Slutrapport

**Inkubator för nationella plattformar för systematisk klinisk validering av AI inom bilddiagnostik mammografi VAI-B**

**Dokumenthistorik**

|  |  |  |  |
| --- | --- | --- | --- |
| **Version** | **Datum** | **Kommentar** | **Handläggare** |
| **0.1** | **9 dec 2024** | **Fösta ifyllnad för VAI** | **Ann-Sofi Mikaelsson** |
| **0.2** | **8 jan 2025** | **Fokus på VAI-B** | **Ann-Sofi Mikaelsson** |
| **0.3** | **20 jan 2025** | **korrigeringar** | **Ann-Sofi Mikaelsson** |
| **0.4** | **2 feb 2025** | **Fredriks korrigeringar** | **Ann-Sofi Mikaelsson** |
| **1.0** | **3 feb 2023** | **Publicerad** | **Ann-Sofi Mikaelsson** |

Innehållsförteckning

[1 Uppfattat resultat i förhållande till förväntningar på projektet\* 3](#_Toc98682739)

[2 Sammanfattning\* 5](#_Toc237446340)

[3 Bakgrund\* 5](#_Toc1175770905)

[3.1 Syfte 5](#_Toc250500373)

[3.2 Historik 6](#_Toc65413626)

[3.3 Projektorganisation 8](#_Toc649080527)

[3.4 Använda metoder 9](#_Toc160898856)

[4 Projektets resultat kontra projektmål\* 9](#_Toc1899716777)

[4.1 Projektresultat kontra projektmål 9](#_Toc1481821937)

[4.2 Tidplan kontra utfall 9](#_Toc124130870)

[5 Erfarenheter av projektarbetet\* 9](#_Toc2059458180)

[5.1 Förslag till förbättringar och rekommendationer 9](#_Toc1453934621)

[6 Övrigt 10](#_Toc1151145392)

[Referenser 10](#_Toc1362687094)

# Uppfattat resultat i förhållande till förväntningar på projektet\*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Måluppfyllnad av projektresultatet?** | | | | |
| **Dåligt** | **Någorlunda** | **Bra** | **Mycket bra** | **Utmärkt** |
| **Höll tidplanen?** | | | | |
| **Dåligt** | **Någorlunda** | **Bra** | **Mycket bra** | **Utmärkt** |
| **Höll ekonomin?** | | | | |
| **Dåligt** | **Någorlunda** | **Bra** | **Mycket bra** | **Utmärkt** |
| **Höll kvaliteten?** | | | | |
| **Dåligt** | **Någorlunda** | **Bra** | **Mycket bra** | **Utmärkt** |
| **Rätt resurser i tillräcklig mängd?** | | | | |
| **Dåligt** | **Någorlunda** | **Bra** | **Mycket bra** | **Utmärkt** |
| **Projektorganisation** | | | | |
| **Dåligt** | **Någorlunda** | **Bra** | **Mycket bra** | **Utmärkt** |
| **Projektstyrning** | | | | |
| **Dåligt** | **Någorlunda** | **Bra** | **Mycket bra** | **Utmärkt** |
| **Kommunikation inom och externt projektet** | | | | |
| **Dåligt** | **Någorlunda** | **Bra** | **Mycket bra** | **Utmärkt** |
| **Din bedömning av framgång i projektet** | | | | |
| **Dåligt** | **Någorlunda** | **Bra** | **Mycket bra** | **Utmärkt** |

# Sammanfattning\*

Projektet har haft tre delprojekt, Valideringsplattorm AI, VAI-B(bröst), VAI-P(patologi), VAI-S (stroke) med egna projektorganisationer men med kunskapsutbyte emellan och gemensamma styrgruppsmöten. Denna slutrapport ägnas till största delen åt VAI-B då övriga har separata rapporter. VAI-B startade först och hann längst både kliniskt och akademiskt. Det är material från 5 regioner, 4 AI leverantörer som har analyserats och rapporter har delgivits de tre huvudregionerna och leverantörerna. En av konklusionerna är att det inte finns en AI lösning som passar alla. Det är viktigt för regionerna att identifiera vilka modaliteter man har och arbetsflöden man vill använda för att se vilken AI algoritm som kommer prestera bäst.

# Bakgrund\*

## Syfte

Den övergripande miljöns mål är att utveckla svensk hälso- och sjukvårds förmåga, över hela landet, att genomföra systematisk klinisk validering av kommersiellt tillgänglig AI-programvara som är tänkt att fungera som beslutsstöd inom bilddiagnostik - såväl inför upphandlingar som vid senare behov att validera förändringar från tillverkarna. Det vi menar med validering är att testa AI-algoritmer på en stor mängd historiska, väl validerade data för att avgöra hur hög träffsäkerheten är - jämfört med andra tillgängliga AI-algoritmer och jämfört med mänskliga bilddiagnostiker. Det är viktigt att de historiska data är representativa, d.v.s. härrör från olika geografiska områden i Sverige, med olika typer av utrustningar, och med olika sammansättning av patientunderlaget avseende ålder, kön och etnicitet.

Innovationsmiljön utgör en inkubator för valideringsplattformar, se figuren nedan. Under den ansökta tidsramen kommer nationella valideringsplattformar att skapas för tre områden där mognaden är hög, och valideringar genomföras som undanröjer hinder för implementering i dessa tre områden. Målsättningen är att de tre valideringsplattformarna efter projektet ska fungera självständigt och uthålligt inom ordinarie vårdverksamhet, samt att svensk sjukvård effektivt ska kunna skapa valideringsplattformar för ytterligare områden efter behov, både inom bilddiagnostik men i hög grad även för andra områden som nyttjar medicinsk teknik.

En bild som visar text, skärmbild, Teckensnitt

Automatiskt genererad beskrivning

Strategin för att uppnå dessa mål är som följer:

* Hälso- och sjukvården ska vara drivande redan under uppbyggnad av inkubator och plattformar, för att möjliggöra övergången till ordinarie verksamhet bortom inkubatortiden.
* I så hög grad som möjligt ska befintliga strukturer nyttjas istället för att skapa nya, vilket ger snabbare startsträcka och lättare väg till translation ut ur projektet.
* Vårdens befintliga kunskapsstyrning i form av Nationella Programområdet (NPO) Medicinsk Diagnostik, och Regionala Cancercenter i samverkan (RCC) utgör ledning av plattformarna.
* Analytic Imaging Diagnostics Arena (AIDA) utgör miljö för kunskapsutbyte och samverkan mellan plattformar och dubbelriktad länk till andra relevanta insatser inom bilddiagnostik och medicinteknik.
* Existerande juridiska strukturer används i största möjliga utsträckning, där huvudhypotesen är att driva plattformarna som kvalitetsregister.
* Existerande teknisk infrastruktur för överföring, lagring och processning av medicinska bilder används i största möjliga utsträckning, där system från AIDA/SciLifeLab, Sectra och Collective Minds Radiology är vidtalade möjliga lösningar.
* Att driva generell utveckling genom spetsiga pilotinsatser inom de mest mogna områdena.
* En hörnsten i miljön är starka forskargrupper med AI-validering i mammografiscreening som specialitet (Karolinska Institutet och Lunds universitet).
* Mammografiscreening är det område inom bilddiagnostik som har de mest mogna och etablerade AI-lösningarna.
* Initialt ska en till två vårdgivare identifieras som pilotklienter inom varje område. Dessa ska väljas bland vårdgivare som står utanför de redan etablerade forskargrupperna, och förväntas bidra med bilder och kliniska data.
* Att använda ett agilt och flexibelt arbetssätt i varje område anpassat till praktiska förutsättningar för genomförbarhet.
* Prioritera beständiga och generella framsteg.
* Enskilda valideringar är inte i fokus i sig, utan är medel för att bygga kunskap och förmåga i form av best practices för organisation, arbetssätt, legal struktur, teknikplattform mm.

## Historik

Nuläget för bilddiagnostisk AI har en tudelad karaktär och de olika sidorna har ytterligare accentuerats under de senaste åren. På den positiva sidan finns att trenden vad gäller potentialen för AI att göra stor nytta i vården är fortsatt mycket stark. Varje vecka publiceras nya forskningsresultat som tydligt visar på möjligheter till säkrare och effektivare diagnostik. För vissa isolerade diagnostiska uppgifter, t ex mammografiscreening, har retrospektiva forskningsstudier gång efter annan etablerat att AI-modeller kan vara lika bra eller bättre än mänskliga bedömningar.

På den negativa sidan är att det fortsatt är en väldigt hög tröskel för AI-innovationer att komma ut och göra nytta i vården. Detta trots de enorma investeringar som görs hos företag med detta som mål, och trots att det numera finns många regulatoriskt godkända produkter på marknaden. Aktuella vetenskapliga översikter från 2021 visar dock att evidensen för att nytta kan realiseras ute i vården i hög grad saknas. Prospektiva studier är extremt få, och studier i verklig klinisk miljö likaså.

Slutsatsen är att det gap som finns mellan möjligheter med AI och faktisk användning i vården fortfarande är mycket stort. De faktorer som brukar lyftas fram, och som är i linje med erfarenheterna i AIDA från den svenska situationen, är sammanfattningsvis att tekniksidan måste få bättre förståelse för sjukvårdens behov och förutsättningar och att sjukvården behöver få bättre förståelse för teknikens möjligheter och begränsningar. Medan fokus ofta ligger på algoritmernas prestanda i forskningsstudier, är det ett brett spektrum aspekter som utgör de svårbemästrade hindren för nyttiggörande. Klinisk validering är en helt central del i att undanröja dessa hinder. Det handlar om förståelse för nuvarande arbetsflöden och förmåga att skapa nya innovativa arbetssätt med AI-verktyg, om förmåga till effektivt och säkert införande med processer för utvärdering och kvalitetssäkring, insikt om hur organisationer behöver förstärkas med nya roller och kompetens, och mycket mer.

Under ledning av Fredrik Strand, medicinsk ledare VAI-B, har man på Karolinska Institutet och Universitetssjukhuset nyligen genomfört en validering av tre AI-algoritmer för mammografi-screening, vilken även framgångsrikt har publicerats i JAMA Oncology. Detta lyckades då teamet där samlat ihop en stor databas med samtliga mammografibilder för samtliga kvinnor i screeningålder 40 till 74 år som drabbats av bröstcancer år 2008 till 2015, samt ett urval mammografier från friska kvinnor. För dessa kvinnor länkades data in från Regionalt Cancercentrum om cancerdiagnoser samt om röntgenläkarnas dokumenterade bedömningar av bilderna. Den aktuella ansökan tar sitt avstamp i denna lyckade validering.

Under ledning av Sophia Zackrisson, vice medicinsk och vetenskaplig ledare för VAI-B har en databas med 15,000 screeningundersökningar (både 2D och 3D mammografi) skapats baserat på en studie med kvinnor 40–74 år, med annoteringar och validering gentemot cancerregister och andra vårdregister samt bedömningar med röntgenläkares granskningar. I detta material har tre valideringsstudier (varav 1 än så länge publicerad, Eur Radiol) gjorts med en AI-leverantör samt av en riskestimeringsmetod baserad på annan leverantör. Delar av materialet har gjorts tillgängligt via AIDAs plattform för internationella samarbeten avseende utveckling och validering av AI-algoritmer, där flera projekt är i diskussionsfas.  För närvarande pågår ett stort arbete med att utvidga databasen att innefatta samtliga 2D mammografier (kliniska och screening) samt 3D mammografier i Malmö från 2004–2018, med matchning mot diverse register, databaser och biobanker, med stöd av bland annat Cancerfonden. Databasen kommer slutligen att innehålla>450 000 undersökningar och beräknas vara klar under 2022/2023.

En bild som visar text, Teckensnitt, skärmbild, cirkel

Automatiskt genererad beskrivning

AIDA är en nationell arena för forskning och innovation kring artificiell intelligens, AI, för medicinsk bildanalys, som startade inom regeringens samverkansprogram 2017. Här möts akademi, sjukvård och industri för att omsätta tekniska framsteg inom AI-teknik till patientnytta i form av kliniskt användbara verktyg. AIDA är en del av Strategiska Innovationsprogrammet Medtech4Health, och Centrum för Medicinsk Bildvetenskap och Visualisering (CMIV) vid Linköpings Universitet står värd för arenan. AIDA har sedan starten bl.a. drivit ca 30 projekt för AI-utveckling, organiserat workshops för fler än 1000 personer, gett fem AI-kurser för bilddiagnostiker och skapat en datahub av bilddata för AI-träning (som under 2021 blir en del av SciLifeLab). Den föreslagna inkubatorn kan ses som en utökning av AIDAs initiativ för kliniska utvärderingar, och kommer ha tät koppling och stora synergier med den befintliga arenamiljön.

Bland de många frågor som Bröstcancerförbundet framgångsrikt driver ur patientperspektivet, finns ett speciellt fokus på AI-frågor. I Bröstcancerrapporten 2020, t.ex., belystes både drabbade kvinnors såväl som bröstradiologers inställning till AI-användning.

## Projektorganisation

**Projektgrupp:**

Claes Lundström har rollen som huvudprojektledare och ansvarar för rapportering till och kontakter med Vinnova. Claes är också Arenaledare för AIDA.

Programutvecklare är Håkan Gustafsson, AI-strateg vid Region Östergötland och biträdande föreståndare vid CMIV vid Linköpings Universitet, som också är värdorganisation för AIDA. Programutvecklingsdelen av projektet leds av AIDA Styrgrupp genom att programutvecklaren rapporterar dit, vilket motiveras av den nära kopplingen till AIDAs övriga verksamhet. AIDA Styrgrupp är således inte ett beslutande organ för delprojekten, vilket gör att ineffektivt dubbelkommando och oklara beslutsvägar undviks.

**VAI-B**

Respektive plattform har en ansvarig med medicinsk och vetenskaplig kompetens. För plattformen för mammografi är detta Fredrik Strand, Karolinska Institutet/Karolinska Universitetssjukhuset och vice forskningsledare är Sophia Zackrisson, Lunds universitet/Skånes Universitetssjukhus. En central roll är plattformarnas seniora projektledare, Ann-Sofi Mikaelsson, Frontit. VAI-Bs biostatistiker Haiko Schurz har en ledande roll i analysfasen av projektet.

Huvudprojektledare: Claes Lundström

Programutvecklare: Håkan Gustafsson

Projektledare: Ann-Sofi Mikaelsson

Forskningsledare: Fredrik Strand

Bitr forskningsledare: Sophia Zackrisson

Biostatistiker: Haiko Schurz

Databasutvecklare: Fernando Cossio Ramirez jan 2022 - okt 2023

IT-arkitekt: Carl Barck Holst dec 2021 - jan2023

Affärsutvecklare: Gerhard Bottma jan - nov 2023

Biostatistik: Klara Solander sept 2023 - aug 2024

Forskningsingenjör: Davida Åström nov 2023 - okt 2024

**Samarbetspartner:**

Bröstcancerförbundet

**Styrgrupp:**

Styrgruppen består av:

Carl Backman (ordf), NPO Medicinsk diagnostik, Västra Götalandsregionen

Jonas Cederberg, NPO Medicinsk diagnostik, Region Västmanland

Katrine Riklund, NPO Medicinsk diagnostik, Region Västerbotten

Johan Ahlgren, RCC i samverkan

Karin Stenström, Representant för Vinnova

## Använda metoder

Styrgruppsmöten total 13 st digitala teamsmöten

Projektverktyget Antura driftat av Region Östergötland.

Teamsyta Region Östergötland

Digitala teamsmöten varannan vecka samt fysiska heldagsmöten 4 ggr per termin

# Projektets resultat kontra projektmål\*

## Projektresultat kontra projektmål

Den förväntade effekten av den föreslagna satsningen är att svenska vårdgivare kommer att kunna genomföra djup förnyelse i form av verktyg och arbetssätt, med stor förbättring av vårdkvalitet och produktivitet, till nytta för sina patienter. AI-stöd kommer att gå från teoretisk vision till vardagsteknik, och Sverige kommer att vara världsledande i att omsätta möjligheterna till vårdnytta på ett säkert och kontrollerat sätt. Vi förväntar oss också att dessa unika plattformar även kan bli en attraktiv plats för internationell testverksamhet, liksom Arvidsjaur har etablerat sig som ett unikt centrum för vintertestning av bilar.

På kort sikt förväntas projektet ge en stark acceleration av införande av AI-verktyg inom de mest mogna områdena. På lång sikt förväntas denna acceleration ha spridit sig till såväl alla bilddiagnostiska områden som all annan medicinteknik.

## Tidplan kontra utfall

Tidplanen finns i Region Östergötlands projektverktyg Antura

Projektet startade november 2021, finansieringen var godkänd för 2.5 år (mars 2024) men då medel kvarstod tillsammans med värdefulla aktiviteter fick projektet tillstånd att nyttja medlen under hela 2024 och slutrapportera januari 2025.

208 aktiviteter som uppföljts i tidplanen

# Erfarenheter av projektarbetet\*

Projektet VAI-B har haft god progress mycket tack vare en bra sammansättning av gruppmedlemmar där alla har prioriterat arbetet. Projektet har både levererat nytta praktiskt i form av rapporter till ingående regioner och AI leverantörer samt publicerat och presenterat forskning på den insamlade data. Regelbundenheten i de digitala mötena kombinerat med heldags fysiska möten har varit en framgångsfaktor. Aktiviteten blir högre i ett fysiskt möte, särskilt när det som diskuteras är mer komplicerat; däremot avrapporteringar och status fungerar utmärkt digitalt.

Vi har haft viss omsättning i gruppen då vissa kompetenser inte har behövts under hela projektet men en kärna om fem personer har varit med hela vägen vilket har varit värdefullt.

## Förslag till förbättringar och rekommendationer

När det kommer till att förändra, förskjuta gränser i existerande juridiska tolkningar, kan man inte nog betona hur mycket kalendertid som behövs. Projektet kom efter många diskussioner med olika parter fram till att den mest rimliga vägen framåt för databasen och erfarenheterna var i form av ett kvalitetsregister. Detta för att kunna ge alla regioner chansen att validera och för att regelverket/lagstiftningen bedömdes passa bäst.

Diskussioner startade med Karolinska sjukhuset eftersom Kvalitetsregister behöver ha en vårdgivare som huvudman. Diskussionerna har varit positiva men stötestenen är hur data ska kunna överföras från Karolinska Institutet, och här är vi ännu inte i mål. Det vore mycket olyckligt ur ett ekonomiskt och tidsperspektiv om data skulle behöva samlas in på nytt av KS pga att KI inte på ett legalt sätt kan föra över informationen.

# Övrigt

-

# Referenser

*Antura, Region Östergötlands projektverktyg, https://projekt.regionostergotland.se*