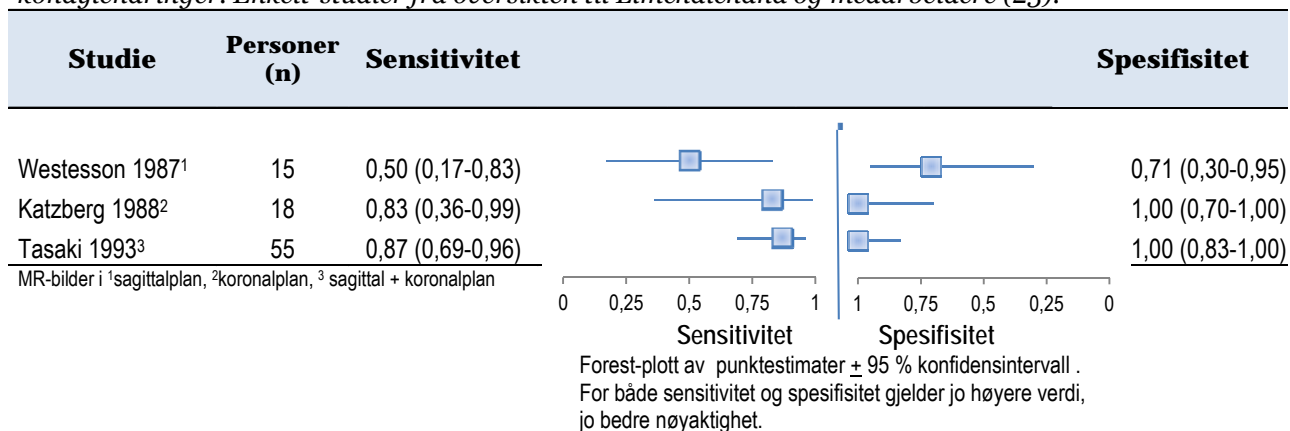


### Bruk av MR for evaluering av ossøs kjeveleddsforandring

Oversiktsforfatterne inkluderte én studie (61) som rapporterte kappa-verdier ved bruk av T1-vektede sagittale MR-bilder for å påvise endringer i kondylens form. Overensstemmelsen var lav, men økte noe med kvaliteten på MR-bildene, og var for høykvalitetsbilder 0,56 for to radiologer, og 0,42 for to tannleger.

Videre inkluderte de tre studier med moderat evidensnivå som undersøkte diagnostisk nøyaktighet. For påvisning av ossøse kondylendringer i kjeveleddet varierte punkttestimatene for sensitivitet fra 0,50 til 0,87 og for spesifisitet fra 0,71 til 1,0 (Figur 8). Punkttestimatene for LR<sup>+</sup> var > 1,75, og for LR<sup>-</sup> < 0,7 (Vedlegg 5; Tabell 11).

**Figur 8.** Diagnostisk nøyaktighet av magnetisk resonanstomografi (MR) ved benyttelse av T1-vektede sagittale og/eller koronale MR-bilder sammenliknet med frysesnitt (post mortem) for påvisning av ossøse kondylendringer. Enkelt-studier fra oversikten til Limchaichana og medarbeidere (23).



### *Bruk av MR for evaluering av benmargsforandringer i kjeveleddet*

Oversiktsforfatterne fant ingen studier som rapporterte grad av enighet mellom observatører.

Data om diagnostisk nøyaktighet ble rapportert i én studie (67) som sammenlignet overensstemmelse mellom MR-bilder (T1-protontetthet, eller T2 -vektet) i koronal og sagittalplanet med benmargsbiopsier fra den mandibulære kondylen, tatt under kirurgi. Studien ble vurdert til å ha et moderat evidensnivå, og rapporterte data fra 50 smertefulle kjeveledd, antatt å skyldes interne kjeveleddsforandringer. Biopsier påviste osteonekrose i ni av 50 kjeveledd. Sju av disse ble korrekt påvist ved MR og to ble feilaktig klassifisert som normale, noe som svarer til sensitivitet på 0,78 (95 % KI 0,40-0,96). Av resterende 41 ledd ble seks feilaktig vurdert som positive på bakgrunn av MR-undersøkelsen. Det vil si at spesifisiteten ble beregnet til 0,84 (95 % KI 0,70-0,94). Positiv og negativ prediktiv verdi ble beregnet til henholdsvis 0,54 (95 % KI 0,26-0,80) og 0,95 (95 % KI 0,80-0,99), og positiv og negativ sannsynlighetsratio til hhv. 4,90 og 0,26.

### *Bruk av MR for evaluering av diskusperforering, sammenvoksing eller bloduttredelse i kjeveleddet*

Forfatterne identifiserte ingen studier med moderat eller høyt evidensnivå. Oversiktsforfatterne identifiserte tre studier med lavt evidensnivå, men resultatene av disse ble ikke referert.

---

## **Diagnostisk nøyaktighet av CT**

---

Ingen av de seks inkluderte systematiske oversiktene hadde undersøkt diagnostisk nøyaktighet av CT.

---

## **Diagnostisk nøyaktighet av ultralyd**

---

Én inkludert systematisk oversikt (26), med totalt 15 primærstudier og 924 deltakere, har undersøkt diagnostisk nøyaktighet av ultralyd sammenliknet med MR for påvisning av diskusforskyvning med eller uten tilbakegang i temporomandibulærleddet. Andelen kvinner var 77 %, og alder varierte fra 13-78 år.

### **Litteratursøk og vurdering av risiko for metodiske skjevheter**

*Li og medarbeidere* (26) søkte etter litteratur i tre ulike databaser: Medline (1948 til april 2011), Embase (1980 til april 2011) og en kinesisk biomedisinsk database (1978 til april 2011). Søketermene bestod av en kombinasjon av fritekstord og MeSH-termer som for eksempel "temporomandibular joint", "temporomandibular disorder" og "ultrasonography". Søketreffene ble gjennomgått og vurdert opp mot seleksjonskriteriene av to uavhengige forfattere. Risiko for metodiske skjevheter ble vurdert ved hjelp av en modifisert QUADAS-sjekkliste som inneholdt følgende elementer:

- Er spekteret av pasientene som inngår i studien representativt for de pasienter som vil gjennomgå undersøkelsen i klinisk praksis?

- Ble de to testene (ultralyd og MR) gjennomført så nært i tid at det er lite sannsynlig at tilstanden endrer seg i mellomtiden?
- Ble testene gjennomført uten differensialverifisering?
- Ble MR-undersøkelsen foretatt blindet?
- Ble ultralydundersøkelsen foretatt blindet?

Foruten å beskrive inkluderte studier gjennomførte oversiktsforfatterne meta-analyser og meta-regresjon med henblikk på å utforske om, og eventuelt hvilke, kliniske faktorer som påvirket den diagnostiske nøyaktigheten til ultralydundersøkelsen.

Litteratursøket ledet til 148 unike treff. Blant disse fant oversiktsforfatterne 15 studier med relevans for sammenligningen av ultralyd og MR for påvisning av diskusforskyvning i temporomandibularleddet, 14 kohortestudier og én kasus-kontroll-studie. Oversiktsforfatterne konkluderte med at seks studier var lite utsatt for metodiske skjevheter (lav risiko), seks hadde uklar risiko for metodiske skjevheter mens tre var svært utsatte for metodiske skjevheter (høy risiko). Resultater ble rapportert for ulike tester.

## **Resultater om diagnostisk nøyaktighet**

### *Påvisning av diskusforskyvning, undersøkt med munnen lukket*

Åtte studier (68-75) sammenlignet den diagnostiske nøyaktigheten av ultralyd versus MR for deteksjon av diskusforskyvning med munnen i lukket posisjon (Figur 9). Som det fremkommer i figuren var det store forskjeller mellom ulike studier. Punkttestimatet for sensitivitet varierte eksempelvis fra 0,50 til 0,93, og spesifisitet fra 0,47 til 0,92. Heterogeniteten ble bekreftet i meta-analyser som viste  $I^2$ -verdier mellom 80 % og 90 %. Verdier over 60-70 % regnes gjerne som uttrykk for vesentlig heterogenitet. Meta-regresjon viste at heterogeniteten ikke kunne forklares av type ultralyd (dynamisk eller statisk;  $p=0,67$ ), dimensjon (2D eller 3D;  $p=0,52$ ) eller ultralydbildet av diskus (hypoekko, isoekko eller hyperekko;  $p=0,42$ ). Sensitivitetsanalyser viste at statisk ultralyd hadde noe høyere sensitivitet enn dynamisk. Punkttestimatene for  $LR^+$  var  $> 1,20$ , og for  $LR^- < 0,78$  (Vedlegg 5; Tabell 12).

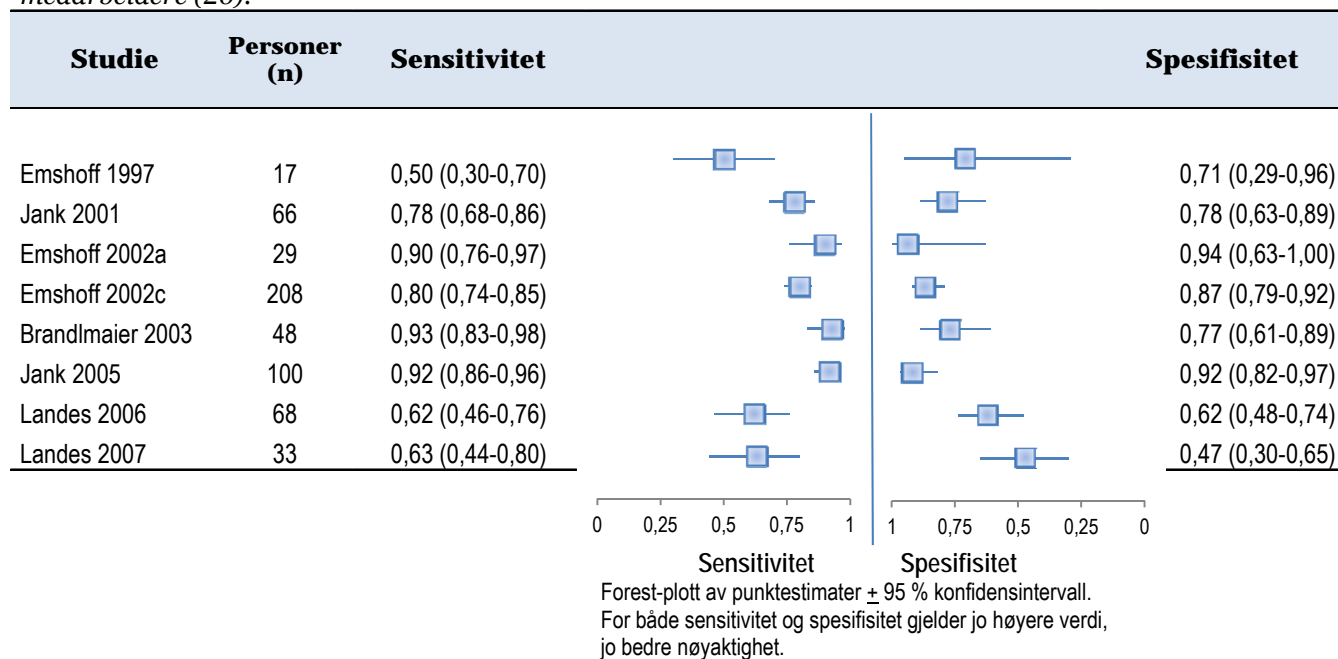
### *Påvisning av intern kjeveleddsforandring, undersøkt med munnen åpen*

Sammenliknet med ultralyd med munnen i lukket posisjon så det ut til at ultralyd med munnen i åpen posisjon kunne assosieres med noe lavere sensitivitet, 0,69 (95 % KI 0,64 -0,74) versus 0,81 (95 % KI 0,77 -0,84) og noe bedre spesifisitet 0,89 (95 % KI 0,86 -0,91) versus 0,78 (95 % KI 0,74 -0,82). Sensitivitetsanalyser viste at dynamisk avbildning hadde noe bedre diagnostiske egenskaper enn statisk avbildning.

### *Diskusforskyvning med versus uten tilbakegang*

Resultatene viste bedre nøyaktighet for diskusforskyvning uten tilbakegang enn for diskusforskyvning med tilbakegang. ROC-analyser resulterte i AUC på henholdsvis 0,97 og 0,83.

**Figur 9.** Diagnostisk nøyaktighet av ultralyd sammenliknet med magnetisk resonanstomografi (MR) for deteksjon av diskusforskyvning med munnen i lukket posisjon. Enkeltstudier i oversikten til Li og medarbeidere (26).



## Diagnostisk nøyaktighet av elektromyografi

Én inkludert systematisk oversikt (24), med to primærstudier (76;77) og 148 deltakere, har undersøkt diagnostisk nøyaktighet av elektromyografi (EMG) for påvisning av myofacial temporomandibulær dysfunksjon. Kjønn- og alderfordeling er oppgitt kun i en av primærstudiene, hvor andelen kvinner var 80 %, og gjennomsnittsalder var 25 år.

### Litteratursøk og vurdering av risiko for metodiske skjevheter

Forfatterne av denne oversikten søkte etter litteratur til og med august 2010, i følgende databaser: Medline, Embase, Evidence-Based Medicine Reviews (inkludert Cochrane databasen), ACP Journal Club, Database of Abstracts of Reviews of Effectiveness, Cochrane Central Register of Controlled Trials, Cochrane Methodology register, Health Technology Assessment, NHS Economic Evaluation Database, Allied and Complementary medicine, Ovid HealthSTAR, Sci Verse Scopus. Søketermene bestod av fritekstord og Mesh-termer som "tempomandibular joint", "temporomandibular disorder", "electromyography" og "EMG", i tillegg til "sensitivity and specificity", "diagnostic accuracy" og "predictive values".

Søketreffene ble gjennomgått og vurdert opp mot seleksjonskriteriene av to uavhengige forfattere. Litteratursøket ledet til 130 unike treff. Blant disse var det til slutt kun to studier (76;77) som tilfredsstilte forfatterens inklusjonskriterier om vurdering av diagnostisk nøyaktighet av EMG for påvisning av TMD, sammenliknet med andre diagnostiske verktøy. Risiko for metodiske skjevheter ble vurdert ved hjelp av QUADAS-sjekkliste (13).



## Resultater om diagnostisk nøyaktighet

Begge studiene som er inkludert i den systematiske oversikten til Al-Saleh og medarbeidere (24) var tverrsnittsstudier hvor resultater fra bruk av EMG ble sammenliknet med resultater fra kliniske undersøkelser for å diagnostisere TMD. Studien inkluderte pasienter som hadde myofasiale smerter, og som ble rekruttert fra universitetsklinikker. Pasienter med intern kjeveleddspatologi eller klikkelyder ble ekskludert. EMG ble registrert ved hjelp av overflateelektroder festet på huden. Elektrodene ble plassert over tre ulike muskelgrupper (hhv. m.frontalis, m.temporalis og m.masseter) på begge sider av ansiktet. Forfatterne beregnet gjennomsnittlig EMG-verdier og standardavvik. Det ble definert en grenseverdi (cut-off score) for hver muskelgruppe. Resultatene fra EMG ble også sammenliknet med resultater fra friske kjønns-, og aldersmatchede kontroller. Oversiktsforfatterne konkluderte med at begge studiene var svært utsatte for metodiske skjevheter (høy risiko). Tabell 5 oppsummerer resultatene om diagnostisk nøyaktighet; Sensitiviteten i den ene studien (76) var lav, og varierte fra 15 % til 40 % mens den var på 69 % i den andre studien (77). Spesifisiteten var høy (95 % - 98 %) i en studie (76) og relativt lav (67 %) i den andre studien (77). I tillegg var det stor variasjon og brede konfidensintervall i prediktive verdier og sannsynlighetsratioer. Forfatterne konkluderte derfor med at det ikke er tilstrekkelig vitenskapelig evidens for å støtte bruk av EMG i diagnostisering av TMD.

**Tabell 5.** Resultater om diagnostisk nøyaktighet av elektromyografi (EMG) sammenliknet med kliniske tester for å diagnostisere myofascial TMD, hentet fra oversikten til Al-Saleh og medarbeidere, 2012 (22).

| Forfatter/<br>År/<br>(antall deltakere) | Muskel                | Sensitivitet<br>95 % KI | Spesifisitet<br>95 % KI | PPV<br>95 % KI    | NPV<br>95 % KI    | LR+<br>95 % KI   | LR-<br>95 % KI    |
|---|-----------------------|-------------------------|-------------------------|-------------------|-------------------|------------------|-------------------|
| Glaros<br>1989<br>(n=40)                | Frontalis<br>Venstre  | 0,15<br>0,34-0,66       | 0,95<br>0,73 – 1,00     | 0,75<br>0,22-0,99 | 0,53<br>0,36-0,69 | 3,0<br>0,3-26,5  | 0,89<br>0,74-1,08 |
|   | Frontalis<br>Høyre    | 0,40<br>0,2-0,64        | 0,95<br>0,73-0,99       | 0,89<br>0,51-0,99 | 0,61<br>0,42-0,78 | 8,0<br>1,1-58,2  | 0,63<br>0,44-0,90 |
|   | Temporalis<br>Venstre | 0,40<br>0,2-0,64        | 0,98<br>0,77-1,0        | 0,94<br>0,54-1,0  | 0,63<br>0,44-0,78 | 16,4<br>1,0-26,9 | 0,61<br>0,43-0,88 |
|   | Temporalis<br>Høyre   | 0,15<br>0,04-0,39       | 0,95<br>0,73-1,0        | 0,75<br>0,22-0,99 | 0,53<br>0,36-0,69 | 3,0<br>0,3-26,5  | 0,9<br>0,74-1,1   |
|   | Masseter<br>Venstre   | 0,40<br>0,2-0,64        | 0,95<br>0,73-0,99       | 0,89<br>0,51-0,99 | 0,61<br>0,42-0,78 | 8,0<br>1,1-8,2   | 0,63<br>0,44-0,90 |
|   | Masseter<br>Høyre     | 0,25<br>0,1-0,5         | 0,95<br>0,73-1,0        | 0,83<br>0,36-0,99 | 0,56<br>0,38-0,72 | 5,0<br>0,8-31,0  | 0,79<br>0,52-1,2  |
| Glaros<br>1997<br>(n=108)               | Gjennomsnittsverdi    | 0,69<br>0,54-0,80       | 0,67<br>0,52-0,79       | 0,67<br>0,53-0,79 | 0,68<br>0,54-0,80 | 2,1<br>1,4-3,1   | 0,47<br>0,31-0,71 |

KI; Konfidensintervall, LR; Likelihood Ratio (sannsynlighetsratio), NPV; negative prediktiv verdi, PPV; positive prediktiv verdi.

---

# Diskusjon

I denne oversikten over systematiske oversikter har vi oppsummert resultater fra seks systematiske oversikter med i alt 67 primærstudier og ca 5400 deltakere. Oversiktene oppsummerer tilgjengelig vitenskapelig dokumentasjon om nøyaktighet av metoder for diagnostikk av ulike former for temporomandibulær dysfunksjon (TMD). Tre av de inkluderte oversiktene omhandler diagnostisk nøyaktighet av kliniske tester (14;25;57). De andre tre fokuserer på henholdsvis magnetisk resonanstomografi (23), ultralyd (26) og elektromyografi (24).

I mangel av en metodisk gullstandard for diagnostisering av TMD-tilstander er det viktig å presisere at funnene i denne oversikten ikke gir grunnlag for å avgjøre hvilke tester som best påviser eller utelukker TMD, men derimot om og i hvor stor grad ulike diagnostiske tester fører til samme konklusjon.

---

## Hva viser resultatene?

---

### MR-avbildning

MR anses ofte som referansestandard i diagnostisering av TMD-tilstander (14;15). Én av de inkluderte oversiktene (23) undersøkte den diagnostiske nøyaktigheten av MR for påvisning av ulike TMD-tilstander (diskusposisjoner, morfologi og benstrukturer) opp mot resultatene av biopsi eller frysesenitt (post mortem). De inkluderte primærstudiene viste gjennomgående god overensstemmelse mellom de to metodene, med høye verdier ( $>0,8$ ) for sensitivitet og spesifisitet. Oversiktforfatterne konkluderer likevel med at det er utilstrekkelig dokumentasjon for diagnostisk nøyaktighet av MR, og peker på et behov for flere valideringsstudier av høy kvalitet (23). I tillegg viste studiene relativt dårlig samsvar (reliabilitet) mellom bedømmingene til ulike observatører, noe som imidlertid ser ut til å bli bedre når kvaliteten på MR-bildene økes.

### Ultralyd

Ultralydapparater er rimeligere og ofte lettere tilgjengelig enn MR. Tross usikker dokumentasjon anses MR gjerne som gjeldende referansestand for påvisning av diskusforskyvning i kjeveleddet, og det er dermed interessant å undersøke om de to avbildningsteknikkene leder frem mot samme resultat. De oppsummerte resultatene viste sensitivitet og spesifisitet rundt 0,8 for ultralyd versus MR, men estimatene varierte til dels betydelig mellom ulike primær-

studier. Oversiktforfatterens konklusjon er at ultralyd kan benyttes som et raskt og preliminært diagnostisk instrument, men at positive funn bør verifiseres ved MR (26).

### **Kliniske tester**

Med kliniske tester menes undersøkelse av kliniske tegn og symptomer på temporomandibulær dysfunksjon, som for eksempel kjeveleddslyder (klikk og krepitasjon), smerte, begrenset gapevne eller kombinasjoner av disse. Vi identifiserte tre systematiske oversikter som undersøkte om kliniske tegn og symptomer kunne verifisere diskusforskyvning, degenerativ kjeveleddsforandring, eller inflammatorisk kjeveleddstilstander påvist ved MR. Tross usikker dokumentasjon anses MR ofte som gjeldende referansestandard for påvisning av disse tilstandene, og det er dermed interessant å undersøke i hvor stor grad de to metodene leder til samme konklusjon.

De tre inkluderte oversiktene viser uforklart heterogenitet i punkttestimatene for sensitivitet og spesifisitet når kliniske enkelttester sammenliknes med funn gjort ved MR-avbildning (14; 22;25). De positive sannsynlighetsratioene (LR+) var med få unntak lavere enn 2,5, og de negative (LR-) høyere enn 0,2, som betyr at kliniske tester har liten eller ingen verdi som diagnostisk markør såfremt MR aksepteres som referansestandard. Av unntak kan vi nevne resiprokalt klikk som kan være egnet til å påvise diskusforskyvning med tilbakegang (LR+ > 10), og kjeveleddssmerte ved assistert maksimal åpning (LR+ >6) og defleksjon (LR+ fra 6,3 til 9,0) som kan være egnet for å påvise diskusforskyvning uten tilbakegang. Et negativt resultat utelukker imidlertid ikke tilstanden, da LR- var > 0,25. Således kan den kliniske nytten virke liten, og siden vi mangler konfidensintervall for punkttestimatene kan vi heller ikke utelukke at dette dreier seg om tilfeldige enkeltfunn. Bruk av testkluster ser ikke ut til å øke den diagnostiske nøyaktigheten, men punkttestimatene varierte betydelig fra studie til studie for både sensitivitet (0,05 - 0,97) og spesifisitet (0,26 - 1,0).

### **Elektromyografi**

Det er vanlig å skille mellom og intraartikulære (herunder diskusforskyvning) og ekstraartikulære TMD-tilstander (som myofasciale smertetilstander) (3). Bildediagnostikk er potensiell nyttig for påvisning av kjeveleddsforandringer, mens ekstraartikulære TMD-tilstander som myofasciale smertetilstander muligens fordrer bruk av andre diagnostiske metoder. Elektromyografi (EMG) er basert på en antakelse om at patologiske og dysfunksjonelle tilstander kan identifiseres gjennom unormal aktivitet i kjeveleddsmuskulaturen. Al-Saleh og medarbeidere (24) inkluderte to tverrsnittstudier hvor resultater av EMG ble sammenliknet med resultater fra kliniske undersøkelser (76;77). Ifølge oversiktforfatterne var begge primærstudiene svært utsatt for metodiske skjevheter i tillegg til at resultatene viste stor variasjon. Det synes derfor ikke å være tilstrekkelig vitenskapelig dokumentasjon for bruk av EMG i diagnostisering av TMD (24).

---

## Utfordringer knyttet til manglende referansestandard

---

Når vi skal evaluere nøyaktigheten til en diagnostisk test er det en fordel dersom vi kan sammenligne resultatet opp mot en fasit, det vil si at det må finnes en gullstandard som gir sikkert svar på om en tilstand er tilstede eller ikke. Med fasit i hånd kan man enkelt sjekke om en nyutviklet test gir samme resultater som gullstandarden. Hvis resultatene overensstemmer godt kan vi konkludere med at den nye testen duger. Dersom avvikene fra gullstandarden blir mange må vi innse at den nyutviklede testen har vesentlige feil og mangler. I en slik idealisert verden reflekterer tallverdiene til sensitivitet, spesifisitet, sannsynlighetsratioer og prediktive verdier den nyutviklede testens egenskaper på en god måte. Virkeligheten er ofte mer komplisert.

I den virkelige verden sitter vi vanligvis ikke på fasiten. Av og til har vi tilgang til diagnostiske verktøy som gjør at vi med ganske stor grad av sikkerhet kan avgjøre om en person har sykdommen eller ikke, andre ganger har vi bare upresise diagnostiske verktøy. Det sier seg selv at vi kan komme skjevt ut hvis vi evaluerer treffsikkerheten til nyutviklede diagnostiske tester opp mot ufullstendige standarder. I slike tilfeller kan tall for sensitivitet, spesifisitet, sannsynlighetsratioer og prediktive verdier fortelle om to måter å diagnostisere på ender med samme resultat, men høye sensitivitets- og spesifisitetsverdier er ikke nødvendigvis et mål i seg selv.

Per i dag synes det ikke å eksistere en omforent referansestandard i diagnostisering av TMD, selv om bruk av MR ofte anses som en slik standard (14;15). I mangel av omforent, eller ufullkommen/imperfekt referansestandard åpnet vi for inklusjon av alle studier som, med utgangspunkt i de samme pasientene, sammenlignet grad av overensstemmelse mellom to eller flere diagnostiske prosedyrer. I tråd med diskusjonen over er det viktig å presisere at de inkluderte oversiktene ikke gir grunnlag for å avgjøre hvilke tester som best påviser eller utelukker ulike TMD-tilstander. De inkluderte oversiktene indikerer derimot om og i hvor stor grad ulike tester og prosedyrer fører til samme konklusjon.

Selv om en god diagnostisk referansestandard mangler, vil det likevel være behov for å kontrollere at en diagnostisk test gir meningsfulle svar (79). Heldigvis finnes det muligheter for å omgå de utfordringene som mangelen av en referansestandard skaper. I en oversiktsartikkel oppsummerer Reitsma og medarbeidere strategier som kan brukes for å bøte på at referansestandardene gir ufullstendige svar (79). Forfatterne viser til at det av og til er hensiktsmessig å flytte fokus fra diagnostisk nøyaktighet (sensitivitet, spesifisitet med mer) til studier som undersøker om et diagnostisk testresultat er klinisk meningsfylt, også kalt valideringsstudier (79). Vi kan argumentere for at kliniske diagnoser først blir viktig idet de har betydning for det videre pasientforløpet gjennom prediksjon av behandlingsrespons eller prognose. Hvis en diagnostisk prosedyre evner å skille ut TMD-pasienter som vil respondere på kirurgi kan vi konkludere med at den diagnostiske prosedyren er valid. Diagnostiske valideringsstudier krever gjerne andre metodiske oppsett enn diagnostiske nøyaktighetsstudier, blant annet kan randomiserte kontrollerte design komme til nytte (80). Ingen av de inkluderte oversiktene

refererer til denne typen valideringsstudier, men siden vi ikke har lett etter primærstudier kan vi ikke utelukke muligheten for at de finnes.

Mangel på en gullstandard, eller i tilfeller hvor man har en imperfekt referansestandard som her, kan medføre at en indekstest blir målt mot en standard som ikke i tilstrekkelig grad gjenspeiler "sannheten", såkalt referansestandardbias (81). Dette vil igjen kunne gi utfordringer med hensyn til hvordan man skal oppsummere de diagnostiske resultatene. I slike tilfeller er det anbefalt at man bør behandle indekstest og referansetest som to alternative metoder, og måle samsvar mellom testresultatene, dvs. kappa-statistikk for dikotome data (81). Vi har imidlertid valgt, siden dette er en oversikt over systematiske oversikter, ikke å benytte slik statistikk, men heller oppsummerer resultatene slik de fremstår i de inkluderte oversiktene.

---

## **Oversikt over oversikter – fordeler og ulemper**

---

### **Hvorfor oversikt over oversikter?**

Da vi startet arbeidet med denne rapporten var artikkeltilfanget ukjent. Utprøvende litteratursøk brakte på det rene at det finnes mange relevante systematiske oversikter, og vi valgte derfor å besvare Helsedirektoratets bestilling med en rapport som oppsummerer systematiske oversikter – også kalt oversikt over oversikter. Tilnærmingen kan sies å være grovmasket, men er velegnet til å gi en overordnet beskrivelse av status innenfor et bredt definert tematisk område. Selv om personer med TMD kan sies å være en avgrenset populasjon er den diagnostiske prosessen og bruken av diagnostiske hjelpemidler preget av stor variasjon. Denne rapporten dekker diagnostiske metoder som spenner fra ulike former for bildediagnostikk via elektromyografi til kliniske tegn som klikkelyder og smerte. Variasjonen i tematiske områder er derfor stor, og vi mener at en oversikt over oversikter er en velegnet inngangsport til problemstillingen.

### **Kan vi stole på resultatene?**

For å avgjøre om vi kan ha tillit til resultatene som presenteres i oversikter over oversikter må vi vurdere kvaliteten på bakenforliggende data – vi tenker da både på kvaliteten av hver enkelt oversikt og kvaliteten til tilgjengelige enkeltstudier. Det er lite sannsynlig at systematiske oversikter av dårlig kvalitet gir korrekte virkelighetsbilder, så i arbeidet med vår rapport kvalitetsvurderte vi alle relevante systematiske oversikter opp mot en forhåndsdefinert sjekkliste. Kunnskapssenterets sjekkliste ([www.kunnskapssenteret.no](http://www.kunnskapssenteret.no)) er utarbeidet med fokus på systematiske oversikter over intervensjonsstudier, og vi valgte derfor å benytte en noe forenklet sjekkliste som baserer seg på sjekklisten som er oppgitt i Kunnskapssenterets metodehåndbok og sjekklisten fra Critical Appraisal Skills Program (vedlegg 2). Vi inkluderte systematiske oversikter som vi vurderte til å være av god og moderat kvalitet. Nedgraderingene til moderat kvalitet skyldtes hovedsakelig manglende rapportering av resultater fra vurdering av risiko for systematiske skjevheter i de inkluderte primærstudiene, eller at litteratursøket etter vår mening ikke var omfattende nok. Oversikter av lav kvalitet ble ekskludert og står oppført i

tabellen over ekskluderte studier (vedlegg 3). Alt i alt føler vi oss ganske sikre på at de inkluderte systematiske oversiktene gir et representativt virkelighetsbilde av relevant og tilgjengelig forskning på feltet.

### **Hva har vi gått glipp av?**

Under visse betingelser kan en oversikt over oversikter gir ufullstendige svar. Spesielt gjelder dette hvis metoden benyttes på fagfelt med stor forskningsaktivitet. I slike tilfeller er det fare for at det forligger primærstudier av nyere dato som ikke dekkes av tilgjengelige systematiske oversikter. Oversiktene vi inkluderte i denne rapporten var stort sett av nyere dato, og det er lite sannsynlig at det er tilkommet nye primærstudier som evner å endre rapportens overordnede konklusjoner. Det er antakelig større fare for at relevante underspørsmål ikke dekkes av rapporten fordi vi ikke fant systematiske oversikter med tilfredsstillende kvalitet. Vi ser for eksempel at ingen av de inkluderte oversiktene undersøker diagnostisk nøyaktighet av CT til tross for at det finnes relevante enkeltstudier på området (se for eksempel referanse (82)).

Hvis enkelte underspørsmål peker seg ut som spesielt interessante kan det være aktuelt å gjennomføre et spesifikt søk etter primærstudier. I den sammenheng er det verdt å nevne at vi allerede har utarbeidet og gjennomført et søk etter primærstudier, men på nåværende tidspunkt har vi ikke gått inn i materialet. Vi føler oss likevel rimelig sikre på at de inkluderte systematiske oversiktene er et godt utgangspunkt for å vurdere eksisterende forskning. Vårt hovedinntrykk er at det er manglende diagnostisk referansestandard og robuste primærstudier som er de største hindringene mot sikre konklusjoner.

---

## **Dokumentasjonens kvalitet**

---

GRADE er et verktøy som brukes for å vurdere kvaliteten til forskningsdokumentasjonen og hvilken tillit vi kan ha til resultatene (83;84). GRADE er utviklet med spesiell vekt på effektspørsmål, men GRADE vinner innpass på stadig nye områder herunder diagnostikkfeltet (85). Bruken av GRADE innenfor diagnostikkdomenet er på utviklingsstadiet, og vi planla derfor ikke å bruke GRADE til å fastslå kvaliteten på dokumentasjonen som presenteres i denne oversikten over oversikter. Vi har likevel her valgt å diskutere resultatene opp mot noen av de samme elementene som benyttes for å vurdere kvaliteten på effektforskning i GRADE.

### **Kvalitet på enkeltstudier**

I denne oversikten har vi ikke selv vurdert den metodiske kvaliteten på primærstudier, isteden baserer vi oss på vurderinger som er gjort av forfatterne av de ulike inkluderte oversiktene. Av de seks inkluderte oversiktene tok fem utgangspunkt i QUADAS (13) for å vurdere kvaliteten på inkluderte studier. Den siste oversikten refererer ikke spesifikt til QUADAS, men forfatterne benytter en sjekklister som har fellestrekk med spørsmålene i QUADAS. Metodisk kvalitet er dermed stort sett bedømt med utgangspunkt i de samme kriteriene, men dette sikrer ikke nødvendigvis en konsistent bedømming. Vurderinger av studiekvalitet baserer seg til en viss grad på subjektive skjønnsmessige vurderinger, og ulike oversiktsforfattere kan ha vekt-

lagt ulike kriterier. På tvers av de oversiktene er det imidlertid et fellestrekk at mange primærstudier har metodiske svakheter som øker risikoen for at resultatene er utsatt for metodiske feilkilder.

## **Konsistens**

Med konsistens menes grad av samsvar mellom primærstudier som søker å besvare samme spørsmål (86). Med unntak av sammenligningen av MR versus frysesenitt der ulike primærstudier overensstemmer relativt godt med hverandre, er det et gjennomgående problem at punkt-estimatene for diagnostisk nøyaktighet i studier som ser på tilnærmet samme problemstilling variere mye på tvers av de inkluderte studiene. Det er ikke vanskelig å tenke seg til faktorer som potensielt kan føre til slik heterogenitet, herunder forskjeller i avbildningsprotokoller, ulik bruk av terskelverdier og ulikheter i pasientpopulasjonen. Så lenge de inkluderte oversiktene ikke evner å forklare de faktiske årsakene til den observerte heterogeniteten bidrar stor heterogenitet til å redusere vår tillit til dokumentasjonen.

## **Presisjon**

I systematiske oversikter som gjør bruk av meta-analyser viser presisjonen gjerne til konfidensintervallet til det sammenslåtte effektestimater, som i sin tur bestemmes av konsistens mellom ulike studier og størrelsen på tilgjengelige enkeltstudie. Vi sier at effektestimater er presist dersom det er basert på et tilstrekkelig antall pasienter samtidig som det resulterende konfidensintervallet er så smalt at det gir grunnlag for en sikker konklusjon (87). Et resultat regnes som upresist hvis konfidensintervallet er så bredt at det er det vanskelig å trekke konklusjoner om den sanne kliniske effekten.

Effektstudier rapporterer gjerne effekt som risikoratio eller endring i gjennomsnitt. Til sammenligning rapporterer diagnostiske nøyaktighetsstudier estimer for sensitivitet, spesifisitet, sannsynlighetsratioer og prediktive verdier (18). I likhet med effektstudier der smale konfidensintervall er tegn på god presisjon ønsker vi smale konfidensintervall også i diagnostiske studier, men i diagnostiske nøyaktighetsstudier må vi forholde seg til flere estimat på en gang. Sensitivitet henger eksempelvis nært sammen med spesifisitet. Bedret sensitivitet fører typisk til at spesifisitet blir dårligere, og fortolkningen av verdiene kan derfor ikke gjøres uavhengig av hverandre. Her må vi også ta stilling til hva som kan anses som en akseptabel terskel for diagnostisk nøyaktighet ved diagnostikk av TMD. RDC/TMD definerte i sin tid sensitivitet  $\geq 0,7$  og spesifisitet  $>0,95$  som akseptable terskler for diagnostisk nøyaktighet (8). Dette betyr, basert på det utvalget man har studert, at man vil oppdage minst 70 % av de som faktisk har sykdommen, og at testen vil være negativ hos minst 95 % av de som faktisk er friske. Dette betyr også at 30 % av de som er syke feilaktig vil bli diagnostisert som friske (falsk negative), og 5 % av de friske vil feilaktig bli diagnostisert som syke (falsk positive). Hvis vi ser dette opp mot rapporterte verdier for sensitivitet og spesifisitet i ulike enkeltstudier er det få enkeltstudier som måler akseptable verdier, og i de tilfeller der konfidensintervall er oppgitt krysser de ofte grensene for det som betraktes som en akseptabel terskel. En tilsvarende mangel på presisjon er å finne igjen i konfidensintervallene for sannsynlighetsra-

tioene. Manglende presisjon kan i mange tilfeller tilskrives at det bakenforliggende pasientmaterialet er lite. I tillegg til brede konfidensintervall i kombinasjon med stor inkonsistens på tvers av inkluderte studiene, er det derfor vanskelig, basert på det tilgjengelige datamaterialet, å trekke sikre konklusjoner.

### **Direkthet**

Spørsmålet om direkthet handler om de inkluderte primærstudiene omhandler deltakere og sammenligninger som er relevante, og dermed om resultatene er overførbare for de spørsmålene vi stiller oss (88). De inkluderte studiene omfattet gjennomgående selekterte pasientgrupper i spesialisthelsetjenesten. De fleste pasientene var kvinner, anslagsvis ca 80 %, og alderen varierte fra 13 til 79 år. Studier om forekomst viser at de fleste som har TMD er kvinner i alderen 20-50 år (89). Andre forhold som kan ha betydning for overførbarheten er at den praktiske gjennomføringen av referanse- og indekstest ikke alltid er tilfredsstillende beskrevet. En del inkluderte studier er også av eldre dato, spesielt gjelder dette en del av studiene som er inkludert i oversikten om diagnostisk nøyaktighet av MR versus biopsi/frysesnitt (90). Siden MR-teknologien er forbedret siden starten av 90-tallet kan vi ikke være helt sikre på overførbarheten til noen av de eldre studiene.



---

# Konklusjon

Oppsummerte resultater om diagnostisk nøyaktighet av ulike tester for påvisning av tempo-romandibulær dysfunksjon (TMD) fra seks systematiske oversikter med i alt 67 primærstudier og over 5400 deltakere (gjennomgående høyselekterte pasientgrupper av ungdom og voksne i spesialisthelsetjenesten) viser at:

- Undersøkelse med magnetisk resonanstomografi (MR), særlig bilder kombinert i sagittal- og koronalplan, ser ut til å ha akseptabel nøyaktighet for å påvise eller utelukke diskusforskyvning, diskusdeformitet og ossøse kondylendringer i kjeveleddet, sammenliknet med frysesnitt (post mortem) eller biopsier.
- Ultralyd ser ut til å ha akseptabel nøyaktighet for å påvise, men ikke utelukke, diskusforskyvning, sammenliknet med MR-undersøkelse.
- Kliniske tester for undersøkelse av kjeveleddslyder, smerte og kjeveleddsbevegelser, ser ikke ut til å ha tilfredsstillende nøyaktighet for å påvise eller utelukke TMD-tilstander, sammenliknet med MR-undersøkelse.
- Elektromyografi ser ikke ut til å ha tilfredsstillende nøyaktighet for å påvise eller utelukke myofascial TMD sammenliknet med kliniske tester.
- Resultatene må tolkes med forsiktighet, fordi
  - primærstudiene i de inkluderte systematiske oversiktene kan ha metodiske svakheter
  - det er stor variasjon i resultatene
  - feilmarginene (konfidensintervallene) i nøyaktighetsmålene er gjennomgående brede eller ikke oppgitt.

---

# Referanser

1. Bjørnland T, Møystad A. Kroniske sykdommer og tilstander i kjeveledd og muskulatur. *Nor Tannlegeforen Tid* 2010;120:896-902.
2. Hylander WL. Functional anatomy and biomechanics of the masticatory apparatus. I: Laskin DM, Greene CS, Hylander WL, editors. *Temporimandibular Disorders. An evidence-based approach to diagnosis and treatment*. Quintessence Publishing Co, Inc; 2006. p. 3-34.
3. Scrivani SJ, Keith DA, Kaban LB. Temporomandibular disorders. *N Engl J Med* 2008;359(25):2693-705.
4. Anastassaki KA, Hugoson A, Magnusson T. Prevalence of symptoms indicative of temporomandibular disorders in adults: cross-sectional epidemiological investigations covering two decades. *Acta Odontol Scand* 2012;70(3):213-23.
5. Gonzalez YM, Mohl ND. Masticatory muscle pain and dysfunction. I: Laskin DM, Greene CS, Hylander WL, editors. *Temporimandibular Disorders. An evidence-based approach to diagnosis and treatment*. Quintessence Publishing Co, Inc; 2006. p. 255-69.
6. Laskin DM. Internal derangements. I: Laskin DM, Greene CS, Hylander WL, editors. *Temporimandibular Disorders. An evidence-based approach to diagnosis and treatment*. Quintessence Publishing Co, Inc; 2006. p. 249-53.
7. Manfredini D, Lobbezoo F. Relationship between bruxism and temporomandibular disorders: A systematic review of literature from 1998 to 2008. *Oral Surgery, Oral Medicine, Oral Pathology, Oral Radiology and Endodontology* 2010;109(6):e26-e50.
8. Dworkin SF, LeResche L. Research diagnostic criteria for temporomandibular disorders: review, criteria, examinations and specifications, critique. *J Craniomandib Disord* 1992;6(4):301-55.
9. Dworkin SF. Research Diagnostic criteria for Temporomandibular Disorders: current status & future relevance. *J Oral Rehabil* 2010;37(10):734-43.
10. Larheim TA. Role of magnetic resonance imaging in the clinical diagnosis of the temporomandibular joint. *Cells Tissues Organs* 2005;180(1):6-21.
11. Suvinen TI, Kempainen P. Review of clinical EMG studies related to muscle and occlusal factors in healthy and TMD subjects. *J Oral Rehabil* 2007;34(9):631-44.
12. Manfredini D, Bucci MB, Montagna F, Guarda-Nardini L. Temporomandibular disorders assessment: medicolegal considerations in the evidence-based era. *J Oral Rehabil* 2011;38(2):101-19.
13. Whiting P, Rutjes AW, Reitsma JB, Bossuyt PM, Kleijnen J. The development of QUADAS: a tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews. *BMC Med Res Methodol* 2003;3:25.

14. Chaput E, Gross A, Stewart R, Nadeau G, Goldsmith CH. The Diagnostic Validity of Clinical Tests in Temporomandibular Internal Derangement: A Systematic Review and Meta-analysis. *Physiotherapy Canada* 2012;64(2):116-34.
15. Larheim T, Westesson P-L. TMJ Imaging. I: Laskin DM, Greene CS, Hylander WL, editors. *Temporimandibular Disorders. An evidence-based approach to diagnosis and treatment.* Quintessence Publishing Co, Inc; 2006. p. 149-79.
16. Schunemann HJ, Oxman AD, Brozek J, Glasziou P, Bossuyt P, Chang S, et al. GRADE: assessing the quality of evidence for diagnostic recommendations. *Evid Based Med* 2008;13(6):162-3.
17. Irwig L, Tosteson AN, Gatsonis C, Lau J, Colditz G, Chalmers TC, et al. Guidelines for meta-analyses evaluating diagnostic tests. *Ann Intern Med* 1994;120(8):667-76.
18. Altman DG, Bland JM. Diagnostic tests. 1: Sensitivity and specificity. *BMJ* 1994;308(6943):1552.
19. Altman DG, Bland JM. Diagnostic tests 2: Predictive values. *BMJ* 1994;309(6947):102.
20. Deeks JJ, Altman DG. Diagnostic tests 4: likelihood ratios. *BMJ* 2004;329(7458):168-9.
21. Jaeschke R, Guyatt GH, Sackett DL. Users' guides to the medical literature. III. How to use an article about a diagnostic test. B. What are the results and will they help me in caring for my patients? The Evidence-Based Medicine Working Group. *JAMA* 1994;271(9):703-7.
22. Koh K, List T, Petersson A, Rohlin M. Relationship between clinical and magnetic resonance imaging diagnoses and findings in degenerative and inflammatory temporomandibular joint diseases: a systematic literature review. *J Orofac Pain* 2009;23(2):123-9.
23. Limchaichana N, Petersson A, Rohlin M. The efficacy of magnetic resonance imaging in the diagnosis of degenerative and inflammatory temporomandibular joint disorders: a systematic literature review. *Oral Surgery, Oral Medicine, Oral Pathology, Oral Radiology, and Endodontics* 2006;102(4):521-36.
24. Al-Saleh MA, Armijo-Olivo S, Flores-Mir C, Thie NM. Electromyography in diagnosing temporomandibular disorders. *J Am Dent Assoc* 2012;143(4):351-62.
25. Reneker J, Paz J, Petrosino C, Cook C. Diagnostic accuracy of clinical tests and signs of temporomandibular joint disorders: a systematic review of the literature. *J Orthop Sports Phys Ther* 2011;41(6):408-16.
26. Li CJ, Su NC, Yang XR, Yang X, Shi ZD, Li LJ. Ultrasonography for Detection of Disc Displacement of Temporomandibular Joint: A Systematic Review and Meta-Analysis. *Journal of Oral and Maxillofacial Surgery* 2012;70(6):1300-9.
27. Barclay P, Hollender LG, Maravilla KR, Truelove EL. Comparison of clinical and magnetic resonance imaging diagnosis in patients with disk displacement in the temporomandibular joint. *Oral Surg Oral Med Oral Pathol Oral Radiol Endod* 1999;88(1):37-43.
28. Bertram S, Rudisch A, Innerhofer K, Pumpel E, Grubwieser G, Emshoff R. Diagnosing TMJ internal derangement and osteoarthritis with magnetic resonance imaging. *J Am Dent Assoc* 2001;132(6):753-61.
29. Emshoff R, Innerhofer K, Rudisch A, Bertram S. The biological concept of "internal derangement and osteoarthrosis": a diagnostic approach in patients with temporomandibular joint pain? *Oral Surg Oral Med Oral Pathol Oral Radiol Endod* 2002;93(1):39-44.
30. Manfredini D, Basso D, Salmaso L, Guarda-Nardini L. Temporomandibular joint click sound and magnetic resonance-depicted disk position: which relationship? *J Dent* 2008;36(4):256-60.

31. Marguelles-Bonnet RE, Carpentier P, Yung JP, Defrennes D, Pharaboz C. Clinical diagnosis compared with findings of magnetic resonance imaging in 242 patients with internal derangement of the TMJ. *J Orofac Pain* 1995;9(3):244-53.
32. Rudisch A, Innerhofer K, Bertram S, Emshoff R. Magnetic resonance imaging findings of internal derangement and effusion in patients with unilateral temporomandibular joint pain. *Oral Surg Oral Med Oral Pathol Oral Radiol Endod* 2001;92(5):566-71.
33. Taskaya-Yilmaz N, Ogutcen-Toller M. Clinical correlation of MRI findings of internal derangements of the temporomandibular joints. *Br J Oral Maxillofac Surg* 2002;40(4):317-21.
34. Usumez S, Oz F, Guray E. Comparison of clinical and magnetic resonance imaging diagnoses in patients with TMD history. *J Oral Rehabil* 2004;31(1):52-6.
35. Holmlund AB, Axelsson S. Temporomandibular arthropathy: correlation between clinical signs and symptoms and arthroscopic findings. *Int J Oral Maxillofac Surg* 1996;25(3):178-81.
36. Manfredini D, Tognini F, Zampa V, Bosco M. Predictive value of clinical findings for temporomandibular joint effusion. *Oral Surg Oral Med Oral Pathol Oral Radiol Endod* 2003;96(5):521-6.
37. Orsini MG, Kuboki T, Terada S, Matsuka Y, Yatani H, Yamashita A. Clinical predictability of temporomandibular joint disc displacement. *J Dent Res* 1999;78(2):650-60.
38. Stegenga B, De Bont LG, van der Kuijl B, Boering G. Classification of temporomandibular joint osteoarthritis and internal derangement. 1. Diagnostic significance of clinical and radiographic symptoms and signs. *Cranio* 1992;10(2):96-106.
39. Yatani H, Suzuki K, Kuboki T, Matsuka Y, Maekawa K, Yamashita A. The validity of clinical examination for diagnosing anterior disk displacement without reduction. *Oral Surg Oral Med Oral Pathol Oral Radiol Endod* 1998;85(6):654-60.
40. Adame CG, Monje F, Offnoz M, Martin-Granizo R. Effusion in magnetic resonance imaging of the temporomandibular joint: a study of 123 joints. *J Oral Maxillofac Surg* 1998;56(3):314-8.
41. Emshoff R, Rudisch A. Validity of clinical diagnostic criteria for temporomandibular disorders: clinical versus magnetic resonance imaging diagnosis of temporomandibular joint internal derangement and osteoarthritis. *Oral Surg Oral Med Oral Pathol Oral Radiol Endod* 2001;91(1):50-5.
42. Emshoff R, Innerhofer K, Rudisch A, Bertram S. Relationship between temporomandibular joint pain and magnetic resonance imaging findings of internal derangement. *Int J Oral Maxillofac Surg* 2001;30(2):118-22.
43. Emshoff R, Brandmaier I, Bertram S, Rudisch A. Magnetic resonance imaging findings of osteoarthritis and effusion in patients with unilateral temporomandibular joint pain. *Int J Oral Maxillofac Surg* 2002;31(6):598-602.
44. Emshoff R, Brandmaier I, Bertram S, Rudisch A. Relative odds of temporomandibular joint pain as a function of magnetic resonance imaging findings of internal derangement, osteoarthritis, effusion, and bone marrow edema. *Oral Surg Oral Med Oral Pathol Oral Radiol Endod* 2003;95(4):437-45.
45. Emshoff R, Brandmaier I, Gerhard S, Strobl H, Bertram S, Rudisch A. Magnetic resonance imaging predictors of temporomandibular joint pain. *J Am Dent Assoc* 2003;134(6):705-14.
46. Guler N, Yatmaz PI, Ataoglu H, Emlik D, Uckan S. Temporomandibular internal derangement: correlation of MRI findings with clinical symptoms of pain and joint sounds in patients with bruxing behaviour. *Dentomaxillofac Radiol* 2003;32(5):304-10.

47. Haley DP, Schiffman EL, Lindgren BR, Anderson Q, Andreasen K. The relationship between clinical and MRI findings in patients with unilateral temporomandibular joint pain. *J Am Dent Assoc* 2001;132(4):476-81.
48. Huddleston Slater JJ, Lobbezoo F, Chen YJ, Naeije M. A comparative study between clinical and instrumental methods for the recognition of internal derangements with a clicking sound on condylar movement. *J Orofac Pain* 2004;18(2):138-47.
49. Katzberg RW, Westesson PL, Tallents RH, Drake CM. Anatomic disorders of the temporomandibular joint disc in asymptomatic subjects. *J Oral Maxillofac Surg* 1996;54(2):147-53.
50. Limchaichana N, Nilsson H, Ekberg EC, Nilner M, Petersson A. Clinical diagnoses and MRI findings in patients with TMD pain. *J Oral Rehabil* 2007;34(4):237-45.
51. Paesani D, Westesson PL, Hatala MP, Tallents RH, Brooks SL. Accuracy of clinical diagnosis for TMJ internal derangement and arthrosis. *Oral Surg Oral Med Oral Pathol* 1992;73(3):360-3.
52. Rammelsberg P, Pospiech PR, Jager L, Pho Duc JM, Bohm AO, Gernet W. Variability of disk position in asymptomatic volunteers and patients with internal derangements of the TMJ. *Oral Surg Oral Med Oral Pathol Oral Radiol Endod* 1997;83(3):393-9.
53. Ribeiro RF, Tallents RH, Katzberg RW, Murphy WC, Moss ME, Magalhaes AC, et al. The prevalence of disc displacement in symptomatic and asymptomatic volunteers aged 6 to 25 years. *J Orofac Pain* 1997;11(1):37-47.
54. Sano T, Westesson PL. Magnetic resonance imaging of the temporomandibular joint. Increased T2 signal in the retrodiskal tissue of painful joints. *Oral Surg Oral Med Oral Pathol Oral Radiol Endod* 1995;79(4):511-6.
55. Shaefer JR, Jackson DL, Schiffman EL, Anderson QN. Pressure-pain thresholds and MRI effusions in TMJ arthralgia. *J Dent Res* 2001;80(10):1935-9.
56. Takahashi T, Nagai H, Seki H, Fukuda M. Relationship between joint effusion, joint pain, and protein levels in joint lavage fluid of patients with internal derangement and osteoarthritis of the temporomandibular joint. *J Oral Maxillofac Surg* 1999;57(10):1187-93.
57. Koh K, List T, Petersson A, Rohlin M. Relationship between clinical and magnetic resonance imaging diagnoses and findings in degenerative and inflammatory temporomandibular joint diseases: a systematic literature review. *Journal of orofacial pain* 2009;23(2):123-9.
58. Nebbe B, Brooks SL, Hatcher D, Hollender LG, Prasad NG, Major PW. Magnetic resonance imaging of the temporomandibular joint: interobserver agreement in subjective classification of disk status. *Oral Surg Oral Med Oral Pathol Oral Radiol Endod* 2000;90(1):102-7.
59. Orsini MG, Terada S, Kuboki T, Matsuka Y, Yamashita A. The influence of observer calibration in temporomandibular joint magnetic resonance imaging diagnosis. *Oral Surg Oral Med Oral Pathol Oral Radiol Endod* 1997;84(1):82-7.
60. Orsini MG, Kuboki T, Terada S, Matsuka Y, Yamashita A, Clark GT. Diagnostic value of 4 criteria to interpret temporomandibular joint normal disk position on magnetic resonance images. *Oral Surg Oral Med Oral Pathol Oral Radiol Endod* 1998;86(4):489-97.
61. Schmitter M, Kress B, Hahnel S, Rammelsberg P. The effect of quality of temporomandibular joint MR images on interrater agreement. *Dentomaxillofac Radiol* 2004;33(4):253-8.
62. Katzberg RW, Westesson PL, Tallents RH, Anderson R, Kurita K, Manzione JV, Jr., et al. Temporomandibular joint: MR assessment of rotational and sideways disk displacements. *Radiology* 1988;169(3):741-8.

63. Schwaighofer BW, Tanaka TT, Klein MV, Sartoris DJ, Resnick D. MR imaging of the temporomandibular joint: a cadaver study of the value of coronal images. *AJR Am J Roentgenol* 1990;154(6):1245-9.
64. Tasaki MM, Westesson PL. Temporomandibular joint: diagnostic accuracy with sagittal and coronal MR imaging. *Radiology* 1993;186(3):723-9.
65. Westesson PL, Katzberg RW, Tallents RH, Sanchez-Woodworth RE, Svensson SA. CT and MR of the temporomandibular joint: comparison with autopsy specimens. *AJR Am J Roentgenol* 1987;148(6):1165-71.
66. Westesson PL, Katzberg RW, Tallents RH, Sanchez-Woodworth RE, Svensson SA, Espeland MA. Temporomandibular joint: comparison of MR images with cryosectional anatomy. *Radiology* 1987;164(1):59-64.
67. Larheim TA, Westesson PL, Hicks DG, Eriksson L, Brown DA. Osteonecrosis of the temporomandibular joint: correlation of magnetic resonance imaging and histology. *J Oral Maxillofac Surg* 1999;57(8):888-98.
68. Brandlmaier I, Rudisch A, Bodner G, Bertram S, Emshoff R. Temporomandibular joint internal derangement: detection with 12.5 MHz ultrasonography. *J Oral Rehabil* 2003;30(8):796-801.
69. Emshoff R, Bertram S, Rudisch A, Gassner R. The diagnostic value of ultrasonography to determine the temporomandibular joint disk position. *Oral Surg Oral Med Oral Pathol Oral Radiol Endod* 1997;84(6):688-96.
70. Emshoff R, Jank S, Rudisch A, Walch C, Bodner G. Error patterns and observer variations in the high-resolution ultrasonography imaging evaluation of the disk position of the temporomandibular joint. *Oral Surg Oral Med Oral Pathol Oral Radiol Endod* 2002;93(3):369-75.
71. Emshoff R, Jank S, Rudisch A, Bodner G. Are high-resolution ultrasonographic signs of disc displacement valid? *J Oral Maxillofac Surg* 2002;60(6):623-8.
72. Jank S, Rudisch A, Bodner G, Brandlmaier I, Gerhard S, Emshoff R. High-resolution ultrasonography of the TMJ: helpful diagnostic approach for patients with TMJ disorders ? *J Craniomaxillofac Surg* 2001;29(6):366-71.
73. Jank S, Emshoff R, Norer B, Missmann M, Nicasi A, Strobl H, et al. Diagnostic quality of dynamic high-resolution ultrasonography of the TMJ--a pilot study. *Int J Oral Maxillofac Surg* 2005;34(2):132-7.
74. Landes CA, Goral W, Mack MG, Sader R. 3-D sonography for diagnosis of osteoarthritis and disk degeneration of the temporomandibular joint, compared with MRI. *Ultrasound Med Biol* 2006;32(5):627-32.
75. Landes CA, Goral WA, Sader R, Mack MG. Three-dimensional versus two-dimensional sonography of the temporomandibular joint in comparison to MRI. *Eur J Radiol* 2007;61(2):235-44.
76. Glaros AG, McGlynn FD, Kapel L. Sensitivity, specificity, and the predictive value of facial electromyographic data in diagnosing myofascial pain-dysfunction. *Cranio* 1989;7(3):189-93.
77. Glaros AG, Glass EG, Brockman D. Electromyographic data from TMD patients with myofascial pain and from matched control subjects: evidence for statistical, not clinical, significance. *J Orofac Pain* 1997;11(2):125-9.

78. Li CJ, Su NC, Yang XR, Yang X, Shi ZD, Li LJ. Ultrasonography for Detection of Disc Displacement of Temporomandibular Joint: A Systematic Review and Meta-Analysis. *Journal of Oral and Maxillofacial Surgery* 2012;70(6):1300-9.
79. Reitsma JB, Rutjes AW, Khan KS, Coomarasamy A, Bossuyt PM. A review of solutions for diagnostic accuracy studies with an imperfect or missing reference standard. *J Clin Epidemiol* 2009;62(8):797-806.
80. Rodger M, Ramsay T, Fergusson D. Diagnostic randomized controlled trials: the final frontier. *Trials* 2012;13:137.
81. Smetana GW, Umscheid CA, Chang S, Matchar DB. Methods guide for authors of systematic reviews of medical tests: a collaboration between the Agency for Healthcare Research and Quality (AHRQ) and the Journal of General Internal Medicine. *J Gen Intern Med* 2012;27 Suppl 1:S1-S3.
82. Liedberg J, Panmekiate S, Petersson A, Rohlin M. Evidence-based evaluation of three imaging methods for the temporomandibular disc. *Dentomaxillofac Radiol* 1996;25(5):234-41.
83. Balshem H, Helfand M, Schunemann HJ, Oxman AD, Kunz R, Brozek J, et al. GRADE guidelines: 3. Rating the quality of evidence. *J Clin Epidemiol* 2011;64(4):401-6.
84. Guyatt GH, Oxman AD, Schunemann HJ, Tugwell P, Knottnerus A. GRADE guidelines: a new series of articles in the Journal of Clinical Epidemiology. *J Clin Epidemiol* 2011;64(4):380-2.
85. Schunemann HJ, Oxman AD, Brozek J, Glasziou P, Jaeschke R, Vist GE, et al. Grading quality of evidence and strength of recommendations for diagnostic tests and strategies. *BMJ* 2008;336(7653):1106-10.
86. Guyatt GH, Oxman AD, Kunz R, Woodcock J, Brozek J, Helfand M, et al. GRADE guidelines: 7. Rating the quality of evidence--inconsistency. *J Clin Epidemiol* 2011;64(12):1294-302.
87. Guyatt GH, Oxman AD, Kunz R, Brozek J, Alonso-Coello P, Rind D, et al. GRADE guidelines 6. Rating the quality of evidence--imprecision. *J Clin Epidemiol* 2011;64(12):1283-93.
88. Guyatt GH, Oxman AD, Kunz R, Woodcock J, Brozek J, Helfand M, et al. GRADE guidelines: 8. Rating the quality of evidence--indirectness. *J Clin Epidemiol* 2011;64(12):1303-10.
89. Bagis B, Ayaz EA, Turgut S, Durkan R, Ozcan M. Gender difference in prevalence of signs and symptoms of temporomandibular joint disorders: a retrospective study on 243 consecutive patients. *Int J Med Sci* 2012;9(7):539-44.
90. Limchaichana N, Petersson A, Rohlin M. The efficacy of magnetic resonance imaging in the diagnosis of degenerative and inflammatory temporomandibular joint disorders: a systematic literature review. *Oral Surgery, Oral Medicine, Oral Pathology, Oral Radiology, and Endodontics* 2006;102(4):521-36.
91. Cakli H, Cingi C, Ay Y, Oghan F, Ozer T, Kaya E. Use of cone beam computed tomography in otolaryngologic treatments. *European Archives of Oto-Rhino-Laryngology* 2012;269(3):711-20.
92. Cummings M, Baldry P. Regional myofascial pain: diagnosis and management. *Best Practice and Research in Clinical Rheumatology* 2007;21(2):367-87.
93. De Senna BR, Dos Santos Silva VK, Franca JP, Marques LS, Pereira LJ. Imaging diagnosis of the temporomandibular joint: Critical review of indications and new perspectives. *Oral Radiology* 2009;25(2):86-98.

94. De Vos W, Casselman J, Swennen GRJ. Cone-beam computerized tomography (CBCT) imaging of the oral and maxillofacial region: A systematic review of the literature. *Int J Oral Maxillofac Surg* 2009;38(6):609-25.
95. Klasser GD, Okeson JP. The clinical usefulness of surface electromyography in the diagnosis and treatment of temporomandibular disorders. *J Am Dent Assoc* 2006;137(6):763-71.
96. Lobbezoo F, Drangsholt M, Peck C, Sato H, Kopp S, Svensson P. Topical review: New insights into the pathology and diagnosis of disorders of the temporomandibular joint. *J Orofac Pain* 2004;18(3):181-91.
97. Lucas N, Macaskill P, Irwig L, Moran R, Bogduk N. Reliability of Physical Examination for Diagnosis of Myofascial Trigger Points A Systematic Review of the Literature. *Clin J Pain* 2009;25(1):80-9.
98. Manfredini D, Guarda-Nardini L, Winocur E, Piccotti F, Ahlberg J, Lobbezoo F. Research diagnostic criteria for temporomandibular disorders: A systematic review of axis i epidemiologic findings. *Oral Surgery, Oral Medicine, Oral Pathology, Oral Radiology and Endodontology* 2011;112(4):453-62.
99. Myburgh C, Larsen AH, Hartvigsen J. A Systematic, Critical Review of Manual Palpation for Identifying Myofascial Trigger Points: Evidence and Clinical Significance. *Archives of Physical Medicine and Rehabilitation* 2008;89(6):1169-76.
100. Naeije M, Kalaykova S, Visscher CM, Lobbezoo F. Evaluation of the Research Diagnostic Criteria for Temporomandibular Disorders for the recognition of an anterior disc displacement with reduction. *J Orofac Pain* 2009;23(4):303-11.
101. Petersson A. What you can and cannot see in TMJ imaging - an overview related to the RDC/TMD diagnostic system. *J Oral Rehabil* 2010;37(10):771-8.
102. Ribeiro-Rotta RF, Marques KDS, Pacheco MJ, Leles CR. Do computed tomography and magnetic resonance imaging add to temporomandibular joint disorder treatment? A systematic review of diagnostic efficacy. *J Oral Rehabil* 2011;38(2):120-35.
103. Saridin CP, Raijmakers PGHM, Tuinzing DB, Becking AG. Bone scintigraphy as a diagnostic method in unilateral hyperactivity of the mandibular condyles: a review and meta-analysis of the literature. *Int J Oral Maxillofac Surg* 2011;40(1):11-7.
104. Toscano P, Defabianis P. Clinical evaluation of temporomandibular disorders in children and adolescents: a review of the literature. *European journal of paediatric dentistry : official journal of European Academy of Paediatric Dentistry* 2009;10(4):188-92.
105. Trovato F, Orlando B, Bosco M. Occlusal features and masticatory muscles activity. A review of electromyographic studies. *Stomatologija / issued by public institution "Odontologijos studija"* 2009;[et(1):26-31.
106. Zito G, Morris ME, Selvaratnam. Characteristics of TMD headache a systematic review. *Physical Therapy Review* 2008;13(5):324-32.
107. Hussain AM, Packota G, Major PW, Flores-Mir C. Role of different imaging modalities in assessment of temporomandibular joint erosions and osteophytes: a systematic review. *Dento maxillo facial radiology* 2008;37(2):63-71.
108. Manfredini D, Guarda-Nardini L. Ultrasonography of the temporomandibular joint: a literature review. *Int J Oral Maxillofac Surg* 2009;38(12):1229-36.
109. Turp JC, Minagi S. Palpation of the lateral pterygoid region in TMD--where is the evidence? *Journal of Dentistry* 2001;29(7):475-83.



110. Turp JC, Arma T, Minagi S. Is the posterior belly of the digastric muscle palpable? A qualitative systematic review of the literature. *Clinical Anatomy* 2005;18(5):318-22.
111. Brooks SL, Westesson PL. Temporomandibular joint: value of coronal MR images. *Radiology* 1993;188(2):317-21.
112. Cavalcanti MG, Lew D, Ishimaru T, Ruprecht A. MR imaging of the temporomandibular joint: a validation experiment in vitro. *Acad Radiol* 1999;6(11):675-9.
113. Dolan EA, Kim HG, Nokes SR, Vogler JB, III. Correlation of magnetic resonance imaging and surgical findings in patients with meniscal perforation. *J Craniomandib Disord* 1989;3(3):174-8.
114. Emshoff R, Jank S, Bertram S, Rudisch A, Bodner G. Disk displacement of the temporomandibular joint: sonography versus MR imaging. *AJR Am J Roentgenol* 2002;178(6):1557-62.
115. Emshoff R, Brandlmaier I, Bodner G, Rudisch A. Condylar erosion and disc displacement: detection with high-resolution ultrasonography. *J Oral Maxillofac Surg* 2003;61(8):877-81.
116. Hansson LG, Westesson PL, Katzberg RW, Tallents RH, Kurita K, Holtas S, et al. MR imaging of the temporomandibular joint: comparison of images of autopsy specimens made at 0.3 T and 1.5 T with anatomic cryosections. *AJR Am J Roentgenol* 1989;152(6):1241-4.
117. Kaya K, Dulgeroglu D, Unsal-Delialioglu S, Babadag M, Tacal T, Barlak A, et al. Diagnostic value of ultrasonography in the evaluation of the temporomandibular joint anterior disc displacement. *J Craniomaxillofac Surg* 2010;38(5):391-5.
118. Kerstens HC, Golding RP, Valk J, van der Kwast WA. Magnetic resonance imaging of partial temporomandibular joint disc displacement. *J Oral Maxillofac Surg* 1989;47(1):25-9.
119. Manfredini D, Tognini F, Melchiorre D, Bazzichi L, Bosco M. Ultrasonography of the temporomandibular joint: comparison of findings in patients with rheumatic diseases and temporomandibular disorders. A preliminary report. *Oral Surg Oral Med Oral Pathol Oral Radiol Endod* 2005;100(4):481-5.
120. Rao VM, Farole A, Karasick D. Temporomandibular joint dysfunction: correlation of MR imaging, arthrography, and arthroscopy. *Radiology* 1990;174(3 Pt 1):663-7.
121. Ren YF, Westesson PL, Isberg A. Magnetic resonance imaging of the temporomandibular joint: value of pseudodynamic images. *Oral Surg Oral Med Oral Pathol Oral Radiol Endod* 1996;81(1):110-23.
122. Santler G, Karcher H, Simbrunner J. MR imaging of the TMJ. MR diagnosis and intraoperative findings. *J Craniomaxillofac Surg* 1993;21(7):284-8.
123. Tognini F, Manfredini D, Melchiorre D, Bosco M. Comparison of ultrasonography and magnetic resonance imaging in the evaluation of temporomandibular joint disc displacement. *J Oral Rehabil* 2005;32(4):248-53.
124. Uysal S, Kansu H, Akhan O, Kansu O. Comparison of ultrasonography with magnetic resonance imaging in the diagnosis of temporomandibular joint internal derangements: a preliminary investigation. *Oral Surg Oral Med Oral Pathol Oral Radiol Endod* 2002;94(1):115-21.
125. Watt-Smith S, Sadler A, Baddeley H, Renton P. Comparison of arthrotomographic and magnetic resonance images of 50 temporomandibular joints with operative findings. *Br J Oral Maxillofac Surg* 1993;31(3):139-43.

126. Yamada I, Murata Y, Shibuya H, Suzuki S. Internal derangements of the temporomandibular joint: comparison of assessment with three-dimensional gradient-echo and spin-echo MRI. *Neuroradiology* 1997;39(9):661-7.

---

# Vedlegg

---

## Vedlegg 1 - Søkestrategier

---

### A. PRESIST SØK

Søket ble opprinnelig designet for å finne valideringsstudier (primærstudier), men ble til slutt avgrenset til systematiske oversikter. Søket var sammensatt av følgende elementer: TMD + diagnostikk eller utredning + valideringsstudier + presist filter for systematiske oversikter.

Søketreff totalt etter dublettkontroll: 58  
Antall treff i hver database ble ikke registrert.

### Ovid MEDLINE, EMBASE, Amed, PsycINFO

Dato: 19.9.2012

1. exp Temporomandibular Joint Disorders/ use prmz or Craniomandibular Disorders/ use prmz
2. \*temporomandibular joint disorder/ use emez
3. exp temporomandibular joint disease/ use amed
4. (Temp?romandibular\* or temp?ro mandibular\* or (tmj adj (disorder\* or disease\* or syndrom\*))).tw.
5. \*Temporomandibular Joint/ use prmz
6. \*temporomandibular joint/ use emez
7. \*Temporomandibular joint/ use amed
8. (craniomandibular\* or cranio mandibular\*).tw.
9. or/1-8
10. diagnosis/ use prmz or exp "diagnostic techniques and procedures"/ use prmz or Physical Examination/ use prmz or Diagnosis, Differential/ use prmz
11. exp \*physical examination/ use emez or exp \*diagnostic test/ use emez or exp \*diagnostic procedures/ use emez or \*diagnostic procedure/ use emez or \*differential diagnosis/ use emez
12. diagnosis/ use amed or exp "diagnostic techniques and procedures"/ use amed or patient assessment/ use amed or exp Physical examination/ use amed or Clinical assessment scales/ use amed
13. diagnosis/ use psych or differential diagnosis/ use psych or physical examination/ use psych or Screening/ use psych
14. (examination\* or screening or (evaluation adj2 patient\*) or assessment\* or assessing or "to assess" or diagnostic\* or diagnosing or "to diagnose" or clinical evaluat\*).tw.
15. (RDC?TMD or RDCTMD or RDC TMD or Research Diagnostic\* Criteria\*).tw.
16. (CDC?TMD or CDCTMD or CDC TMD or Clinical\* Diagnostic\* Criteria\*).tw.
17. (Helkimo\* or Cinical dysfunction index or Clinical\* articular\* disc\* derangement\* or Wilkes).tw.

18. (diagnostic adj2 (criteria\* or value\*)).tw.
19. ((clinical adj (test\* or scale\* or subscale\*)) or TMJ scale\*).tw.
20. or/10-19
21. 9 and 20
22. exp Temporomandibular Joint Disorders/di use prmz or exp \*Temporomandibular Joint Disorders/di use emez
23. 21 or 22
24. "reproducibility of results"/ use prmz or "sensitivity and specificity"/ use prmz or validation studies/ use prmz
25. exp diagnostic accuracy/ use emez or exp diagnostic test accuracy study/ use emez or exp diagnostic value/ use emez or exp instrument validation/ use emez or exp "sensitivity and specificity"/ use emez or exp reproducibility/ use emez or validation study/ use emez
26. test reliability/ use psych or interrater reliability/ use psych or test validity/ use psych or psychometrics/ use psych
27. (validat\* or validity or reliability or accuracy or sensitiv\* or specificit\*).tw.
28. (kappa or inter-rater reliability or inter-rater agreement or interrater reliability or interrater agreement).tw.
29. (false negative\* or false positive\* or discriminative capacit\*).tw.
30. or/24-29
31. 23 and 30
32. (animals/ or animal/ or nonhuman/ or animal experiment/) not (human/ or humans/)
33. (rats or animal\*).ti.
34. 32 or 33
35. 31 not 34
36. 35 use emez
37. limit 36 to embase
38. 35 use prmz,amed,psych
39. 37 or 38
40. remove duplicates from 39
41. ((systematic\* adj2 review\*) or meta-analysis).pt,mp.
42. 40 and 41

## **Cochrane Library (Cochrane Database of Systematic Reviews, DARE, HTA)**

Dato: 19.9.2012

- #1 (Temp?romandibular\* or (temp?ro next mandibular\*)):ti,ab,kw
- #2 (tmj next (disorder\* or disease\* or syndrom\*)):ti,ab,kw
- #3 MeSH descriptor: [Temporomandibular Joint Disorders] explode all trees
- #4 MeSH descriptor: [Temporomandibular Joint] explode all trees
- #5 MeSH descriptor: [Craniomandibular Disorders] explode all trees
- #6 (craniomandibular\* or (cranio next mandibular\*)):ti,ab,kw
- #7 #1 or #2 or #3 or #4 or #5 or #6
- #8 MeSH descriptor: [Diagnosis] this term only
- #9 MeSH descriptor: [Diagnostic Techniques and Procedures] explode all trees
- #10 MeSH descriptor: [Diagnosis, Differential] explode all trees
- #11 (examination\* or (evaluat\* near/2 patient\*) or (clinical\* next evaluation\*) or assessment\* or assessing or "to assess" or diagnostic\* or diagnosing or "to diagnose" or screening\*):ti,ab,kw
- #12 (RDC?TMD or RDCTMD or (RDC next TMD) or (Research next Diagnostic\* next Criteria\* next for next Temp?romandibular next Disorder\*) or (Research next Diagnostic\* next Criteria\* next for next TMD\*)):ti,ab,kw
- #13 (CDC?TMD or CDCTMD or (CDC next TMD) or (Clinical\* next Diagnostic\* next Criteria\* next for next Temp?romandibular next Disorder\*) or (Clinical\* next Diagnostic\* next Criteria\* next for next TMD\*)):ti,ab,kw

- #14 (Helkimo\* or (Cinical next dysfunction next index) or (Clinical\* next articular\* next disc\* next derangement\*) or Wilkes):ti,ab,kw
- #15 (diagnostic\* near/2 (criteria\* or value\*)):ti,ab,kw or (differential diagnos\*):ti,ab,kw
- #16 (tests or scale\* or subscale\* or instrument\*):ti,ab,kw
- #17 #8 or #9 or #10 or #11 or #12 or #13 or #14 or #15 or #16
- #18 #7 and #17
- #19 MeSH descriptor: [Temporomandibular Joint Disorders] explode all trees and with qualifiers: [Diagnosis - DI]
- #20 #18 or #19
- #21 MeSH descriptor: [Reproducibility of Results] explode all trees
- #22 MeSH descriptor: [Sensitivity and Specificity] explode all trees
- #23 MeSH descriptor: [Validation Studies as Topic] explode all trees
- #24 MeSH descriptor: [Validation Studies] explode all trees
- #25 (validat\* or validity or reliability or accuracy or sensitiv\* or specificit\*):ti,ab,kw
- #26 (kappa or "inter-rater reliability" or "inter-rater agreement" or "interrater reliability" or "interrater agreement"):ti,ab,kw
- #27 #21 or #22 or #23 or #24 or #25 or #26

### PubMed

Dato: 19.9.2012

Søket ble avgrenset til referanser registrert som "Epub ahead of print".

(temporomandibular OR craniomandibular) AND (diagnostic OR diagnose) AND (validity OR validation OR accuracy OR reliability OR specificity) AND (systematic review or meta-analysis) publisher[sb]

### CRD Databases

Dato: 19.9.2012

- 1 MeSH DESCRIPTOR Temporomandibular Joint Disorders EXPLODE ALL TREES
- 2 MeSH DESCRIPTOR Temporomandibular Joint EXPLODE ALL TREES
- 3 MeSH DESCRIPTOR Craniomandibular Disorders EXPLODE ALL TREES
- 4 Temporomandibular or craniomandibular or "cranio mandibular"
- 5 MeSH DESCRIPTOR Diagnostic Techniques and Procedures EXPLODE ALL TREES
- 6 MeSH DESCRIPTOR Diagnosis EXPLODE ALL TREES
- 7 MeSH DESCRIPTOR Physical Examination EXPLODE ALL TREES
- 8 diagnostic\*
- 9 #1 OR #2 OR #3 OR #4
- 10 #5 OR #6 OR #7 OR #8
- 11 #9 AND #10

### ISI web of Science

Dato: 19.9.2012

Topic=(temporomandibular or craniomandibular or "cranio mandibular") AND Topic=(diagnostic\*) AND Topic=(validation or validity or reliability or accuracy or sensitiv\* or specificit\*) AND Topic=(systematic review or meta-analysis)

Databases=Science Citation Index, Social Sciences Citation Index -1975-present

Timespan=All Years

Lemmatization=Off

## **B. BREDT SØK**

Søket var sammensatt av følgende elementer: TMD og tilgrensende diagnoser + filter for systematiske oversikter.

Søketreff totalt: 615

Søketreff etter dublettkontroll: 453

I Reference Manager ble alle referanser som var publisert før 2000 slettet. Antall referanser med publikasjonsår f.o.m. 2000: 406

### **PubMed**

Dato: 25.4.2012

Søketreff: 5

Søk etter artikler registrert som "epub ahead of print"

(temporomandibular joint disorder OR temporomandibular OR temperomandibular OR craniomandibular disorder OR craniomandibular OR tmj OR tmd OR cmd) AND (systematic review OR meta-analysis OR (review AND (pubmed OR medline OR embase))) AND publisher [sb]

### **CRD Databases (HTA, DARE, NHS EED)**

Dato: 25.4.2012

Søketreff: 45

Quick search: temporomandibular joint disorder OR craniomandibular disorder OR tmj OR tmd OR cmd

### **Prospero** (pågående systematiske oversikter)

Dato: 25.4.2012

Søketreff: 0

All fields: temporomandibular OR craniomandibular OR tmj OR tmd OR cmd

### **Cochrane Library**

Dato: 25.4.2012

Søketreff: Cochrane Reviews 14, DARE 34, HTA 4, Economic evaluations 2

- #1 MeSH descriptor Temporomandibular Joint Disorders explode all trees
- #2 MeSH descriptor Craniomandibular Disorders explode all trees
- #3 ((Temp?romandibular or (temp?ro next mandibular\*) or craniomandibular\* or (cranio next mandibular\*) or tmj or tmd or cmd)):ti,ab,kw
- #4 (((myofascial or myofacial) next pain) and (jaw\* or face\* or facial\*)):ti,ab,kw
- #5 (Costen\* next Syndrome):ti,ab,kw
- #6 MeSH descriptor Facial Neuralgia explode all trees
- #7 (((facial next neuralgia\*) or ((craniofascial or craniofacial or "cranio fascial" or "cranio facial") next (dis\* or syndrom\* or pain)))):ti,ab,kw
- #8 MeSH descriptor Temporomandibular Joint explode all trees
- #9 (((jaw\* or mandibular) next dislocation\*) or "locked jaw"):ti,ab,kw
- #10 (#1 OR #2 OR #3 OR #4 OR #5 OR #6 OR #7 OR #8 OR #9)
- #11 <nothing>, from 2000 to 2012
- #12 (#10 AND #11)

## **Ovid MEDLINE, EMBASE, Amed, PsycINFO**

Dato: 25.4.2012

AMED (Allied and Complementary Medicine) 1985 to April 2012

Embase 1980 to 2012 Week 16

Ovid MEDLINE In-Process & Other Non-Indexed Citations and Ovid MEDLINE 1946 to Present

Ovid PsycInfo 1806 to April Week 3 2012

Søketreff: 378

1. exp Temporomandibular Joint Disorders/ use prmz or Craniomandibular Disorders/ use prmz
2. temporomandibular joint disorder/ use emez or myofascial pain/ use emez
3. exp temporomandibular joint disease/ use amed
4. (Temp?romandibular\* or temp?ro mandibular\* or craniomandibular\* or cranio mandibular\* or myofacial pain or tmj\* or tmd\* or cmd\*).tw.
5. Costen\* Syndrome.tw.
6. Facial Neuralgia/ use prmz
7. (facial neuralgia\* or ((craniofacial or cranio facial) adj2 (dis\* or syndrom\* or pain))).tw.
8. exp \*Temporomandibular Joint/ use prmz
9. \*temporomandibular joint/ use emez
10. Myofascial Pain Syndromes/ use amed
11. or/1-10
12. ((systematic\* adj2 review\*) or meta-analy\*).pt,mp.
13. review.pt. and (medline or embase or pubmed or search\* database\* or systematic\* search\*).tw.
14. 12 or 13
15. 11 and 14
16. remove duplicates from 15

## **BMJ Clinical Evidence**

Dato: 25.4.2012

Søkt via BMJ Best Practice. Ingen informasjon under "Temporomandibular joint syndrome"

## **UptoDate**

Dato: 25.4.2012

En artikkel: "Temporomandibular joint dysfunction syndrome"

[http://www.uptodate.com/contents/temporomandibular-joint-dysfunction-synrome?source=search\\_result&search=temporomandibular+joint+pain&selectedTitle=1~150](http://www.uptodate.com/contents/temporomandibular-joint-dysfunction-synrome?source=search_result&search=temporomandibular+joint+pain&selectedTitle=1~150)

## **ISI Science Citation Index**

Dato: 25.4.2012

Søketreff: 117

Topic=(temporomandibular OR craniomandibular OR temperomandibular OR tempuromandibular OR tmj OR tmd OR cmd) AND Topic=("systematic review" or "meta-analysis")

Databases=SCI-EXPANDED, SSCI, A&HCI Timespan=All Years

Lemmatization=Off

**CINAHL**

Dato: 25.4.2012

Søketreff: 15

S10 S4 and S8 Limiters - Exclude MEDLINE records  
S9 S4 and S8  
S8 S5 or S6 or S7  
S7 TI ( "systematic review" or "meta-analysis" ) OR AB ( "systematic review" or "meta- analysis" )  
S6 (MH "Meta Analysis")  
S5 (MH "Systematic Review")  
S4 S1 or S2 or S3  
S3 TI ( temporomandibular or tempuromandibular or temperomandibular or temporo mandibular or tempuro mandibular or temporo mandibular or craniomandibular or cranio mandibular or tmj or cmd or tmd ) OR AB ( temporomandibular or tempuromandibular or craniomandibular or tmj or cmd or tmd )  
S2 (MH "Craniomandibular Disorders+")  
S1 (MH "Temporomandibular Joint Syndrome") OR (MH "Temporomandibular Joint Diseases+") OR (MH "Temporomandibular Joint")



---

## Vedlegg 2 - Sjekkliste for vurdering av metodisk kvalitet

---

### 1) Er kriteriene for inklusjon av studier klare?

- Ble populasjon, tiltak, kontroll og utfallsmål presisert?
- Ble relevante studiedesign inkludert?

### 2) Ble det utført et tilfredsstillende litteratursøk?

- Ble det søkt i relevante databaser?
- Var tidsrommet for søkene oppgitt?
- Var benyttede søkeord oppgitt?
- Ble referanselister i de inkluderte studiene gjennomført?
- Ble ressurspersoner på fagfeltet kontaktet?
- Ble det søkt etter upublisert materiale?
- Ble det søkt etter studier publisert på andre språk enn engelsk?

### 3) Ble det sikret mot systematiske skjevheter (bias) ved seleksjon av studier?

- Ble eksplisitte seleksjonskriterier brukt?
- Ble seleksjon av studier utført av minst to personer uavhengig av hverandre?
- Er inkluderte studier listet opp?
- Er ekskluderte studier listet opp?
- Er seleksjonsprosessen beskrevet med diagram/figur?

### 4) Ble risiko for systematiske feil (bias) på de inkluderte studiene vurdert ved bruk av relevante kriterier?

- Ble det presisert hvilke kriterier man har benyttet for å vurdere risiko for systematiske feil?
- Var en slik vurdering foretatt av minst to personer uavhengig av hverandre?
- Er slike vurderinger oppsummert for hver inkluderte studie og hvert enkelt kriterium?
- Ble manglende informasjon i publiserte studier hentet inn?

### 5) Er resultatene fra studiene sammenfattet på en tilfredsstillende måte?

- Ble dataekstraksjonen utført av minst to personer uavhengig av hverandre?
- Er de inkluderte studiers karakteristika med hensyn til populasjon, tiltak, kontroll og endepunkter oppsummert i tabeller?
- Er resultater av de inkluderte studiene oppsummert i tabeller?
- Er eventuelle variasjoner i resultatene forklart og diskutert?

### 6) Er metodene som ble brukt da resultatene ble sammenfattet, klart beskrevet?

- Hvordan ble dataene tolket (f.eks. deskriptiv, "vote counting" basert på tendens eller statistisk signifikans, beskrivelse av effektstørrelser, metaanalyse, metaregresjon)?
- Hvordan ble studiene vektlagt i analysen (f. eks. lik vekt ("ved vote counting"), kvaliteten eller studiedesignet, invers varians (meta-analyse), antall deltakere)?
- Hvis en meta-analyse ble utført, er populasjonene og tiltakene i studiene "like nok" til å slå sammen?

### 7) Er graden av heterogenitet beskrevet?

- Ble det sørget for at inkluderte studier var "like nok" til å slå sammen eller ble fornuftig oppdelt i homogene grupper, eller at det ble konkludert med at det ikke var fornuftig å slå sammen de inkluderte studiene?
- Er graden av viktige forskjeller i resultatene til studiene diskutert?
- Hvis en meta-analyse ble utført, er  $I_2$ , chi square test for heterogenitet eller annen statistikk rapportert?

## **8) Er forfatterens konklusjoner støttet av data og/eller analysen som er rapportert i oversikten?**

### **Samlet vurdering av hver systematiske oversikt er "høy", "moderat" eller "mangelfull" med følgende til grunn:**

- Høy kvalitet: Brukes hvis alle eller de fleste kriteriene fra sjekklisten er oppfylt. Dersom noen av kriteriene ikke er oppfylt, må det være veldig lite sannsynlig at studiens konklusjon blir påvirket.
- Moderat kvalitet: Brukes hvis noen av kriteriene fra sjekklisten ikke er oppfylt og/eller der kriteriene ikke er tilfredsstillende beskrevet. Samlet vurdering tilsier at det er lite sannsynlig at studiens konklusjon påvirkes.
- Mangelfull: Brukes hvis få eller ingen kriterier i sjekklisten er oppfylt og/eller ikke er tilfredsstillende beskrevet. Samlet vurdering tilsier at det er sannsynlig at studiens konklusjon kan forandres.

#### **Denne sjekklisten er utformet ved hjelp av:**

1. Sjekkliste for systematiske oversikter fra Håndboka 2009
2. SURE checklist (adapted from the checklist used to prepare SUPPORT summaries (<http://www.support-collaboration.org/summaries.htm>) and revised to incorporate additional considerations from the AMSTAR checklist)
3. Sjekkliste for vurdering av en oversiktsartikkel (basert på Guyatt G, Rennie D. The Evidence-Based Medicine Working Group. Users' Guides to the Medical literature. JAMA Archives Journals, AMA Press, 2002. og Critical Appraisal Skills Programme (<http://www.phru.nhs.uk/Pages/PHD/CASP.htm>))

### Vedlegg 3 - Ekskluderte studier

#### *Oversikt over ekskluderte studier innhentet i fulltekst*

| Førsteforfatter/<br>Årstall | Diagnostisk test som evalueres   | Eksklusjonsgrunn          |
|-----------------------------|--|---------------------------|
| Cacli 2012 (91)             | Cone beam computed tomography (CBCT)                                   | Ikke systematisk oversikt |
| Cummings 2007 (92)          | Triggerpunkter   | Ikke systematisk oversikt |
| De Senna 2009 (93)          | Radiologiske undersøkelser   | Ikke systematisk oversikt |
| De Vos 2009 (94)            | -  | Prevalensstudie           |
| Klasser 2006 (95)           | Elektromyografi (EMG)  | Ikke systematisk oversikt |
| Lobbezoo 2004 (96)          | Ulike typer  | Ikke systematisk oversikt |
| Lucas 2009 (97)             | -  | Ikke diagnostikk av TMD   |
| Manfredini 2011 (12)        | Ulike typer  | Oversikt over oversikter  |
| Manfredini 2011 (98)        | -  | Prevalensstudie           |
| Manfredini 2010 (7)         | -  | Ikke diagnostikk av TMD   |
| Myburgh 2008 (99)           | -  | Ikke diagnostikk av TMD   |
| Naeiji 2009 (100)           | Research Diagnostic Criteria for Temporomandibular Disorders (RDC/TMD) | Ikke systematisk oversikt |
| Peterson 2010 (101)         | Radiologiske undersøkelser   | Ikke systematisk oversikt |
| Ribeiro-Rotta 2011 (102)    | Radiologiske undersøkelser   | Oversikt over oversikter  |
| Saridin 2011 (103)          | -  | Ikke diagnostikk av TMD   |
| Suvinen 2007 (11)           | -  | Ikke diagnostikk av TMD   |
| Toscano 2009 (104)          | -  | Prevalensstudie           |
| Trovato 2009 (105)          | EMG  | Ikke systematisk oversikt |
| Zito 2008 (106)             | -  | Ikke diagnostikk av TMD   |

#### *Systematiske oversikter ekskludert etter vurdering av metodologisk kvalitet*

|  |                               |                           |
|--|-------------------------------|---------------------------|
| Hussain 2008 (107)                     | Radiologiske undersøkelser    | Lav metodologisk kvalitet |
| Manfredini & Guarda-Nardini 2009 (108) | Ultralyd                      | Lav metodologisk kvalitet |
| Türp & Minagi 2001 (109)               | Muskelpalpasjon (m.pterygoid) | Lav metodologisk kvalitet |
| Türp 2005 (110)                        | Muskelpalpasjon (m.digastric) | Lav metodologisk kvalitet |

## Vedlegg 4 - Primærstudier som inngår i inkluderte systematiske oversikter

| Systematisk oversikt           | Al-Saleh<br>2012 | Chaput<br>2012 | Li<br>2012 | Reneker<br>2011 | Koh<br>2009 | Limchaichana<br>2006 | Overlapp<br>(n) |
|--------------------------------|------------------|----------------|------------|-----------------|-------------|----------------------|-----------------|
| Primærstudier                  |                  |                |            |                 |             |                      |                 |
| Adame 1998 (40)                |                  |                |            |                 | x           |                      |                 |
| Barclay 1999 (27)              |                  | x              |            |                 | x           |                      | 2               |
| Bertram 2001 (28)              |                  | x              |            |                 | x           |                      | 2               |
| Brandlmaier 2003 (68)          |                  |                | x          |                 |             |                      |                 |
| Brooks 1993 (111)              |                  |                |            |                 |             | x                    |                 |
| Cavalcanti 1999 (112)          |                  |                |            |                 |             | x                    |                 |
| Dolan 1989 (113)               |                  |                |            |                 |             | x                    |                 |
| Emshoff 1997 (69)              |                  |                | x          |                 |             |                      |                 |
| Emshoff 2001 (42)              |                  |                |            |                 | x           |                      |                 |
| Emshoff 2001 (41)              |                  |                |            |                 | x           |                      |                 |
| Emshoff 2002 (29)              |                  | x              |            |                 |             |                      |                 |
| Emshoff 2002 (71)              |                  |                | x          |                 |             |                      |                 |
| Emshoff 2002 (114)             |                  |                | x          |                 |             |                      |                 |
| Emshoff 2002 (70)              |                  |                | x          |                 |             |                      |                 |
| Emshoff 2002 (43)              |                  |                |            |                 | x           |                      |                 |
| Emshoff 2003 (115)             |                  |                | x          |                 |             |                      |                 |
| Emshoff 2003 (44)              |                  |                |            |                 | x           |                      |                 |
| Emshoff 2003 (45)              |                  |                |            |                 | x           |                      |                 |
| Glaros 1989 (76)               | x                |                |            |                 |             |                      |                 |
| Glaros 1997 (77)               | x                |                |            |                 |             |                      |                 |
| Guler 2003 (46)                |                  |                |            |                 | x           |                      |                 |
| Haley 2001 (47)                |                  |                |            |                 | x           |                      |                 |
| Hansson 1989 (116)             |                  |                |            |                 |             | x                    |                 |
| Holmlund og Axelsson 1996 (35) |                  |                |            | x               |             |                      |                 |
| Huddleston 2004 (48)           |                  |                |            |                 | x           |                      |                 |
| Jank 2001 (72)                 |                  |                | x          |                 |             |                      |                 |
| Jank 2005 (73)                 |                  |                | x          |                 |             |                      |                 |
| Katzberg 1988 (62)             |                  |                |            |                 |             | x                    |                 |
| Katzberg 1996 (49)             |                  |                |            |                 | x           |                      |                 |
| Kaya 2010 (117)                |                  |                | x          |                 |             |                      |                 |
| Kerstens 1989 (118)            |                  |                |            |                 |             | x                    |                 |
| Landes 2006 (74)               |                  |                | x          |                 |             |                      |                 |
| Landes 2007 (75)               |                  |                | x          |                 |             |                      |                 |
| Larheim 1999 (67)              |                  |                |            |                 |             | x                    |                 |
| Limchaichana 2007 (50)         |                  |                |            |                 | x           |                      |                 |
| Manfredini 2003 (36)           |                  |                |            | x               |             |                      |                 |
| Manfredini 2005 (119)          |                  |                | x          |                 |             |                      |                 |
| Manfredini 2008 (30)           |                  | x              |            | x               |             |                      | 2               |
| Marguelles-Bonnet 1995 (31)    |                  | x              |            |                 | x           |                      | 2               |

|  |   |   |    |   |    |    |   |
|--|---|---|----|---|----|----|---|
| Nebbe 2000 (58)                            |   |   |    |   |    | x  |   |
| Orsini 1997 (59)                           |   |   |    |   |    | x  |   |
| Orsini 1998 (60)                           |   |   |    |   |    | x  |   |
| Orsini 1999 (37)                           |   |   |    | x | x  |    | 2 |
| Paesani 1992 (51)                          |   |   |    |   | x  |    |   |
| Rammelsberg 1997 (52)                      |   |   |    |   | x  |    |   |
| Rao 1990 (120)                             |   |   |    |   |    | x  |   |
| Ren 1996 (121)                             |   |   |    |   |    | x  |   |
| Ribeiro 1997 (53)                          |   |   |    |   | x  |    |   |
| Rudisch 2001 (32)                          |   | x |    |   | x  |    | 2 |
| Sano 1995 (54)                             |   |   |    |   | x  |    |   |
| Santler 1993 (122)                         |   |   |    |   |    | x  |   |
| Schmitter 2004 (61)                        |   |   |    |   |    | x  |   |
| Schwaighofer 1990 (63)                     |   |   |    |   |    | x  |   |
| Shaefer 2001 (55)                          |   |   |    |   | x  |    |   |
| Shaidi 2008                                |   |   | x  |   |    |    |   |
| Stegenga 1992 (38)                         |   |   |    | x |    |    |   |
| Takahashi 1999 (56)                        |   |   |    |   | x  |    |   |
| Tasaki 1993 (64)                           |   |   |    |   |    | x  |   |
| Taskaya-Yilmaz og Ogutcen-Toller 2002 (33) | x |   |    |   |    |    |   |
| Tognini 2005 (123)                         |   |   | x  |   |    |    |   |
| Üşümez 2004 (34)                           |   | x |    | x | x  |    | 3 |
| Uysal 2002 (124)                           |   |   | x  |   |    |    |   |
| Watt-Smith 1993 (125)                      |   |   |    |   |    | x  |   |
| Westesson 1987 (65)                        |   |   |    |   |    | x  |   |
| Westesson 1987 (66)                        |   |   |    |   |    | x  |   |
| Yamada 1997 (126)                          |   |   |    |   |    | x  |   |
| Yatani 1998 (39)                           |   |   |    |   | x  |    |   |
| Antall                                     | 2 | 8 | 15 | 7 | 23 | 20 |   |

## Vedlegg 5 – Tabeller, diagnostisk nøyaktighet

**Tabell 6.** Påvisning av diskusforskyvning med tilbakegang ved bruk av ulike kliniske enkelttester eller test-kluster som indekstest, sammenliknet med referansestandard MR.

| Studie                            | Pretest sannsynlighet | Antall ledd | PPV              | NPV              | LR+              | LR-              |
|-----------------------------------|-----------------------|-------------|------------------|------------------|------------------|------------------|
| <b>Klikk</b>                      |                       |             |                  |                  |                  |                  |
| <i>Chaput 2012</i>                |                       |             |                  |                  |                  |                  |
| Manfredini, 2008 <sup>1</sup>     | 0,35                  | 255         | 0,42 (0,32-0,57) | 0,70 (0,62-0,77) | 1,34 (0,98-1,83) | 0,82 (0,66-1,03) |
| Uşümez, 2004 <sup>2</sup>         | 0,60                  | 63          | 0,63 (0,49-0,75) | 0,56 (0,23-0,85) | 1,12 (0,89-1,40) | 0,53 (0,16-1,77) |
| Barclay, 1999 <sup>1</sup>        | 0,78                  | 50          | 0,89 (0,73-0,99) | 0,50 (0,24-0,76) | 2,26 (1,02-5,00) | 0,28 (0,13-0,63) |
| <i>Reneker 2011</i>               |                       |             |                  |                  |                  |                  |
| Manfredini, 2008 <sup>3</sup>     | -                     | 388         | -                | -                | 1,12             | 0,92             |
| Orsini, 1999 <sup>3</sup>         | -                     | -           | -                | -                | 3,01             | 0,59             |
| Stegenga, 1992 <sup>3</sup>       | -                     | -           | -                | -                | 7,10             | 0,32             |
| Stegenga, 1992 <sup>4</sup>       | -                     | -           | -                | -                | 15,20            | 0,25             |
| Uşümez, 2004 <sup>3</sup>         | -                     | -           | -                | -                | 1,50             | 0,26             |
| <b>Krepitasjon</b>                |                       |             |                  |                  |                  |                  |
| <i>Chaput 2012</i>                |                       |             |                  |                  |                  |                  |
| Uşümez, 2004                      | 0,60                  | 63          | 0,57 (0,20-0,88) | 0,39 (0,27-0,53) | 0,88 (0,21-3,59) | 1,02 (0,85-1,22) |
| <i>Reneker 2011</i>               |                       |             |                  |                  |                  |                  |
| Uşümez, 2004                      | -                     | -           | -                | -                | 0,29             | 1,39             |
| Yatani, 1998                      | -                     | -           | -                | -                | 0,21             | 1,08             |
| <b>Smerte</b>                     |                       |             |                  |                  |                  |                  |
| <i>Chaput 2012</i>                |                       |             |                  |                  |                  |                  |
| Uşümez, 2004 <sup>1</sup>         | 0,60                  | 63          | 0,66 (0,52-0,77) | 1,00 (0,46-1,00) | 1,25 (1,03-1,52) | NA               |
| Bertram, 2001 <sup>2</sup>        | 0,46                  | -           | -                | -                | 1,32 (0,90-1,95) | 0,84 (0,65-1,08) |
| Rudisch, 2001 <sup>2</sup>        | 0,36                  | 47          | 0,53 (0,29-0,76) | 0,73 (0,54-0,87) | 1,99 (0,94-4,18) | 0,64 (0,37-1,11) |
| Taskaya-Yylmaz, 2002 <sup>3</sup> | 1,00                  | 131         | -                | -                | -                | -                |
| Taskaya-Yylmaz, 2002 <sup>4</sup> | 1,00                  | 131         | -                | -                | -                | -                |
| <i>Reneker 2011</i>               |                       |             |                  |                  |                  |                  |
| Uşümez, 2004 <sup>1</sup>         | -                     | -           | -                | -                | 1,14             | 0,00             |
| Stegenga, 1992 <sup>1</sup>       | -                     | -           | -                | -                | 0,64             | 1,51             |
| Stegenga, 1992 <sup>5</sup>       | -                     | -           | -                | -                | 0,64             | 1,81             |
| Yatani, 1998 <sup>5</sup>         | -                     | -           | -                | -                | 0,56             | 1,19             |
| Stegenga, 1992 <sup>6</sup>       | -                     | -           | -                | -                | 0,66             | 1,83             |
| Orsini, 1999 <sup>7</sup>         | -                     | -           | -                | -                | 1,04             | 0,84             |

<sup>1</sup>Muskel- og kjeveleddssmerte ved palpasjon, <sup>2</sup>Kjeveleddssmerte ved palpasjon/funksjon/åpning, <sup>3</sup>kjeveleddssmerte uten nærmere spesifisering, <sup>4</sup>øresmerte, <sup>5</sup>kjeveleddssmerte ved åpning, <sup>6</sup>kjeveleddssmerte ved assistert maksimal åpning, <sup>7</sup>ingen kjeveleddssmerte ved assistert maksimal åpning. PPV; Positiv prediktiv verdi, NPV; negativ prediktiv verdi, LR+; Positiv sannsynlighetsratio (likelihood ratio), LR-; negativ sannsynlighetsratio.

**Tabell 7. Påvisning av diskusforskyvning uten tilbakegang ved bruk av ulike kliniske enkelttester og test-kluster som indekstest, sammenliknet med referansestandard MR.**

| Studie   | Pretest sannsynlighet | Antall ledd | PPV              | NPV              | LR+               | LR-              |
|--|-----------------------|-------------|------------------|------------------|-------------------|------------------|
| <b>Klikk</b>   |                       |             |                  |                  |                   |                  |
| <i>Chaput 2012</i>   |                       |             |                  |                  |                   |                  |
| Manfredini, 2008 <sup>1</sup>  | 0,45                  | 298         | 0,54 (0,45-0,63) | 0,61 (0,54-0,69) | 1,44 (1,09-1,90)  | 0,77 (0,63-0,94) |
| Uşümez, 2004 <sup>2</sup>  | 0,40                  | 42          | 0,20 (0,08-0,41) | 0,29 (0,11-0,56) | 0,37 (0,17-0,79)  | 3,53 (1,52-8,19) |
| Barclay, 1999 <sup>1</sup>   | 0,62                  | 29          | -                | -                | 1,53 (0,63-3,70)  | 0,70 (0,35-1,38) |
| <i>Reneker 2011</i>  |                       |             |                  |                  |                   |                  |
| Manfredini, 2008 <sup>3</sup>  | -                     | -           | -                | -                | 1,28              | 0,83             |
| Stegenga, 1992 <sup>3</sup>  | -                     | -           | -                | -                | 0,17              | 2,25             |
| Stegenga, 1992 <sup>4</sup>  | -                     | -           | -                | -                | 0,08              | 1,85             |
| Uşümez, 2004 <sup>3</sup>  | -                     | -           | -                | -                | 0,34              | 4,93             |
| <sup>1</sup> Klikk ved palpasjon, <sup>2</sup> klikk ved lytting (auskulasjon), <sup>3</sup> klikk ikke nærmere definert, <sup>4</sup> resiprokalt klikk |                       |             |                  |                  |                   |                  |
| <b>Krepitasjon</b>   |                       |             |                  |                  |                   |                  |
| <i>Chaput 2012</i>   |                       |             |                  |                  |                   |                  |
| Uşümez, 2004   | 0,40                  | 42          | 0,80 (0,51-0,95) | 0,81 (0,61-0,93) | 5,88 (1,95-17,76) | 0,33 (0,16-0,71) |
| <i>Reneker 2011</i>  |                       |             |                  |                  |                   |                  |
| Uşümez, 2004   | -                     | -           | -                | -                | 6,36              | 0,33             |
| Yatani, 1998   | -                     | -           | -                | -                | 5,30              | 0,86             |
| <b>Smerte</b>  |                       |             |                  |                  |                   |                  |
| <i>Chaput 2012</i>   |                       |             |                  |                  |                   |                  |
| Uşümez, 2004 <sup>1</sup>  | 0,40                  | 42          | 0,46 (0,30-0,63) | 1,00 (0,46-1,00) | 1,25 (1,03-1,52)  | -                |
| Taskaya-Yylmaz, 2002 <sup>3</sup>  | 1,00                  | -           | -                | -                | -                 | -                |
| Taskaya-Yylmaz, 2002 <sup>4</sup>  | 1,00                  | -           | -                | -                | -                 | -                |
| Rudisch 2001 <sup>2</sup>  | 0,54                  | 65          | 0,75 (0,56-0,88) | 0,67 (0,48-0,81) | 2,57 (1,36-4,85)  | 0,43 (0,25-0,73) |
| Bertram 2001 <sup>2</sup>  | 0,54                  | 188         | 0,70 (0,60-0,79) | 0,63 (0,52-0,73) | 2,01 (1,45-2,78)  | 0,49 (0,25-0,67) |
| <i>Reneker 2011</i>  |                       |             |                  |                  |                   |                  |
| Uşümez, 2004 <sup>1</sup>  | -                     | -           | -                | -                | 1,09              | 0,00             |
| Stegenga, 1992 <sup>1</sup>  | -                     | -           | -                | -                | 2,00              | 0,51             |
| Stegenga, 1992 <sup>5</sup>  | -                     | -           | -                | -                | 1,72              | 0,46             |
| Yatani, 1998 <sup>5</sup>  | -                     | -           | -                | -                | 4,90              | 0,47             |
| Stegenga, 1992 <sup>6</sup>  | -                     | -           | -                | -                | 1,81              | 0,39             |
| Orsini, 1999 <sup>6</sup>  | -                     | -           | -                | -                | 6,02              | 0,49             |

<sup>1</sup>Muskel- og kjeveleddssmerter ved palpasjon, <sup>2</sup>Kjeveleddssmerter ved palpasjon/funksjon/åpning, <sup>3</sup>kjeveleddssmerter uten nærmere spesifisering, <sup>4</sup>øresmerter, <sup>5</sup>kjeveleddssmerter ved åpning, <sup>6</sup>kjeveleddssmerter ved assistert maksimal åpning, <sup>7</sup>ingen kjeveleddssmerter ved assistert maksimal åpning, PPV; Positiv prediktiv verdi, NPV; negativ prediktiv verdi, LR+; Positiv sannsynlighetsratio (likelihood ratio), LR-; negativ sannsynlighetsratio.

**Tabell 8.** Påvisning av degenerativ kjeveleddsforandring eller inflammatoriske kjeveleddstilstander ved bruk av ulike kliniske enkelttester og test-kluster som indekstest, sammenliknet med referansestandard MR eller artroskopi.

| Studie                                  | Pretest sannsynlighet | Antall ledd | PPV | NPV | LR+  | LR-  |
|---|-----------------------|-------------|-----|-----|------|------|
| <b>Klikk</b>                            |                       |             |     |     |      |      |
| <i>Reneker 2011</i>                     |                       |             |     |     |      |      |
| Manfredini, 2003 <sup>1a</sup>          | -                     | -           | -   | -   | 1,40 | 0,61 |
| <b>Krepitasjon</b>                      |                       |             |     |     |      |      |
| <i>Reneker 2011</i>                     |                       |             |     |     |      |      |
| Manfredini, 2003 <sup>1a</sup>          | -                     | -           | -   | -   | 1,21 | 0,50 |
| Holmlund & Axelsson, 1996 <sup>b</sup>  | -                     | -           | -   | -   | 3,21 | 0,64 |
| Holmlund & Axelsson, 1996 <sup>c</sup>  | -                     | -           | -   | -   | 4,19 | 0,39 |
| <b>Kjeveleddsbevegelse</b>              |                       |             |     |     |      |      |
| <i>Reneker 2011</i>                     |                       |             |     |     |      |      |
| Holmlund & Axelsson, 1996 <sup>1a</sup> | -                     | -           | -   | -   | 1,50 | 0,25 |
| <b>Smerte</b>                           |                       |             |     |     |      |      |
| <i>Reneker 2011</i>                     |                       |             |     |     |      |      |
| Manfredini, 2003 <sup>1a</sup>          | -                     | -           | -   | -   | 2,25 | 0,24 |
| Holmlund & Axelsson, 1996 <sup>1b</sup> | -                     | -           | -   | -   | 1,38 | 0,33 |
| Manfredini, 2003 <sup>2a</sup>          | -                     | -           | -   | -   | 2,08 | 0,30 |
| Manfredini, 2003 <sup>3a</sup>          | -                     | -           | -   | -   | 0,95 | 4,13 |
| Manfredini, 2003 <sup>4a</sup>          | -                     | -           | -   | -   | 1,32 | 0,60 |

<sup>1</sup>Smerte ved posterior/lateral palpasjon, <sup>2</sup>Kjeveleddssmerte ved åpning, <sup>3</sup>Kjeveleddssmerte ved assistert maksimal åpning, <sup>4</sup>Kjeveleddssmerte ved bevegelsesmotstand, <sup>a</sup>TMD-klassifisering; Effusjon, <sup>b</sup>TMD-klassifisering; Synovitt

PPV; Positiv prediktiv verdi, NPV; negativ prediktiv verdi, LR+; Positiv sannsynlighetsratio (likelihood ratio), LR-; negativ sannsynlighetsratio.



**Tabell 9.** Diagnostisk nøyaktighet av MR sammenliknet med frysensnitt foretatt post mortem for påvisning av diskusposisjon

| Studie               | n  | Pretest sannsynlighet | LR+                | LR-               | PPV                 | NPV                 |
|----------------------|----|-----------------------|--------------------|-------------------|---------------------|---------------------|
| Westesson<br>1987    | 15 | 0,47                  | 2,3<br>(0,9-5,9)   | 0,2<br>(0,0-1,6)  | 0,67<br>(0,31-0,91) | 0,83<br>(0,36-0,99) |
| Katzberg<br>1988     | 18 | 0,44                  | 4,4<br>(1,2-15,5)  | 0,2<br>(0,0-1,0)  | 0,78<br>(0,40-0,96) | 0,89<br>(0,51-0,99) |
| Schwaighofer<br>1990 | 22 | 0,36                  | 10,5<br>(1,5-72,4) | 0,27<br>(0,1-0,9) | 0,86<br>(0,42-0,99) | 0,87<br>(0,58-0,98) |
| Tasaki<br>1993       | 55 | 0,56                  | ∞                  | 0,10<br>(0,0-0,3) | 1,0<br>(0,85-1,0)   | 0,89<br>(0,70-0,97) |

∞ (Uendelig, nevneren = null ettersom det ikke ble identifisert falske negative)

95 % konfidensintervall er beregnet på bakgrunn av tall som er rapportert i oversiktsartikkelen.

PPV; Positiv prediktiv verdi, NPV; negativ prediktiv verdi, LR+; Positiv sannsynlighetsratio (likelihood ratio), LR-; negativ sannsynlighetsratio.

**Tabell 10.** Diagnostisk nøyaktighet av MR sammenliknet med frysensnitt foretatt post mortem for påvisning av diskusdeformitet.

| Studie            | n  | Pretest sannsynlighet | LR+              | LR-               | PPV                 | NPV                 |
|-------------------|----|-----------------------|------------------|-------------------|---------------------|---------------------|
| Westesson<br>1987 | 14 | 0,43                  | 1,7<br>(0,8-3,6) | 0,33<br>(0,0-2,5) | 0,56<br>(0,23-0,85) | 0,80<br>(0,30-0,99) |
| Tasaki<br>1993    | 55 | 0,27                  | ∞                | 0,20<br>(0,1-0,6) | 1,00<br>(0,70-1,00) | 0,93<br>(0,80-0,98) |

∞ (Uendelig, nevneren = null ettersom det ikke ble identifisert falske negative)

95 % konfidensintervall er beregnet på bakgrunn av tall som er rapportert i oversiktsartikkelen.

PPV; Positiv prediktiv verdi, NPV; negativ prediktiv verdi, LR+; Positiv sannsynlighetsratio (likelihood ratio), LR-; negativ sannsynlighetsratio.

**Tabell 11.** Diagnostisk nøyaktighet av MR ved benyttelse av T1-vektede sagittale bilder, sammenliknet med frysesnitt (post mortem) påvisning av ossøse kondylendringer

| Studie            | n  | Pretest sannsynlighet | LR+               | LR-               | PPV                 | NPV                 |
|-------------------|----|-----------------------|-------------------|-------------------|---------------------|---------------------|
| Westesson<br>1987 | 15 | 0,53                  | 1,75<br>(0,4-6,8) | 0,70<br>(0,3-1,6) | 0,67<br>(0,24-0,94) | 0,56<br>(0,23-0,85) |
| Katzberg<br>1988  | 18 | 0,33                  | ∞                 | 0,17<br>(0,0-1,0) | 1,00<br>(0,46-1,00) | 0,92<br>(0,62-1,00) |
| Tasaki<br>1993    | 55 | 0,56                  | ∞                 | 0,13<br>(0,1-0,3) | 1,00<br>(0,84-1,00) | 0,86<br>(0,66-0,95) |

∞ (Uendelig, nevneren = null ettersom det ikke ble identifisert falske negative)

95 % konfidensintervall er beregnet på bakgrunn av tall som er rapportert i oversiktsartikkelen.

PPV; Positiv prediktiv verdi, NPV; negativ prediktiv verdi, LR+; Positiv sannsynlighetsratio (likelihood ratio), LR-; negativ sannsynlighetsratio.

**Tabell 12.** Diagnostisk nøyaktighet av ultralyd sammenliknet med MR for deteksjon av intern kjeveleddsforandring med munnen i lukket posisjon.

| Studie              | n   | Pretest sannsynlighet | LR+<br>(95 % KI)       | LR-<br>(95 % KI)    | PPV | NPV |
|---------------------|-----|-----------------------|------------------------|---------------------|-----|-----|
| Emshoff<br>1997     | 17  | -                     | 1,75<br>(0,51-6,00)    | 0,70<br>(0,38-1,28) | -   | -   |
| Jank<br>2001        | 66  | -                     | 3,52<br>(2,01-6,14)    | 0,28<br>(0,18-0,43) | -   | -   |
| Emshoff<br>2002     | 29  | -                     | 16,20<br>(2,40-109,13) | 0,11<br>(0,04-0,27) | -   | -   |
| Emshoff<br>2002c    | 208 | -                     | 5,97<br>(3,77-9,47)    | 0,23<br>(0,18-0,31) | -   | -   |
| Brandlmaier<br>2003 | 48  | -                     | 4,03<br>(2,26-7,18)    | 0,09<br>(0,03-0,24) | -   | -   |
| Jank<br>2005        | 100 | -                     | 11,41<br>(4,92-26,48)  | 0,09<br>(0,05-0,15) | -   | -   |
| Landes<br>2006      | 68  | -                     | 1,63<br>(1,09-2,42)    | 0,62<br>(0,41-0,93) | -   | -   |
| Landes<br>2007      | 33  | -                     | 1,20<br>(0,79-1,81)    | 0,78<br>(0,43-1,39) | -   | -   |

PPV; Positiv prediktiv verdi, NPV; negativ prediktiv verdi, LR+; Positiv sannsynlighetsratio (likelihood ratio), LR-; negativ sannsynlighetsratio.

## Vedlegg 6 - Begrep og forkortelser

| Begrep/forkortelse                      | Forklaring  |
|---|---|
| <b>Artikulær</b>                        | Det som er inne i ledd, eventuelt brakt dit ved innsprøyting.   |
| <b>Artritt</b>                          | Betennelse i ledd.  |
| <b>Artroskop</b>                        | Kikkertinstrument som brukes til å se inn i leddhuler.  |
| <b>Artroskopi</b>                       | Lukket kirurgisk metode. Utføres under generell anestesi, hvor et artroskop føres inn gjennom et snitt i huden. En avløpskanyle plasseres gjennom huden ved siden av og like under innstikkstedet for artroskopet. Det øvre kjeveleddskammeret gjennomskylles og kan deretter optisk inspiseres.  |
| <b>Artrose</b>                          | Degenerativ leddforandring.   |
| <b>Bruksisme</b>                        | Betegnelse på overdreven gnisning, skjæring eller pressing av tenner.   |
| <b>Closed lock</b>                      | Når diskusforskyvningen ikke går tilbake til normal posisjon i løpet av gapebevegelsen kalles tilstanden gjerne "closed lock".  |
| <b>Coronalplan</b>                      | = Frontalplan.  |
| <b>CT</b>                               | Computertomografi. Røntgenologisk undersøkelsesmetode som kan fremstille tverrsnittsbilder av enhver kroppsdel.   |
| <b>Defleksjon</b>                       | Kjeven avviker til en side fra midtstilling, uten tilbakegang til midtstilling ved åpning av munnen.  |
| <b>Degenerativ kjeveleddsforandring</b> | Nedbryting av leddbrusk.  |
| <b>Deviasjon</b>                        | Kjeven avviker til en side fra midtstilling, med tilbakegang til midtstilling ved åpning av munnen.   |
| <b>Diskus</b>                           | Bruskskiver i ledd.   |
| <b>Effektestimater</b>                  | Mål for effekt, f.eks. gjennomsnitt, frekvens, prosent, relativ risiko, odds ratio, "numbers needed to treat to benefit," standardisert gjennomsnittlig forskjell eller vektet gjennomsnittlig forskjell.   |
| <b>Elektromyografi</b>                  | EMG - metode til registrering av den elektriske aktivitet i musklene.   |
| <b>Ekstraartikulær</b>                  | Noe som ligger utenfor leddhulen eller de strukturene som bygger opp denne.   |
| <b>EMG</b>                              | Elektromyografi   |
| <b>Frontalplan</b>                      | Vertikal linje som går parallelt med pannens plan og som ligger på kroppens fremside (kroppens «frontalside»).  |
| <b>Frysesnitt</b>                       | En patologisk laboratorieprosedyre (kryosnitt)  |
| <b>Heterogen</b>                        | Ulik, uensartet. Populasjoner eller studier er heterogene når gir et uensartet eller ulikt uttrykk, noe som betyr at de er forskjellige fra hverandre.  |
| <b>Intern kjeveleddsforandring</b>      | Forskyvning av kjeveleddsdiskus fra dens normale posisjon. Fremoverforskyvning av diskusen med tilbakegang (el. reduksjon) er definert som diskusforskyvning ved lukket munn som går tilbake (med el. uten klikk) til normal posisjon i løpet av en gapebevegelse. Når diskusforskyvningen ikke går tilbake til normal posisjon i løpet av gapebevegelsen kalles tilstanden gjerne "closed lock". |
| <b>Intraartikulær</b>                   | I eller innenfor leddet.  |
| <b>Intrakapsulær</b>                    | Det som er innenfor leddkapselen.   |
| <b>Kondyl</b>                           | Knoke, leddknoke, benfremspring i enden av en knokkel, utgjør en del av et ledd.  |
| <b>Mandibula</b>                        | Underkjeve  |
| <b>Maxilla</b>                          | Overkjeve   |
| <b>Meta-analyse</b>                     | Statistiske teknikk for å integrere resultatene av inkluderte studier.  |
| <b>Morfologi</b>                        | Læren om struktur, oppbygning og form på vev, organer og individer. Brukes også som fellesbetegnelse på normal og patologisk ana-   |

|                                 |   |
|---------------------------------|---|
|                                 | tomi.   |
| <b>MR</b>                       | Magnetisk resonanstomografi. Metode til fremstilling av snittbilder av kroppen.   |
| <b>Okklusjon</b>                | Betegnelse på sambitt av tennene i over- og underkjeven .   |
| <b>Orofacial</b>                | I eller rundt ansiktet.   |
| <b>Os mandibulare</b>           | Underkjevebenet.  |
| <b>Osseøs</b>                   | Benet; som har med benvev å gjøre.  |
| <b>Os temporale</b>             | Tinningbeinet.  |
| <b>Osteotomi</b>                | Operativ gjennomsaging eller gjennommeisling av en knokkel, evt. fjernelse av en del av en knokkel.   |
| <b>Prediktiv verdi</b>          | Begrep brukt ved beskrivelsen av diagnostiske tester eller screeningmetoder, om sannsynligheten for at en person med utslag på testen (positiv test) virkelig er syk eller virkelig har den undersøkte risikofaktoren (positiv prediktiv verdi), eller om sannsynligheten for at en person uten utslag på testen virkelig er frisk (negativ prediktiv verdi). |
| <b>Reliabilitet</b>             | Pålitelighet. I hvilken grad man får samme resultater når en måling eller undersøkelse gjentas under identiske forhold.   |
| <b>Sagittalplan</b>             | Et tenkt plan gjennom kroppens midtlinje som deler kroppen i en høyre og en venstre halvdel.  |
| <b>Sannsynlighetsratio</b>      | Sannsynligheten for å få et visst prøvesvar hos en pasient som har den sykdommen vi mistenker, dividert med sannsynligheten for å akkurat det samme prøvesvaret hos en person som ikke har sykdommen.   |
| <b>Sensitivitet</b>             | Brukes om sannsynligheten for at man ved hjelp av en test eller undersøkelsesmetode for en viss sykdom kan påvise sykdommen hos dem som virkelig har sykdommen. Jo større sensitivitet, jo mindre er sannsynligheten for falskt negative prøveresultater.   |
| <b>Spesifisitet</b>             | Et uttrykk for sannsynligheten for at man ved hjelp av en test eller undersøkelsesmetode for en viss sykdom kan klassifisere som friske de som virkelig ikke har sykdommen. Jo større spesifisitet, jo mindre er sannsynligheten for falskt positive prøveresultater (prøveresultat som feilaktig sier at personen er syk).                                   |
| <b>Synovitt</b>                 | Betennelse i synovialhinnen, den innerste hinnen i kapselen rundt et ledd.  |
| <b>Temporomandibulær-leddet</b> | Kjeveleddet   |
| <b>TMD</b>                      | Temporomandibular Disorder  |
| <b>Translasjonsbevegelse</b>    | Fremoverglidning  |
| <b>Trismus</b>                  | Krampaktig sammentrekning av tyggemusklene, slik at det blir umulig å åpne munnen normalt.  |
| <b>T-vektet</b>                 | Begrep som benyttes ved MR-undersøkelse. MR-sekvenser ved undersøkelser av skjelett og bløtdeler vil i utgangspunktet være T1-vektede og T2-vektede. T1-vektet sekvens gir først og fremst en anatomisk fremstilling av strukturene, men kan også vise patologiske forhold. T2-vektet sekvens er mer sensitiv for patologiske forandringer.                   |
| <b>Validitet</b>                | Gyldighet. I hvilken grad man ut fra resultatene av et forsøk eller en studie kan trekke gyldige slutninger om det man har satt seg som formål å undersøke.   |

Nasjonalt kunnskapssenter for helsetjenesten  
Postboks 7004, St. Olavs plass  
N-0130 Oslo  
(+47) 23 25 50 00  
www.kunnskapssenteret.no  
Rapport: ISBN 978-82-8121-535-1 ISSN 1890-1298

nr 10-2013



kunnskapssenteret

utelukke myofascial TMD sammenliknet med kliniske tester. •Resultatene må tolkes med forsiktighet, fordi primærstudiene i de inkluderte systematiske oversiktene kan ha metodiske svakheter, det er stor variasjon i resultatene, og feilmarginene i nøyaktighetsmålene er gjennomgående brede eller ikke oppgitt.