



# PROTOKOL

INDLEDENDE EVALUERING AF GENERATIV AI TIL OPSUMMERING AF VIDEOKONSULTATION I ALMEN PRAKSIS

## FORMÅL

At udføre indledende evaluering af accept, behov og effektivitet af en generativ AI-løsning til opsummering af videokonsultation i form af et journalnotat i almen praksis.

## METODE

Metoden for den indledende evaluering af AI-løsningen består af fire dele; observation, tidsmåling, tænke-højt og interview. Forløbet af metoderne kan ses i figur 1.

**Observation:** Evalueringen starter med, at lægen afholder en normal videokonsultation med en samtykkende patient. Denne videokonsultation observeres af en forsker fra Center for Almen medicin (CAM AAU), hvor der føres notat omkring indhold i videokonsultation til senere sammenligning og analyse af det AI-genererede journalnotat.

Efter afslutning af videokonsultationen kan der ske en af to ting – enten måling af tiden det tager lægen at skrive sit eget journalnotat eller udlevering af det AI-genererede journalnotat, som lægen skal modificere indholdet ved tænke-højt metoden.

- **Måling af tid:** Lægen skriver et journalnotat for den netop overståede videokonsultation. Tiden det tager for lægen at skrive journalnotatet måles og noteres. Journalnotatet printes og afleveres til forskeren fra CAM AAU.
- **Tænke-højt:** Lægen får udleveret det AI-genererede journalnotat. Hvis lægen har haft til opgave at skrive sit eget journalnotat, så sammenlignes det AI-genererede journalnotat med lægens eget journalnotat og lægen modificerer indholdet, så det stemmer overens med, hvad lægen mener ville være et korrekt journalnotat. Hvis lægen har fået udleveret det AI-genererede journalnotat umiddelbart efter videokonsultationen, så skal lægen modificere indholdet til det menes at afspejle videokonsultationen. Tiden det tager for lægen at modificere det AI-genererede journalnotat måles og noteres. Metoden 'Tænke-højt' indbefatter, at lægen taler sine umiddelbare tanker højt mens lægen modificerer det AI-genererede journalnotat.

Omkring halvdelen af deltagerne skal igennem både måling af tid og tænke-højt. Den anden halvdel af lægerne skal kun igennem tænke-højt, hvor det AI-genererede journalnotat modificeres. Grunden hertil er at måle den sparede tid ved at anvende det AI-genererede journalnotat, som modificeres til ønsket indhold sammenlignet med, at lægen selv skriver journalnotatet.

**Interview:** Efter gennemgang af tænke-højt metoden, vil lægen indgå i et semistruktureret interview med forskeren fra CAM AAU. Interviewet vil starte med at indsamle umiddelbare reaktioner på det AI-genererede journalnotat. Derefter vil der spørges ind til lægens tanker om accept af det AI-genererede journalnotat og behovet for den pågældende løsning i organisationen.



FIGUR 1: ILLUSTRATION AF FORLØBET FOR EVALUERINGEN AF AI-LØSNINGEN.

## DATAINDSAMLING

### DELTAGERE

Til udførelse af evalueringen er der behov for planlægning af omkring 20 videokonsultationer. Der er ønske om, at der rekrutteres 10-12 klinikker, hvor en til to læger pr. klinik deltager i evalueringen med efterfølgende interview til måling af accept og behov. Dog skal der gerne indsamles data fra så mange videokonsultationer som muligt, udover de 20 videokonsultationer, som skal bruges til måling af effektivitet. Dette kræver en aftale om deling af data fra videokonsultationer udover de

Det planlægges, at evalueringen vil tage omkring to timer pr. læge. Den beregnede tid omfatter alt fra videokonsultation til afslutning af det semistrukturerede interview.

Derudover er der også behov for patienter med planlagte videokonsultationer. Patienterne skal give samtykke til, at deres videokonsultation og journalnotat må indgå i evalueringen.

#### SETTING

Det planlægges, at evalueringen kan foregå i de deltagende lægers egen klinik, hvor forskeren fra CAM AAU vil møde op på et aftalt tidspunkt.

#### UDSTYR

Til evaluering er der behov for følgende udstyr:

- AI-løsningen udviklet af Corti og Trifork skal være tilgængelig for de involverede praktiserende læger og fungere på en sikker måde.
- stopur til måling af tiden det tager lægen at skrive journalnotatet selv, samt tiden det tager at modificere det AI-genererede journalnotat.
- diktafon til optagelse af tænke-højt metoden og det semistrukturerede interview.
- semistruktureret interviewguide til vejledning af interview.

#### EFFEKT MÅL

Til måling af effekt af AI-løsningen til generering af journalnotat, vil der blive målt på accept, behov og effektivitet.

#### ACCEPT

Til måling af accept blandt de praktiserende læger, skal det kvalitativt vurderes om løsningen lever op til de krav, der eksisterer for et godt journalnotat. Accept vil blive målt på følgende parametre:

- Oplevet brugbarhed. Gør løsningen det den skal og oplever den praktiserende læge, at løsningen har den nødvendige kvalitet?
- Oplevet brugervenlighed. Oplever de praktiserende læger, at løsningen er nem at anvende? Finder de praktiserende læger løsningen intuitiv?
- Intention om brug. Kan de praktiserende læger forestille sig at fortsætte med at bruge løsningen? Hvad er sandsynligheden for, at løsningen vil blive brugt, hvis implementeret?

#### BEHOV

Til måling af behov, skal de praktiserende læger kvalitativt vurdere behovet for løsningen, og hvordan løsningen vil passe ind i organisationen. Behovet for løsningen vil blive målt på følgende parametre:

- Efterspørgsel. Oplever de praktiserende læger, at der er behov for løsningen? Hvilke problemer i klinikken ville løsningen kunne bidrage til at løse?
- Hensigtsmæssighed. Oplever de praktiserende læger, at løsningen er passende ift. Det formål, som det skal leve op til? Har de idéer til, hvordan løsningen ellers kan bruges? Hvilke potentialer og udfordringer forestiller de praktiserende læger sig, at der kan følge med løsningen?
- Organisationskultur. Passer løsningen ind i den organisation, som eksisterer i almen praksis? Ville løsningen også kunne bruges af andet personale i klinikken?

#### EFFEKTIVITET

Til evaluering af effektivitet af AI-løsningen, skal de følgende data gøres tilgængelige for CAM AAU:

- Lydfil af videokonsultation

- Automatisk transskription af videokonsultation med tidsstempel.
- Det AI-genererede journalnotat med tidsstempel, når det vises til den praktiserende læge.
- Lægens eget journalnotat med tidsstempel, når det gemmes.

## ETIK

Etisk komité skal kontaktes for at få afklaret behovet for etisk godkendelse af evalueringen. Det formodes dog, at evalueringen ikke er anmeldelsespligtig, da evalueringen er baseret på observation og interview, og der ikke indgår menneskeligt biologisk materiale. Projektet vil også blive anmeldt til den forskningsetiske komité ved Aalborg Universitet.

Evalueringen vil blive anmeldt som forskningsprojekt til kontraktenheden ved Aalborg Universitet i overensstemmelse med regler om GDPR. Derudover vil der blive indsamlet informeret samtykke fra deltagende læge om tilladelse til at bruge data og udsagn i evalueringen. Derudover, skal der også indsamles informeret samtykke fra patienten om tilladelse til at forskeren fra CAM AAU må observere videokonsultationen samt tilladelse til deling af lydfile med videokonsultation, journalnotatet skrevet af lægen og det AI-genererede journalnotat.

## OPBEVARING AF DATA

Efter dataindsamling vil alle lydfile blive uploadet til en sikker fildelingsserver leveret af Aalborg Universitet i overensstemmelse med GDPR. Alle interviewdata vil blive pseudonymiseret ved upload til serveren med dannelse af en nøgle til sikker referencehåndtering.

Kvalitativt data indsamlet i forbindelse med tænke-højt metoden og de semistrukturerede interviews vil blive transskriberet med det formål at kunne blive analyseret. Transskriptionen vil blive udført af en AI-løsning udbudt af Aalborg Universitet kaldet *"Transscriber"*.

## ANALYSE

### ACCEPT OG BEHOV

Til analyse af accept og behov vil der blive gennemført en tematisk analyse. Analysen vil have fokus på at identificere de vigtigste elementer med betydning for accept af AI-løsningen, og hvilke behov der eksisterer for at løsningen vil kunne eksistere videre i almen praksis.

En tematisk analyse foregår på den måde, at der først skal opnås dybdegående kendskab til det indsamlede kvalitative data. Herefter skal data kodes med henblik på at finde udsagn som understøtter formålet med evalueringen. Når kodning af data er fuldført, skal det forsøges at findes mønstre i koderne, som går på tværs af data fra alle deltagerne. Koderne bliver dermed efterhånden syntetiseret og samlet i tværgående temaer, som kan fremsætte de praktiserende lægers væsentligste pointer i relation til accept og behov ved AI-løsningen til generering af journalnotat i videokonsultationen.

### EFFEKTIVITET

For at evaluere effektivitet af AI-løsningen for hver videokonsultation, skal der første gennemføres en vurdering af de potentielle ændringer i den AI-genereret journalnotat:

- Type af ændring:
  - Formatering og grammatik
  - Upræcis information
  - Forkert information
  - Manglede information

- Hallucination
- Klinisk signifikans (ingen, lav, høj)

Yderligere skal tiden det tager at gennemgå og gemme det AI-genererede journalnotat måles.

Baseret på vurderingen beskrevet ovenfor, skal der gennemføres statistisk analyse for at få en estimering af fejltypen og tidsbesparelse, med statistisk signifikans og konfidensinterval.

## TIDSPLAN

I nedenstående tabel fremgår tidsplanen for evalueringen.

	Sep	Okt	Nov	Dec	Jan	Feb	Mar	Apr
Forberedelse								
Dataindsamling								
Analyse								
Afrapportering								