

Statusartikel

Kunstig intelligens og føtalmedicinsk ultralydskanning

Caroline Taksøe-Vester^{1, 2, 3}, Zahra Bashir^{1, 3, 4}, Emilie Pi Fogtmann Sejer^{1, 2, 3}, Mary Le Ngo^{1, 3, 4}, Lars Henning Pedersen^{5, 6}, Mikkel Friis⁷, Olav B. Petersen^{1, 2}, Aasa Feragen⁸, Mads Nielsen⁹, Anders Nymark Christensen⁸ & Martin G. Tolsgaard^{1, 2, 3}

1) Institut for Klinisk Medicin, Københavns Universitet, 2) Center for Føtalmedicin, Afdeling for Gynækologi, Fertilitet og Fødsler, Københavns Universitetshospital – Rigshospitalet, 3) Copenhagen Academy of Medical Education and Simulation, Københavns Universitetshospital – Rigshospitalet, 4) Center for Føtalmedicin, Afdeling for Obstetrik, Sjællands Universitetshospital – Slagelse Hospital, 5) Institut for Klinisk Medicin, Aarhus Universitet, 6) Kvindesygdomme og Fødsler, Aarhus Universitetshospital, 7) NordSim, Aalborg Universitet, 8) DTU Compute, Danmarks Tekniske Universitet, 9) Datalogisk Institut, Københavns Universitet

Ugeskr Læger 2025;187:V10240748. doi: 10.61409/V10240748

HOVEDBUDSKABER

- Kunstig intelligens (AI) kan forbedre ultralyddiagnostik ved at automatisere processer og støtte klinikere.
- Begrænsninger ved AI og regulatoriske udfordringer kan hindre integration i klinikken.
- Før AI kan anvendes til diagnosticering, er der behov for implementeringsstudier for at sikre succesfuld klinisk anvendelse.

Ultralyd (UL)-skanning er en grundsten i føtalmedicin og danner basis for diagnostik og monitorering af fostre. Derfor er uddannelsen inden for dette subspecialiserede område også fokuseret på at udvikle UL-kompetencer. UL-skanning er forbundet med lange læringskurver [1], som bliver yderligere udfordret i føtalmedicin grundet fosterets varierende position samt bevægelser i løbet af undersøgelsen. Den enkelte klinikers ekspertise og erfaring afspejles i detektionsrater af føtale misdannelser [2, 3], hvilket afspejles i højere detektionsrater på højt specialiserede centre [4]. Andre faktorer såsom overvægt eller præeksisterende medicinske tilstande som f.eks. diabetes kan også give forskelle i detektionsrater [5, 6]. Således vil kvaliteten af diagnostikken i svangreomsorgen potentielt afhænge af, hvem man er, hvor og af hvem man bliver behandlet.

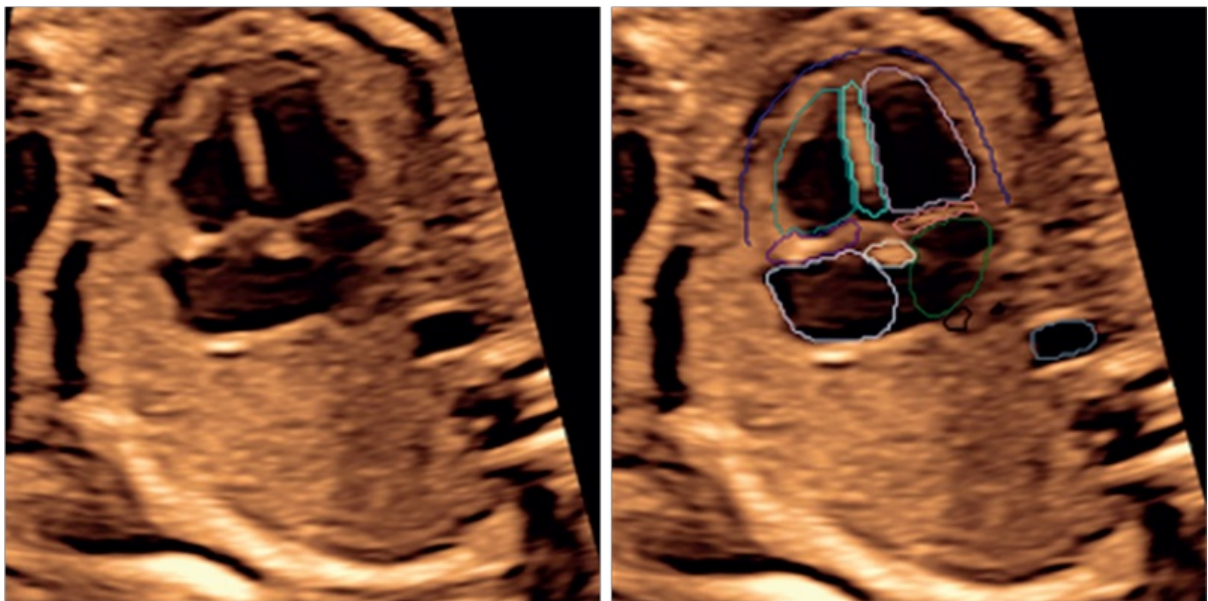
I de senere år er simulationsbaseret UL-oplæring blevet mere udbredt inden for gynækologi, obstetrik og føtalmedicin i forbindelse med oplæring af nye læger, jordemødre og sonografer [7]. Simulationsbaseret UL-oplæring giver et kontrolleret og uforstyrret læringsmiljø, hvor hånd-øje-koordination, systematisk tilgang til undersøgelserne og genkendelse af normal anatomi kan trænes. Det har vist sig, at initial simulationsbaseret træning har reduceret behovet for supervision i den efterfølgende kliniske oplæring [8]. Men læringen fortsætter langt efter simulationskurserne, og her rummer kunstig intelligens (AI) et stort potentiale i føtalmedicin. AI kan støtte klinikerne i den daglige praksis, optimere workflowet ved at automatisere billedlagring og målinger, give feedback på skanningernes komplementhed og dermed frigive tid til patientcentreret behandling og samtale. Desuden kan AI potentielt bidrage til at ensrette diagnostikken på tværs af centre ved at objektivisere undersøgelsen til en vis grad og evaluere, hvornår kvaliteten af UL-billeder er gode nok til at kunne udelukke patologi og sikre optimal føtaldiagnostik. På andre områder såsom brystkræftscreening er implementeringen af

AI i klinisk praksis allerede nået længere. Her har AI vist sig at kunne reducere radiologernes arbejdsbyrde og samtidig forbedre screeningens kvalitet [9]. Men før man lader sig begejstre og overestimerer den nye teknologis potentiale, bør udfordringer og begrænsninger ved implementering af denne nye teknologi vurderes og undersøges nærmere.

Hvor er vi i dag?

Der er udviklet en lang række avancerede AI-modeller inden for føtalmedicin. Disse modeller kan bl.a. genkende standardplaner og udføre automatiske mål, hvilket er nyttigt til vægtskanning, føtalneurosonografi og føtalekkokardiografi [10, 11] Se **Figur 1** for et eksempel på automatiske optegninger. Der er ligeledes udviklet modeller til automatisk at screene for distinkte hjertemisdannelser [12] eller til at identificere mere subtile forskelle i hjerteanatomien ved hjælp af hjerteopmålinger, som f.eks. er relevant ved screening for medfødte hjertemisdannelser såsom coarctation af aorta, som ellers kan give diagnostiske udfordringer [13]. Til gavn i lavindkomstlande med begrænset adgang til UL-ekspertise, er der udviklet AI-modeller til at undersøge placentalplacering med henblik på at vurdere risikoen for placenta prævia [14, 15] samt til automatisk bestemmelse af gestationsalder ved UL-undersøgelser udført af personer uden træning i UL-skanning [16].

FIGUR 1 AI-genereret optegning af relevant anatomi på et firekammerbillede.



Antallet af publikationer om AI i føtalmedicin er steget markant inden for de seneste år. Selv om mange automatiserede modeller har vist imponerende resultater i retrospektive studier, er der kun få studier, der har undersøgt deres gavnlige effekter i klinisk praksis – hvem drager nytte af dem og til hvilket formål? I en prospektiv validering af en model er det ikke overraskende blevet klart, at billedkvaliteten påvirker modellens ydeevne [17]. Dette er vigtigt at tage i betragtning ved implementeringen af disse modeller, da dårlig billedkvalitet er en almindelig udfordring i føtalmedicin på grund af variation i f.eks. fosterposition, BMI og operatørfaring. Implementeringen af AI-modeller kan kun blive succesfuld, hvis man forstår modellens begrænsninger, og hvis udviklerne har tænkt over, hvordan disse udfordringer kan håndteres.

Hvor skal vi hen?

Inden for føtalmedicin vil klinikerens arbejde ikke kunne overtages af AI-modeller. Den menneskelige kontakt og empati vil ikke kunne erstattes, og derudover er AI-modeller langt fra at kunne levere kritisk tænkning tilpasset den enkelte patient. AI derimod excellerer på områder, hvor mennesker ofte har begrænsninger: Det kan analysere store mængder data hurtigt og præcist, udføre opgaver konsistent og objektivt, behandle information effektivt og automatisere rutinemæssige opgaver [18]. På denne måde komplementerer klinikerne og AI hinanden godt og kan sammen skabe en mere effektiv og præcis diagnostik.

For at opnå dette samarbejde i klinisk praksis er der behov for yderligere forskning, der undersøger, hvordan AI-modeller bedst kan integreres i klinikken. Flere studier tyder på, at AI-modeller ikke opnår de bedste resultater på egen hånd, men fungerer bedst i et samarbejde med klinikerne [19-21]. Målet bør være, at AI fungerer som en støtte og aflastning frem for at være en distraktion under arbejdet eller en direkte erstatning. For at nå frem til brugbare integrerede modeller skal vi begynde at udvikle og teste modeller i klinisk praksis og tage højde for den heterogene effekt, AI-modeller kan have på forskellige klinikerne [22]. En form for personaliseret AI-assistance, der hjælper novicen med det basale og aflaster eksperten med tidskrævende rutinearbejde, kan være vejen frem. Sidst, men ikke mindst, skal der være fokus på udviklingen i kliniske færdigheder efter oplæring med og brug af AI-værktøjer i dagligdagen, så vi ikke ser ind i en fremtid med mindre kompetente eksperter, der er afhængige af AI-systemer for at kunne udføre deres arbejde.

Det er et naturligt første skridt at udvikle modeller baseret på retrospektive data. Når man har en fungerende model, kan man derefter teste den i klinikken. Forhåbentlig er den store mængde af publikationer om AI-modeller kun begyndelsen på en udvikling, der vil føre til anvendelige og praktiske værktøjer i sundhedsvæsenet.

Udfordringer på vejen

Til trods for forskeres gode intention om og interesse i at implementere AI-modeller i klinisk praksis er den store implementeringsbarriere de regulatoriske udfordringer på området. De nuværende barrierer skyldes bl.a. strenge regler og krav fra myndigheder om medical device regulation. Da AI-teknologi er relativt ny inden for medicinske og juridiske rammer, er myndighederne også ved at finde deres vej, hvilket resulterer i, at nye reguleringstiltag løbende bliver indført. Det betyder dog også, at der er langt fra udvikling af en model, der virker bedre end eksisterende praksis, til implementering i klinisk drift [23]. Et meget groft estimat på, hvad det koster at få en ny AI-model CE-mærket, dvs. regulatorisk godkendt i Europa, er 10+ mio. kr. De ovennævnte tiltag er selvfølgelig med de bedste intentioner for øje, nemlig at sikre patienterne og beskytte deres personfølsomme data. EU's »artificial intelligence act« [24] understøtter dette ved at indføre en risikobaseret tilgang til AI, hvor balancen mellem AI-autonomi og krav til menneskelig kontrol afhænger af applikationens risiko.

Med risikostratificering, patientsikkerhed samt fastholdelse af klinikerens autonomi og ansvar in mente bør målet med AI-modeller til læring og diagnostik formentlig ikke være at skabe fuldautomatiserede systemer som selvkørende biler, men snarere at udvikle værktøjer, der understøtter samarbejdet mellem AI og klinikerne. Et eksempel kunne være AI-modeller, der vurderer kompletheden af en skanning og sandsynligheden for, at der er tale om en normal undersøgelse. Samarbejdet kan også bestå i lavpraktisk, at klinikerne gennemgår og godkender anatomiske optegninger, som AI-modellen har foretaget, før de benyttes til den videre diagnostik. Denne form for samarbejde mellem menneske og AI skaber gennemsigtighed i modellens beslutningsproces og sikrer, at klinikerne bevarer kontrollen over både skanningen og den endelige konklusion. Et andet oplagt eksempel er forbedret diagnostik og detektion inden for områder, hvor vores eksisterende modeller er dårlige – f.eks. estimering af fostervægt eller truende præterm fødsel. Forskning inden for AI-modeller har også fokuseret

på at udvikle modeller, der kan give forklaringer, XAI (explainable AI) [25, 26], og på at bevæge sig væk fra de såkaldte »black box« AI-modeller, hvor det er uklart, hvad der driver modellens beslutninger. Hidtil har man forsøgt at forklare AI-modellers beslutninger ved hjælp af »heat maps«, der fremhæver de områder af et billede, som påvirker beslutningen. Nye metoder inden for XAI arbejder på at integrere koncepter i modellen såsom anatomiske landemærker, der kan bruges til at forklare klinikerens, hvorfor et billede f.eks. ikke er tilstrækkeligt godt, hvis det pågældende område ikke er tydeligt visualiseret [25].

Den transparente tilgang til samarbejdet mellem AI og kliniker hjælper også med at adressere de etiske overvejelser om ansvar i tilfælde af fejldiagnostik, da klinikerens har mulighed for at interagere med og kritisk evaluere AI-modellens beslutninger. Det er særlig vigtigt i føtalmedicin, hvor screeningsresultater kan føre til svære dilemmaer for patienten, f.eks. om man ønsker at fortsætte eller afbryde en graviditet, hvis der er alvorlige misdannelser.

Derudover hviler et betydeligt ansvar på udviklerne af AI-modellerne, da en model kun er så god som den data, modellen er trænet på. AI-modeller har ofte svært ved at generalisere til forskellige populationer [27], hvilket gør det afgørende, at modeller testes på forskellige datasæt [28, 29], og at deres begrænsninger er klart defineret og forstået før klinisk implementering. Derfor er det vigtigt, at vi værner om vores gode nationale registre og villigheden til at dele data, da gode AI-modeller gerne skal kunne håndtere alle typer af patienter, inklusive sjældne cases og forskellige demografiske karakteristika.

Som for andre områder kommer AI ikke til at erstatte klinikere i den føtalmedicinske UL-diagnostik, men klinikere, der bruger AI, kommer nok med tiden til at erstatte klinikere, der ikke bruger AI [30]. Men adgangen til data, implementeringsstudier og systematisk klinisk evaluering af kliniske AI-støtteredskeer kommer til at være essentielle for, at vi i fremtiden kan opnå den størst mulige værdi af AI i føtalmedicin – til gavn for både klinikere og de gravide.

Korrespondance *Caroline Taksøe-Vester*. E-mail: caroline.amalie.taksoee-vester@regionh.dk

Antaget 31. januar 2025

Publiceret på ugeskriftet.dk 28. april 2025

Interessekonflikter Der er anført potentielle interessekonflikter. Forfatterens ICMJE-formularer er tilgængelige sammen med artiklen på ugeskriftet.dk

Referencer findes i artiklen publiceret på ugeskriftet.dk

Artikelreference Ugeskr Læger 2025;187:V10240748

doi 10.61409/V10240748

Open Access under Creative Commons License [CC BY-NC-ND 4.0](https://creativecommons.org/licenses/by-nc-nd/4.0/)

SUMMARY

Artificial intelligence and ultrasound in fetal medicine

Ultrasound is essential in fetal medicine for diagnosing and monitoring, but it requires extensive training. Artificial intelligence (AI) shows a great promise in enhancing the clinical training and practice, by improving workflow and standardising diagnostics. Despite its potential, AI's limitations and regulatory challenges must be addressed before full integration. Research should focus on implementation to achieve successful use in clinical practice. Ultimately, AI is expected to support clinicians rather than replace them, improving diagnostic accuracy and efficiency in fetal medicine.

REFERENCER

1. Tolsgaard MG, Ringsted C, Dreisler E et al. Sustained effect of simulation-based ultrasound training on clinical performance: a randomized trial. *Ultrasound Obstet Gynecol.* 2015;46(3):312-8. <https://doi.org/10.1002/uog.14780>
2. Tegnander E, Eik-Nes SH. The examiner's ultrasound experience has a significant impact on the detection rate of congenital heart defects at the second-trimester fetal examination. *Ultrasound Obstet Gynecol.* 2006;28(1):8-14. <https://doi.org/10.1002/uog.2804>
3. Corcoran S, Briggs K, O'Connor H et al. Prenatal detection of major congenital heart disease – optimising resources to improve outcomes. *Eur J Obstet Gynecol Reprod Biol.* 2016;203:260-3. <https://doi.org/10.1016/j.ejogrb.2016.06.008>
4. Lytzen R, Vejlsstrup N, Bjerre J et al. Live-born major congenital heart disease in Denmark incidence, detection rate, and termination of pregnancy rate from 1996 to 2013. *JAMA Cardiol.* 2018;3(9):829-837. <https://doi.org/10.1001/jamacardio.2018.2009>
5. Dashe JS, McIntire DD, Twickler DM. Effect of maternal obesity on the ultrasound detection of anomalous fetuses. *Obstet Gynecol.* 2009;113(5):1001-1007. <https://doi.org/10.1097/AOG.0b013e3181a1d2f5>
6. Tsai PJS, Loichinger M, Zalud I. Obesity and the challenges of ultrasound fetal abnormality diagnosis. *Best Pract Res Clin Obstet Gynaecol.* 2015;29(3):320-7. <https://doi.org/10.1016/j.bpobgyn.2014.08.011>
7. Taksøe-Vester C, Dyre L, Schroll J et al. Simulation-based ultrasound training in obstetrics and gynecology: a systematic review and meta-analysis. *Ultraschall Med.* 2021;42(6):e42-e54. <https://doi.org/10.1055/a-1300-1680>
8. Tolsgaard MG, Ringsted C, Rosthøj S et al. The effects of simulation-based transvaginal ultrasound training on quality and efficiency of care: a multicenter single-blind randomized trial. *Ann Surg.* 2017;265(3):630-637. <https://doi.org/10.1097/SLA.0000000000001656>
9. Lauritzen AD, Lillholm M, Lyng E et al. Early indicators of the impact of using AI in mammography screening for breast cancer. *Radiology.* 2024;311(3):e232479. <https://doi.org/10.1148/radiol.232479>
10. Ramirez Zegarra R, Ghi T. Use of artificial intelligence and deep learning in fetal ultrasound imaging. *Ultrasound Obstet Gynecol.* 2023;62(2):185-194. <https://doi.org/10.1002/uog.26130>
11. Horgan R, Nehme L, Abuhamad A. Artificial intelligence in obstetric ultrasound: a scoping review. *Prenat Diagn.* 2023;43(9):1176-1219. <https://doi.org/10.1002/pd.6411>
12. Arnaout R, Curran L, Zhao Y et al. An ensemble of neural networks provides expert-level prenatal detection of complex congenital heart disease. *Nat Med.* 2021;27(5):882-891. <https://doi.org/10.1038/s41591-021-01342-5>
13. Taksøe Vester CA, Mikolaj K, Petersen OBB et al. Role of AI& assisted automated cardiac biometrics in screening for fetal coarctation of aorta. *Ultrasound Obstet Gynecol.* 2024;64(1):36-43. <https://doi.org/10.1002/uog.27608>
14. Saavedra AC, Arroyo J, Tamayo L et al. Automatic ultrasound assessment of placenta previa during the third trimester for rural areas. I: IEEE International Ultrasonics Symposium, IUS. Vol 2020-September, 2020. <https://doi.org/10.1109/IUS46767.2020.9251764>
15. Schilpand M, Neff C, van Dillen J et al. Automatic placenta localization from ultrasound imaging in a resource-limited setting using a predefined ultrasound acquisition protocol and deep learning. *Ultrasound Med Biol.* 2022;48(4):663-674. <https://doi.org/10.1016/j.ultrasmedbio.2021.12.006>
16. Stringer JSA, Pokaprakarn T, Prieto JC et al. Diagnostic accuracy of an integrated AI tool to estimate gestational age from blind ultrasound sweeps. *JAMA.* 2024;332(8):649-657. <https://doi.org/10.1001/jama.2024.10770>
17. Taksoee-Vester CA, Mikolaj K, Bashir Z et al. AI supported fetal echocardiography with quality assessment. *Sci Rep.* 2024;14(1):5809. <https://doi.org/10.1038/s41598-024-56476-6>
18. Tolsgaard MG, Feragen A, Grierson L. Leveraging education science for AI-clinician collaboration in the patient care ecosystem. *Ann Surg Open.* 2024;5(2):e442. <https://doi.org/10.1097/AS9.0000000000000442>
19. Day TG, Matthew J, Budd S et al. Artificial intelligence to assist in the screening fetal anomaly ultrasound scan (PROMETHEUS): a randomised controlled trial. *medRxiv* 2024.05.23.24307329. <https://doi.org/10.1002/uog.28023>
20. Tschandl P, Rinner C, Apalla Z et al. Human-computer collaboration for skin cancer recognition. *Nat Med.* 2020;26(8):1229-1234. <https://doi.org/10.1038/s41591-020-0942-0>
21. Lin M, Zhou Q, Lei T et al. Deep learning system improved detection efficacy of fetal intracranial malformations in a randomized controlled trial. *NPJ Digit Med.* 2023;6(1):191. <https://doi.org/10.1038/s41746-023-00932-6>

22. Yu F, Moehring A, Banerjee O et al. Heterogeneity and predictors of the effects of AI assistance on radiologists. *Nat Med*. 2024;30(3):837-849. <https://doi.org/10.1038/s41591-024-02850-w>
23. Gerdes A, Fasterholdt I, Rasmussen BSB. AI ethics in healthcare. *Ugeskr Læger*. 2024;186(28):V09230600. <https://doi.org/10.61409/V09230600>
24. European Union. Regulation (EU) 2024/1689 of the European Parliament and the Council of 13 June 2024 laying down harmonised rules on artificial intelligence and amending regulations (EC) no 300/2008, (EU) no 167/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (artificial intelligence act) (text with EEA relevance), 2024. <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A32024R1689> (12. dec 2024)
25. Lin M, Feragen A, Bashir Z et al. I saw, I conceived, I concluded: progressive concepts as bottlenecks. *arXiv:2211.10630v1*. (online 19. nov 2022). <https://doi.org/10.48550/arXiv.2211.10630>
26. Olsen MDS, Amsdorf J, Lin M et al. Unsupervised detection of fetal brain anomalies using denoising diffusion models. *arXiv:2408.03654*. (online 7. aug 2024). <https://doi.org/10.48550/arXiv.2408.03654>
27. Ternov NK, Christensen AN, Kampen PJT et al. Generalizability and usefulness of artificial intelligence for skin cancer diagnostics: an algorithm validation study. *JEADV Clinical Practice*. 2022;1(4). <https://doi.org/10.1002/jvc2.59>
28. Tolsgaard MG, Svendsen MBS, Thybo JK et al. Does artificial intelligence for classifying ultrasound imaging generalize between different populations and contexts? *Ultrasound Obstet Gynecol*. 2021;57(2):342-343. <https://doi.org/10.1002/uog.23546>
29. Sendra-Balcells C, Campello VM, Torrents-Barrena J et al. Generalisability of fetal ultrasound deep learning models to low-resource imaging settings in five African countries. *arXiv:2209.09610*. (online 20. sep 2022). <http://arxiv.org/abs/2209.09610>
30. Drukker L. Real-time identification of fetal anomalies on ultrasound using artificial intelligence: what's next? *Ultrasound Obstet Gynecol*. 2022;59(3):285-287. <https://doi.org/10.1002/uog.24869>