

TLV

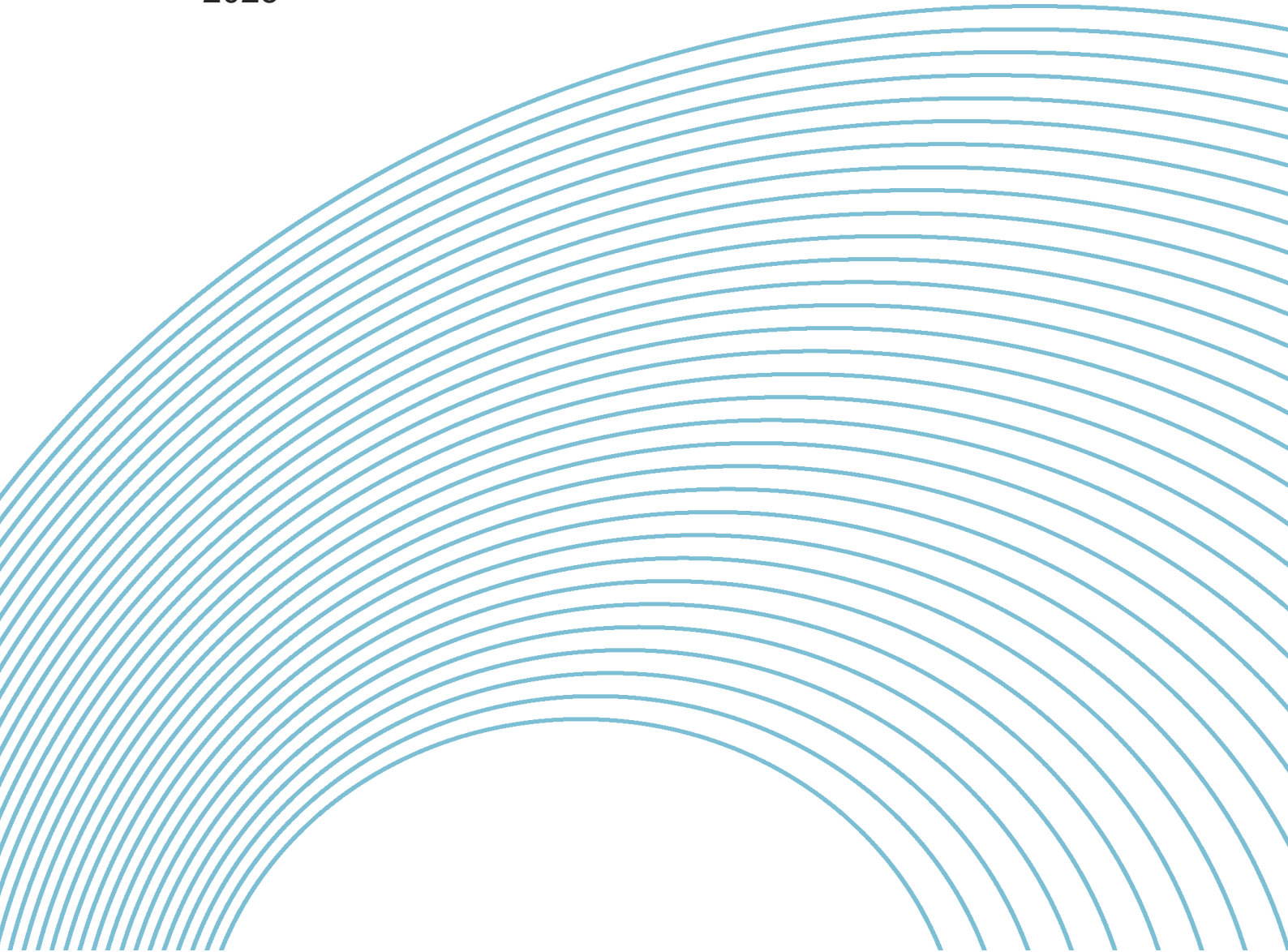
TANDVÅRDS- OCH

LÄKEMEDELSFÖRMÅNSVERKET

Behovsstyrd tandvård

– En modell för framtiden

2025



Tandvårds- och läkemedelsförmånsverket, TLV, är en statlig myndighet som beslutar om vilka läkemedel och förbrukningsartiklar som ska ingå i läkemedelsförmånerna, som i vardagligt tal kallas högkostnadsskyddet. TLV beslutar vilken tandvård som ska omfattas av högkostnadsskyddet i det statliga tandvårdsstödet och bedriver tillsyn över apoteksmarknaden.

TLV:s vision är mesta möjliga hälsa för skattepengarna.

Citera gärna Tandvårds- och läkemedelsförmånsverkets rapporter, men glöm inte att uppge källa: Behovsstyrd tandvård – En modell för framtiden, 2025, Tandvårds- och läkemedelsförmånsverket.

Tandvårds- och läkemedelsförmånsverket, 2025
registrator@tlv.se

Kontaktperson: Elena Borsci

Diarienummer: 02460/2024

Postadress: Box 22520, 104 22 Stockholm
Besöksadress: Fleminggatan 14, Stockholm
Telefon: 08 568 420 50
www.tlv.se

Förord

TLV avrapporterade i mars 2024 ett regeringsuppdrag om att stödja Socialstyrelsens arbete med att bedöma förutsättningarna för ett införande av en nationell modell för riskbedömning inom tandvården, med förslag på en modellprototyp baserad på historisk tandvårdskonsumtion. I juni 2024 fick vi ett nytt regeringsuppdrag om att stödja Socialstyrelsens arbete med att utveckla en nationell riskbedömningsmodell.

Tandhälsan är god för stora delar av befolkningen, men det finns utrymme för förbättringar. Genom att utnyttja de senaste metoderna inom prediktiv modellering kan tandvården i Sverige bli ännu bättre på att identifiera patienter i riskzon och erbjuda förebyggande behandling. På så sätt kan patienters hälsa förbättras samtidigt med optimering av tandvårdens resurser.

I rapporten beskriver TLV olika metoder för att göra riskbedömningar av tandvårdens patienter samt för- och nackdelar med dessa. TLV står fast vid att riskbedömningar bör vara automatiserade och inkludera både historisk tandvårdskonsumtion och information om patientens befintliga tandhälsa.

Agneta Karlsson
Generaldirektör, TLV

Innehåll

Förord	3
Sammanfattning	6
Summary	8
1 Inledning	10
1.1 Bakgrund	10
1.2 Fortsatt uppdrag	11
1.3 Grundläggande förutsättningar	11
1.4 Dialog med andra myndigheter	12
2 Vägval – innovativ riskbedömningsmodell för att möta framtida utmaningar	13
2.1 Munhälsoprofil	14
2.2 Datorstödda beslut är framtiden	14
3 Resultat med prediktionsmodeller	16
3.1 Vilken data som används	16
3.1.1 Datakällor	16
3.1.2 Urval av data	17
3.1.3 Vilka utfallsmått och riskindikatorer har vi valt och varför?	18
3.1.4 Datahantering och transformationer	20
3.2 Vi testar tre modeller	21
3.3 Hur modellerna utvärderas	21
3.3.1 Metod 1: Kalibreringskurva	22
3.3.2 Metod 2: ROC-kurva	24
3.3.3 Metod 3: Realiserad relativrisk	27
3.3.4 Jämförelse av modellerna	28
4 Balansen mellan manuella och automatiserade bedömningar	29
4.1 Svagheter med långtgående manuell riskbedömning	29
5 Framtiden: från prototyp till produktion	32
5.1 Förbättring av modellen som kan göras på kort sikt	32
5.1.1 Förbättrad tidsindelning	32
5.1.2 Modellera ordningen av åtgärder	32
5.1.3 Alternativa utfallsvariabler	33
5.1.4 Data från fler patienter och fler år	33
5.1.5 Inkludera fler variabler i analysen	34
5.2 Förbättring över längre tid som kräver lagändring	34
5.2.1 Vilken information har Försäkringskassan idag?	34
5.2.2 Utforska de praktiska förutsättningarna för sammanvägning av munhälsoprofilen	35

5.3	Hur man kan förklara ett automatiskt beslut för patienten	36
5.3.1	Vad är marginaleffekter	36
5.4	Interoperabilitet och produktionssättning av prediktionsmodellen	37
5.4.1	Praktiska förutsättningar för att införa riskbedömningsmodellen ...	38
5.4.2	Utveckling och interoperabilitet	38
5.4.3	Tekniska krav på plattformen för interoperabilitet.....	39
5.4.4	Uppdateringsprocess för riskbedömningsmodellen.....	39
5.5	Kontinuerlig anpassning efter det nya tandvårdsstödet.....	40
	Projektorganisation	42
	Referenser	43
	Bilaga 1. Termer och begrepp.....	45
	Bilaga 2: Hur vi byggde våra djupinlärningsmodeller	50
	Bilaga 3: Definition av utfallsmått.....	51
	Bilaga 4: Ytterligare modellutvärdering.....	55

Sammanfattning

Regeringen har gett Tandvårds- och läkemedelsförmånsverket (TLV) i uppdrag att stödja Socialstyrelsen med att utveckla en nationell modell för riskbedömning. Målet med modellen är att dela in patienter i olika grupper baserat på deras risk för framtida problem med munhälsan. Riskbedömningen kan användas för att bestämma hur ofta patienter behöver gå på kontroll och vilken tandvård som de behöver utifrån deras individuella behov.

Metoderna för hur patienternas risk ska bedömas och hur prioriteringen mellan patienter i olika riskgrupper ska ske bör väljas utifrån vetenskaplig evidens och beprövad erfarenhet. Hänsyn måste också tas till budgetmässiga ramar och praktiska förutsättningar. Det är viktigt att modellen fungerar som ett stöd för tandläkare och tandhygienister, så att bedömningar av patienterna kan göras objektivt och enhetligt över landet.

En riskbedömningsmodell som ska införas för vuxentandvård (och i förlängningen kanske även barn- och ungdomstandvården) måste klara de stora volymer av patienter som varje år passerar svensk tandvård. Tre miljoner basundersökningar utförs varje år i vuxentandvården, och de flesta av dem skulle behöva inkludera en riskbedömning. Modellen behöver även vara användarvänlig, kunna följas upp och inte vara möjlig att manipulera.

TLV har tagit fram en riskbedömningsmodell som bygger på automatiskt beslutsstöd med hjälp av artificiell intelligens (AI). En sådan modell kan hantera ett stort antal bedömningar till en låg administrativ kostnad för både Försäkringskassan och för vårdgivarna. TLV föreslår att befintlig IT- infrastruktur och de omfattande data som finns i tandvårdsstödet används för att bygga en träffsäker riskbedömningsmodell.

En AI-styrd riskbedömningsmodell kan ta hänsyn till mer information än vad en enskild behandlare har tillgång till. Modellen kan också upptäcka sjukdomar tidigt, vilket ökar chansen att patienter får behovsanpassad behandling i tid. Välgrundade beslut och tidiga insatser kan leda till bättre tandhälsa för hela befolkningen. Vår bedömning är att en AI-styrd riskbedömningsmodell är en samhällsekonomiskt effektiv investering.

TLV presenterar en djupinlärningsmodell som i 84 fall av 100 korrekt kan skilja på en patient med högt vårdbehov från en med lågt vårdbehov. Två andra modeller som vi tog fram har även de god träffsäkerhet. Resultatet kan förbättras med förfinade prediktionsmodeller och ytterligare data från exempelvis patientjournaler, läkemedelsregister och andra databaser. Resultaten från prediktionsmodellerna kan göras tydliga för både behandlare och patienter, så att de förstår besluten som fattas.

För att dra nytta av den kunskap om artificiell intelligens (AI) som finns i Sverige samarbetar TLV med forskningsvärlden och Life Science-sektorn. Vi ser en stor potential i att vidareutveckla AI-styrda beslutsmodeller för att skapa en mer effektiv och behovsanpassad tandvård som skulle gynna både patienter och vårdgivare.

Summary

The Dental and Pharmaceuticals Benefits Agency (TLV) has been tasked by the Swedish government to support the National Board of Health and Welfare in the development of a national risk assessment model for dental care. The aim of the model is to divide patients into groups based on their risk for future problems with dental health. This risk assessment can be used to determine the frequency of dental check-ups and the dental care required to satisfy the unique needs of each patient.

The methods for how the patients' risk should be determined and how priorities between patients in different risk groups should be made ought to be based on scientific evidence and proven experience. Consideration must be taken to budgets and practical feasibility. It is important that the model functions as a support for dentists and dental hygienists to ensure that the assessment of patients is objective and consistent throughout the country.

A risk assessment model to be implemented in adult dental care (and in the long run also the child and youth dental care) must be able to handle the large volume of patients who pass through Swedish dental care every year. Three million basic examinations are made annually in the adult dental care and most of them would need to include a risk assessment. The model also needs to be user friendly, able to be followed-up on and resistant to manipulation.

The risk assessment model delivered by TLV is an automated decision support system backed by Artificial Intelligence (AI). Such a model can handle numerous assessments at a low administrative cost for both the Swedish Social Insurance Agency and the dental care providers. TLV proposes that the current IT-infrastructure and extensive data available for the dental care benefits is used to build a risk assessment model with high accuracy.

An AI-operated risk assessment model can take more information into account than any single practitioner would have access to. The model could also discover diseases at an early stage, increasing the chances that the patient is able to receive the treatment they need in time. Decisions based on sound evidence and early interventions can improve the dental health of the entire population. As such, in our assessment, an AI-operated risk assessment model would constitute an effective investment for society.

TLV presents a deep learning model that can accurately differentiate between a patient with high dental care needs and one with low dental care needs in 84 out of 100 instances. Two other models that we developed also had acceptable accuracy. The results can be improved with refined prediction models and further data from, for example, patient journals, pharmaceutical registries and other databases. The results from the prediction models can be clarified for both practitioners and patients, so that they understand the decisions being made.

To utilise the knowledge about the Artificial Intelligence (AI) that exist in Sweden, TLV collaborates with academia and the Life Science sector. We see great potential in further developing AI-operated decision models to create more efficient and needs-based dental care that would benefit both patients and providers.

1 Inledning

1.1 Bakgrund

Inom tandvården har riskbedömningar länge varit en viktig del av arbetet. Det finns många olika system, både i Sverige och internationellt, som har tagits fram för att förutsäga hur olika orala sjukdomar kan utvecklas. Vissa av dessa system har utvärderats vetenskapligt, men det saknas fortfarande enhetlighet. Det behövs även forskning om effekterna av den odontologiska riskbedömning som rutinmässigt görs för varje patient.¹ Det kan också vara skillnad på bedömningar utförda av tandläkare och tandhygienister eftersom de har olika formell kompetens.

En annan viktig faktor som ofta förbises är tillgången till historiska data om patienten. Trots att patienttrörligheten är relativt låg, är det av flera anledningar ovanligt att en patient går till samma tandläkare under hela sitt liv. Bristen på tillgängliga tandläkare kan vara en bidragande faktor. Det innebär att tandläkaren ofta har begränsad information om patientens sjukdomshistoria.

För att hantera akuta behov räcker det ofta att bedöma patienternas nuvarande tandstatus och lyssna på deras berättelse. För att kunna välja rätt behandling, eller ingen behandling alls, är det viktigt att observera tandrelaterade sjukdomar över en längre tid. I Sverige prioriterar tandvården icke-invasiva behandlingar, vilket innebär att man till exempel undviker att borra i tänderna om mindre kariesangrepp inte har blivit större över flera år. Det är vanligt att kalla patienter till basundersökning med individuella intervall även om de är helt friska. I andra länder söker patienter ofta tandvård endast vid akuta problem. Detta är ett skäl till att tandhälsan i Sverige ständigt förbättras.²

Ett nationellt verktyg för att bedöma patientens risk samt tandvårdsbehov kan fungera som stöd för dagens behandlare. Det skulle stärka den individuella bedömningen, göra den jämlik och också bidra till en ännu bättre tandhälsa.

I juni 2023 fick TLV i uppdrag av regeringen att stödja Socialstyrelsen med att bedöma möjligheterna att införa en nationell modell för riskbedömning inom tandvården. Målet med denna modell är att kunna förutse risken för framtida sjukdomar och se till att tandvården prioriteras utifrån varje patients behov. I mars 2024 presenterade TLV en modellprototyp som delar in patienter i tre grupper: friska, de med medelrisk eller med hög risk.³ Denna indelning baseras på tidigare tandvårdsdata och återspeglar svenska tandläkares bedömningar sedan det statliga tandvårdsstödet infördes 2008. Vår analys visade att det finns ett starkt samband

¹ Statens beredning för medicinsk och social utvärdering (SBU), SBU. *Prediktionsmodeller för karies och parodontit*, 2024.

² Socialstyrelsen, Statistik om tandhälsa, 2023.

³ TLV, *Riskbedömning i tandvården – En nationell modell baserad på historisk tandvårdskonsumtion*, dnr 00763/2024, 2024.

mellan de sjukdomar som omfattas av det statliga tandvårdsstödet och patientens tandhälsa. Detta gör det möjligt att bedöma risker på individnivå. Vårt arbete har visat att friska patienter går till tandläkaren oftare än vad som rekommenderas enligt Socialstyrelsens nationella riktlinjer, och de som har hög risk går för sällan. Analysen visade också att yngre personer oftare tillhör den friska gruppen, medan äldre, särskilt de över 65 år, ofta tillhör den grupp som har hög risk.

I sin rapport bedömde Socialstyrelsen förutsättningarna för att införa en nationell modell för riskbedömning inom tandvården i Sverige som goda.⁴ Socialstyrelsen har undersökt hur riskbedömningar används inom tandvården idag. De har konstaterat att riskbedömning är en vanlig metod, men det finns inte tillräckligt med vetenskapligt stöd för just riskbedömningsprocessen. Det är något som man behöver ta hänsyn till i framtida arbete. Socialstyrelsen har i sin rapport uppmärksammat att många vårdgivare som deltog i undersökningen anser att tandläkaren borde kunna justera riskbedömningar manuellt. Tandläkaren skulle själv kunna ändra bedömningen av en patientens risk för tandproblem. Socialstyrelsen anser också att riskbedömningar inte ska kopplas till det statliga tandvårdsstödet vilket skulle kunna leda till onödig stress för tandläkaren. Dessutom menade Socialstyrelsen att vissa grupper, som ofta har större behov av behandling, bör prioriteras.

1.2 Fortsatt uppdrag

TLV fick i juni 2024 ett fortsättningsuppdrag att stödja Socialstyrelsen i deras uppdrag att fortsätta utveckla den nationella modellen för riskbedömning inom tandvården.⁵

I uppdragsbeskrivningen förtydligade regeringen att riskbedömningen bör inkludera flera faktorer som påverkar utvecklingen av tandhälsan och ska kombineras med behandlarens kliniska bedömning. Den slutliga riskbedömningen ska sedan ligga till grund för beslut om revisionsintervallen och den tandvård som patienten ska erbjudas. En färdig modell bör leda till enhetliga bedömningar i tandvården, mer rättvis fördelning av tandvårdsresurser och en stärkt kontroll över det statliga tandvårdsstödet.

1.3 Grundläggande förutsättningar

När myndigheterna utvecklar framtida riskbedömningar för tandhälsa är det viktigt att det inte bara fokuseras på de faktorer som påverkar hur tandhälsan förändras. Hänsyn behöver också tas till andra aspekter som kan påverka den övergripande orala hälsan. Detta innebär att myndigheterna gemensamt måste se till helheten och inte bara enskilda delar vid bedömning av riskerna för tandhälsa i framtiden. Även patienternas ekonomi och statens subvention har stor betydelse för

⁴ Socialstyrelsen, *Förutsättningar för en nationell modell för riskbedömning inom tandvården*, 2024, s. 72.

⁵ Socialdepartementet, *Uppdrag att utveckla en nationell modell för riskbedömning inom tandvården*, S2024/01260 (delvis).

tandhälsan. Modellen behöver också ta hänsyn till behandlarens bedömning i viss omfattning. Genom att väga samman allt detta kan vi bättre förstå och bidra till att förbättra den orala hälsan hos befolkningen. TLV tolkar detta som att en modell för riskbedömning bör innehålla flera olika delar. Dessa delar ska tillsammans ge en så heltäckande riskbedömning som möjligt. Den kliniska bedömning som behandlaren gör är en viktig del av den slutliga bedömningen, men den får inte vara den enda del som systemet bygger på.

TLV tolkar regeringens fortsatta uppdrag som att riskbedömningsmodellen ska kopplas ihop med fördelningen av det statliga tandvårdsstödet. I praktiken betyder det att riskbedömningen kommer att vara en grund för revisionsintervallens längd och vilken tandvård patienterna ska få tandvårdsersättning för. Riskbedömningen blir därför en viktig del i beslutsprocessen för att se till att patienterna får rätt vård vid rätt tidpunkt.

TLV ser uppdraget som en möjlighet att bidra till att alla i Sverige får en god tandhälsa. Genom att införa ett effektivt riskbedömningssystem ges förutsättningar att identifiera och hantera risker i tid, vilket i sin tur kan leda till bättre tandvård och förebyggande insatser. Vårt mål är att bidra till att alla, oavsett var de bor eller deras ekonomiska situation, har tillgång till den tandvård de behöver för att upprätthålla en god munhälsa.

1.4 Dialog med andra myndigheter

TLV har haft ett flertal möten med Socialstyrelsen, Försäkringskassan och Statens beredning för medicinsk och social utvärdering (SBU) för att utarbeta gemensamma strategier. SBU har tidigare konstaterat att det saknas tillräckligt med vetenskapligt underlag för att avgöra om prediktionsmodeller för karies och parodontit är tillförlitliga. Myndigheten har utökat frågeställningen och undersöker det vetenskapliga underlaget inom bettavvikelse, bettfysiologi och munslemhinneförändringar.⁶ Försäkringskassan understryker vikten av att riskbedömningar ska kunna kontrolleras. Myndigheten förespråkar automatisering av denna process.

Socialstyrelsen har samlat en bred grupp experter som arbetar med att hitta viktiga faktorer som bör ingå i framtida modell för riskbedömning. Detta konsensusarbete är både viktigt och omfattande och lägger grunden för att vi tillsammans ska kunna utveckla en tillförlitlig modell för riskbedömning i framtiden.

⁶ Statens beredning för medicinsk och social utvärdering (SBU), *SBU. Prediktionsmodeller för karies och parodontit*, 2024.

2 Vägval – innovativ riskbedömningsmodell för att möta framtida utmaningar

Riskbedömningsmodellen har som huvudsyfte att identifiera vilka patienter som har störst behov av tandvård. Modellen adresserar problemet med att behandlarnas tid inte alltid läggs på de med störst behov. Med en standardiserad och välfungerande modell kan både staten och vårdgivarna förbättra prioriteringarna och se till att resurserna används där de gör mest nytta.

För att hitta patienter med hög risk för framtida tandsjukdom behöver vi använda all den erfarenhet vi har samlat på oss och det är viktigt att tänka i nya och kreativa banor. Det finns många olika sätt att bedöma risker, och det är viktigt att välja rätt verktyg. Valet av riskbedömningsmodell bör grundas på de specifika mål och syften som uppdraget avser att uppfylla. Eftersom tandvården har unika data ända från 2008 anser TLV att det bästa sättet att uppnå målen är genom att använda de senaste framstegen inom artificiell intelligens (AI).

AI är ett verktyg som alla andra med både fördelar och nackdelar. Utvecklingen inom området går mycket snabbt, och det är viktigt att den svenska staten bidrar till att AI används på ett säkert och ansvarsfullt sätt i vården, men samtidigt inte missar de möjligheter som finns att få ut mer hälsa för vårdpengarna.

AI-kommissionens rapport, som syftar till att forma Sveriges nationella AI-strategi, belyser potentialen i att använda historiska data och aktuella trender. Ett exempel är att AI kan hjälpa till att förutsäga naturkatastrofer så att myndigheterna kan vara beredda och planera åtgärder i tid.⁷ På samma sätt bedömer TLV att tandvården i framtiden kan använda AI för att förutsäga utvecklingen av munhälsoproblem. Det skulle göra det möjligt att agera förebyggande i ett tidigt skede och därmed förbättra den allmänna munhälsan.

Det är dock viktigt att användning av AI sker på ett kontrollerat och etiskt sätt för att garantera att individens integritet skyddas. För att säkerställa att AI-verktyg inom tandvården främst gynnar patienterna, måste staten ta ledningen i utvecklingen. Annars riskerar vi att förlora kontrollen över vilka faktorer och incitament som styr dessa verktyg. I kommersiella system kan AI-teknik främst användas för att maximera ersättningar, snarare än att gynna patienterna.

När det gäller det statliga tandvårdsstödet, som bygger på tillit, är målet med AI-stödd riskbedömning inte att ersätta tandläkaren. Syftet är i stället att ge tandläkaren bättre underlag för att kunna fatta de bästa möjliga besluten. Vidare

⁷ SOU 2025:12 *AI-kommissionens Färdplan för Sverige*, 2024, s. 50.

blir det mer rättvist för patienterna när alla i hela Sverige riskgrupperas på samma sätt. AI-styrd riskgruppering gör det också enklare att ge en skriftlig motivering till varför patienten fick den riskgrupp hen fick.

2.1 Munhälsoprofil

Utredningen om tiotandvård föreslår i sitt betänkande att ersätta ålder som kriterium för förstärkt högkostnadsskydd med en behovsbaserad riskbedömning.⁸ För att göra det föreslår utredningen att det införs något som kallas munhälsoprofil. Munhälsoprofilen ska fånga patientens behov av att återställa förmågan att tugga och tala utan hinder, samt bedöma patientens risk för framtida munsjukdomar.

Munhälsoprofilering handlar om att använda information från den aktuella undersökningen av patientens munhälsa. Det innebär också att det tas hänsyn till den tidigare tandvård patienten fått.

Målet med denna förändring är att säkerställa att de statliga resurserna för tandvård används på bästa sätt och når de personer som har störst behov. Att bedöma munhälsa är en komplicerad process som påverkas av många olika faktorer. För att göra det på ett effektivt sätt behöver vi noggrant utveckla och utvärdera våra metoder. Automatisk riskbedömning ger oss möjlighet att kontinuerligt analysera data och därmed identifiera potentiella risker för tandhälsan i ett tidigt skede. Den automatiska riskbedömningen blir en central komponent i bedömningen av patientens munhälsoprofil.

2.2 Datorstödda beslut är framtiden

Att göra en bedömning av patientens framtida munhälsa kan vara utmanande. Tidsbrist, begränsad erfarenhet och brist på information om patientens tidigare sjukdomshistoria kan göra det svårt att göra en enhetlig bedömning, trots att de flesta tandläkarmottagningar har välgrundade riktlinjer. En modell som enbart bygger på tillit till tandläkarna kan dessutom öka risken för välfärdsbrott. Inspektionen för socialförsäkringen (ISF) påtalar denna problematik i sitt svar på remissen ”Tiotandvård – ett förstärkt högkostnadsskydd för tandvård”. ISF anser att det är mycket viktigt att Försäkringskassan får bättre möjligheter att kontrollera utbetalningar av statligt tandvårdsstöd, särskilt om en tandvårdsreform genomförs. Det är synnerligen viktigt eftersom den föreslagna utformningen av tandvården kan leda till överbehandling och överanvändning av stödet för både patienter och vårdgivare. Anledningen är att patientens låga kostnader minskar incitamenten att begränsa åtgärder och behandlingar. Försäkringskassans arbete med efterhandskontroller är redan nu tidskrävande, och myndigheten i större utsträckning skulle behöva upptäcka risker för felaktiga utbetalningar innan utbetalning sker.⁹ Det finns en gräns för hur mycket resurser som kan fördelas till kontroller. Detta ökar betydelsen av att finna en adekvat lösning på problemet.

⁸ SOU 2024:70 *Tiotandvård – ett förstärkt högkostnadsskydd för tandvård*, 2024, s. 25.

⁹ Inspektionen för socialförsäkringen, *remissvar SOU 2024:70 Tiotandvård - ett förstärkt högkostnadsskydd för tandvård*, S2024/01892, s. 2 och 7.

Därför är det viktigt att använda den information som redan samlas in av tandvården för att skapa en effektiv modell, bland annat en del av den information som registreras i patientjournalen. TLV är eniga med Socialstyrelsen om att patientens aktuella munstatus, inklusive karies, parodontit, tidigare utförda behandlingar, bettfysiologi men även munvårdsvanor och allmänhälsa är sådan information som ska ingå i en riskbedömningsmodell. Man kan överväga att i en myndighetsföreskrift reglera vilka faktorer som ska föras in i en riskbedömningsmodell och på så sätt öka kontroll och transparens. Detta ska dock inte blandas ihop med regler för subvention.

Vi behöver välja en modell som både uppfyller de övergripande målen och som är praktisk att använda i en vårdssituation. TLV har noga övervägt de olika alternativ som finns. En modell som kan uppfylla dagens mål och även vara hållbar i framtiden bör vara automatiserad, lätt att vidareutveckla och gå att kontrollera.

Teknikutveckling med AI-stöd innebär att det inte är nödvändigt att välja mellan automatiska modeller och tillitsbaserade riskbedömningsmodeller som enbart tar hänsyn till tandläkarens bedömning. Det är möjligt att kombinera dessa olika bedömningar i automatiserade system. Sådana system kan dra nytta av den senaste tekniken och den erfarenhet som varje behandlare har.

TLV har samarbetat med experter inom biostatistik, life science och data science. Dessa experter har spetskompetens inom forskning om prediktionsmodeller inom hälso- och sjukvården. Tillsammans med experterna har vi kommit fram till att tandvården i Sverige har unika förutsättningar att skapa en fungerande prediktionsmodell. Även om tandvården på vissa sätt skiljer sig från hälso- och sjukvården finns det gemensamma drag som gör att en sådan prediktionsmodell kan byggas på liknande sätt. Det statliga tandvårdsstödet har lagt en perfekt grund för datainsamling. Vi har idag en stor mängd strukturerade data från över 150 miljoner tandvårdsbesök sedan 2008. Det underlättar arbetet och gör det möjligt att använda prediktionsmodeller med hög träffsäkerhet.

TLV utgår ifrån att om det införs en nationell riskbedömningsmodell kommer den också användas för att bestämma patientens bidrag i statliga tandvårdsstödet. Patienten har enligt lag rätt att få en förklaring till hur ett myndighetsbeslut gick till.¹⁰ Vi har därför från början strävat mot att de prediktionsmodeller vi tar fram ska kunna förklaras för allmänheten.

¹⁰ 32 § Förvaltningslag (2017:900).

3 Resultat med prediktionsmodeller

För att kunna sammanfatta hela behandlingshistoriken på ett praktiskt sätt använder TLV prediktionsmodeller. En prediktionsmodell är ett statistiskt verktyg som hjälper oss att analysera data systematiskt och göra välgrundade förutsägelser. Målet med modellerna är att förutsäga risken för att en patient i framtiden ska behöva omfattande tandvård.

I detta kapitel visar vi att det är möjligt att skapa prediktionsmodeller som med god träffsäkerhet kan förutsäga vilka patienter som kommer att behöva omfattande tandvård i framtiden. Med träffsäkerhet menar vi att modellen ska kunna skilja mellan patienter med stora och små tandvårdsbehov så exakt som möjligt. Vi strävar efter att förbättra denna träffsäkerhet så mycket det går. Här beskrivs också de metoder som används för att välja och utvärdera dessa modeller inom tandvården. Vi förklarar också hur modellernas träffsäkerhet säkerställer att de ska kunna användas som grund för beslut om tandvårdsstöd. Det läggs också särskild vikt vid att förklara hur vi mäter och utvärderar modellernas förmåga att identifiera patienter med hög risk som har behov av mycket tandvård.

3.1 Vilken data som används

3.1.1 Datakällor

Sverige har ett unikt utgångsläge för att utveckla en prediktionsmodell för riskbedömning inom tandvården. Det beror på att Försäkringskassan sedan 2008 har samlat in detaljerad information om samtliga utförda tandvårdsåtgärder inom det statliga tandvårdsstödet, vilket omfattar ca tio miljoner tandvårdsbesök årligen. Denna databas innehåller information om såväl behandlingar och diagnoser som patientspecifika faktorer, vilket möjliggör utvecklingen av en träffsäker prediktionsmodell baserad på verkliga vårddata.

Tandvårdsstödet som etablerades 2008 ställde krav på en samlad och löpande registrering av tandvårdskonsumtionen på individnivå som skulle rapporteras i ett elektroniskt system. Vårdgivare fick skyldighet att lämna uppgifter om all utförd ersättningsberättigande tandvård. Förutom uppgifter om datum för slutförda åtgärder och vårdgivarens pris skulle dessutom tandnummer, tandposition samt uppgifter om bakgrund till utförd vård rapporteras. För att hantera dessa uppgifter i databaser skapades tillståndserier som grundade sig på de diagnoser som förekommer i tandvården enligt den internationella sjukdomsklassifikationen ICD-10. Underavsnitt K00-K14 till ICD-10 ("sjukdomar i munhåla, spottkörtlar och käkar") utgör grunden till de tillstånd som tandvårdsstödet tillämpar men även andra delar av ICD 10 användes. Det var dock tvunget att förenkla de cirka tusen diagnoser för munhålan som anges i kodverket för att kunna hantera det utan att belasta vårdgivarna administrativt. Andra tillstånd behövde utvecklas för att kunna betala korrekt ersättning, till exempel tandlöshet indelades beroende på dess

omfattning. Tillstånd som inte utgår från en diagnos i strikt mening behövde också skapas, till exempel trasig protes eller en implantatskruv som gått sönder.

Åtgärderna och deras koder fick likna de som tidigare tandvårdsstöd hanterat och som var välkända i Sverige och internationellt. Arbetsmoment och material som ingick i åtgärderna hade som utgångspunkt att de skulle vara kostnadseffektiva och vävnadsbesparande vid tillämpning. Åtgärdernas kostnader beräknades enligt detta.

Uppgifter om tandantal och intakta tänder ska lämnas till Tandhälsoregistret som ett grovt mått på tandhälsa. Att det enbart blev tandantal beror på att regeringen vid den tidpunkten bedömde att det skulle bli administrativt betungande att kräva in fler uppgifter.¹¹ Regeringen lämnade dock öppet att om det finns skäl kan det övervägas om vårdgivarna i framtiden ska åläggas skyldighet att lämna ytterligare uppgifter.

Sammantaget finns omfattande strukturerade datauppgifter sedan år 2008 om befolkningens tandvårdskonsumtion. Styrkan är detaljeringsgraden och att det förmodligen inte sker en underrapportering eftersom den är förknippad med ersättning. I Sverige finns därmed datauppgifter med hög kvalitet om tandvårdskonsumtion och produktion.

Dessa datauppgifter kan användas för olika syften:

- Följa produktion av tandvård över landet.
- Fördelning geografiskt och demografiskt, per vårdgivare, mottagning och behandlare.
- Följa konsumtion av tandvård på grupp- eller individnivå vad gäller patienter.
- Ge en bild av tandhälsan i den vuxna befolkningen och de framtida vårdbehoven.
- Uppgifter om tandantal och tandförluster tillsammans med ålder ger en god uppfattning om framtida behov av kostsam rehabilitering.

Vårdgivares nyttjande av tandvårdsstödet kan följas. Indikationer på avvikelser mot normal användning kan upptäckas genom olika algoritmer.

De stora datamängderna i kombination med nya AI verktyg kan medföra att ytterligare ny kunskap om tandvårdskonsumtion och tandvårdsproduktion kan tas fram.

3.1.2 Urval av data

När man utvecklar modeller för att förutsäga risker för dålig tandhälsa är det viktigt att testa dem på rätt sätt. Detta gör vi genom att dela upp data i två delar som innehåller olika patienter och olika tidsperioder. Den första delen, träningsdata,

¹¹ Prop. 2007/08:49 *Statligt tandvårdsstöd*, s. 100-101.

använder vi för att träna själva modellen. Den andra delen, testdata, använder vi för att kontrollera hur bra modellen fungerar när den möter nya patienter som den inte sett tidigare. Det här är ett vanligt sätt att dela upp data inom maskininlärning. Att skilja på data för utveckling och utvärdering är av stor vikt för att säkerställa modellernas tillförlitlighet i klinisk praxis, där den kommer att tillämpas på framtida patienter under varierande förhållanden.

För att säkerställa att en prediktionsmodell är träffsäker är det avgörande att utvärdera den på data som inte använts vid utvecklingen av modellen. Genom att använda data från olika patientgrupper undviker vi att modellen blir överanpassad till särskilda egenskaper hos den ursprungliga patientgruppen som faktiskt inte korrelerar med tandhälsa i hela befolkningen. Detta skulle kunna leda till missvisande resultat vid praktisk användning.

Figur 1 Behandlingsperiodens och uppföljningsperiodens längd.



Vid klinisk driftsättning kommer riskbedömningsmodellen att behöva förutsäga utfall ett år i framtiden. Därför testas hur väl modellen kan prediktera utfall genom att använda historiska data för att förutsäga framtida händelser. Figur 1 visar längden på behandlingsperioden och uppföljningsperioden för både träningsdata och testdata. För att träna modellen användes data för 61 050 patienter från åren januari 2016 – augusti 2022 för att förutsäga dålig tandhälsa under september 2022 – augusti 2023. För att testa modellen användes data för 15 057 andra patienter från perioden januari 2017 – augusti 2023 för att förutsäga dålig tandhälsa under september 2023 – augusti 2024. Tränings- och testperioden var olika eftersom vi ville undvika att prediktionsmodell anpassar sig för mycket till de data som används vid träningen.

3.1.3 Vilka utfallsmått och riskindikatorer har vi valt och varför?

För att utveckla en tillförlitlig prediktionsmodell för riskbedömning inom tandvården använder vi flera informationskällor. Den viktigaste informationskällan är patientens tidigare tandvårdskonsumtion, där vi analyserar både antal och typ av tandvårdsåtgärder samt tidsåtgången för dessa. Hänsyn tas också till åtgärder som utfördes under covid-19-pandemin (mars 2020 till februari 2021) eftersom tandvårdskonsumtionen under denna period gick ner kraftigt jämfört med tidigare år. Patientens tandhälsa mäts genom hur många tänder som skadades eller förlorades samt antalet kvarvarande och intakta tänder vid det sista

undersökningstillfället. Modellen inkluderar även demografiska faktorer som kön, ålder och civilstånd, där det särskilt beaktas att ålderns påverkan på tandhälsan kan vara olika för män och kvinnor. Information om patientens hemkommun, där kommunerna är indelade i tre huvudgrupper:

- Storstad eller storstadsnära kommun.
- Större stad eller kommun nära större stad.
- Mindre stad/tätort eller landsbygdskommun.

Slutligen tar modellen hänsyn till om patienten huvudsakligen besöker privata tandvårdskliniker eller Folk tandvården. All denna information samlas in under behandlingsperioden och används som riskindikatorer.

För att kunna skapa träffsäkra prediktionsmodeller behövs utfallsmått under uppföljningsperioden som:

- Kan mätas med de åtgärder som fanns i tandvårdsstödet för 2023–2024.
- Ingår i en behandling av betydande omfattning och som inte görs i onödan, för att undvika att vi överskattar patientens vårdbehov.

Vilka riskindikatorer som ska ingå i modellen är avgörande för dess utveckling. Identifieringen av lämpliga indikatorer bör göras specifikt för varje oral sjukdom, varvid dessa sedan bör balanseras mot varandra och andra relevanta riskfaktorer såsom allmänhälsa, munvårdsvanor, medicinering och socioekonomiska förhållanden. Tandlossningssjukdom (parodontit) är ett exempel där flera faktorer påverkar prognosen, såsom när sjukdomen debuterar, hur snabbt stödjevådnaden går förlorad och hur sjukdomen utvecklas över tid.¹²

Socialstyrelsen arbetar med att identifiera de faktorer som har störst inverkan på tandhälsan. Samtidigt utvecklar TLV en automatiserad modell för att använda dessa faktorer och maximera modellens förmåga att förutsäga sjukdomens utveckling. I detta arbete använder Socialstyrelsen sig av mätbara och lätt verifierbara variabler. På sikt kommer det även att behövas mer nyanserade och svåråmätbara variabler, så kallade "mjuka" variabler. Dessa kan vara mer subjektiva, men de kan också vara avgörande för att upptäcka sjukdomar i ett tidigt skede.

För denna rapport valdes två typer av utfallsmått som indikerar dålig tandhälsa. De är baserade på åtgärder och tillstånd i tandvårdsstödet:

- Det första utfallsmåttet var om patienten hade fått omfattande vård för parodontit, periimplantit eller andra sjukdomar i vävnaderna runt tänder eller implantat. Lite mer än två procent av alla patienter behövde sådan vård under uppföljningsåret.
- Det andra utfallsmåttet var om patienten behövde dra ut eller rotfylla minst en tand, där utdragningar av visdomständer inte räknades med. Drygt sex

¹² Álfheiður Ástvaldsdóttir, m fl, *ÅRSRAPPORT 2023 Svenskt Kvalitetsregister för Karies och Parodontit*, SKaPa, s. 44.

procent av patienterna behövde denna typ av behandling under uppföljningsåret.

De exakta definitionerna av dessa utfallsmått finns i bilaga 2 i denna rapport. Våra prediktionsmodeller beskriver således risk som sannolikheten att behöva dessa typer av behandlingar. Det skapades också ett tredje utfallsmått baserat på de första två, som visar om patienten behövde parodontal behandling eller förlorade en tand.

I våra modeller väljer vi olika strategier för att hantera den naturliga fördelningen av utfallsvariablerna, där tandförlust förekommer i sex procent av fallen och parodontit i två procent. För de logistiska modellerna behålls den ursprungliga fördelningen för att modellens prediktioner ska spegla den verkliga risken i befolkningen, vilket underlättar tolkningen av resultaten. För våra djupinlärningsmodeller viktas däremot observationerna i omvänd proportion till utfallet eftersom dessa modeller är mer känsliga för obalanserade utfall. För att säkerställa att samtliga modeller utvärderas på ett korrekt sätt använder vi metoder som är särskilt lämpade för obalanserade utfall då de tar hänsyn till den låga förekomsten av positiva utfall (se Bilaga 4).

Även riskindikatorerna är obalanserade. Majoriteten av våra riskindikatorer fångar hur ofta en patient fått en viss åtgärd (eller åtgärdsgrupp), antingen mätt som antalet åtgärder eller hur den totala tiden för åtgärderna. Med undantag av basundersökningarna är det få åtgärder som utförs på mer än en liten del av patienterna. För att hantera obalansen har vi därför skapat två typer av riskindikatorer för varje åtgärd eller åtgärdsgrupp.

Den första typen av riskindikator visar om en patient har fått inga eller några av en viss åtgärd. Den andra typen mäter hur många gånger en åtgärd gjorts. Genom att inkludera två riskindikatorer för varje åtgärd kan vi se om skillnaden mellan att inte ha fått någon åtgärd alls och att ha fått en åtgärd är annorlunda än skillnaden mellan att ha fått en och två åtgärder.

Vi har också gjort en justering där vi tar bort riskindikatorer som färre än 100 patienter har fått, vilket tar hand om de mest obalanserade fallen.

3.1.4 Datahantering och transformationer

För att säkerställa en hög kvalitet på de data som används i riskbedömningsmodellen har data bearbetats i flera steg. I analysen av tandförluster och parodontit exkluderades visdomständer, eftersom dessa ofta tas bort av andra skäl än dålig tandhälsa.

För att underlätta jämförelser mellan olika variabler har alla variabler normaliserats så att de bara tar värden mellan noll och ett. Detta steg är nödvändigt för att träningen ska leda till träffsäkra modeller, och är ett etablerat arbets sätt inom maskininlärning.

Kombinationer av åtgärder och tillstånd som förekommer färre än hundra gånger bland de 60 000 patienternas behandlingar exkluderades. Detta gjordes för att

minska risken att träningsdata innehåller kombinationer som inte förekommer i testdata, samt för att undvika felaktigt beräknade värden för dessa riskindikatorer. Alla värden för antal åtgärder och tidsåtgång till den 95:e percentilen kapades, vilket minskar risken att enstaka extremvärden får oproportionerligt stor påverkan på modellens resultat.

Slutligen gjordes en kvalitetskontroll för att identifiera och korrigera uppenbara felregistreringar av antalet kvarvarande tänder som ökar över tid för en patient. Dessa förberedande åtgärder skapar ett stabilt och tillförlitligt underlag för riskbedömningsmodellen.

3.2 Vi testar tre modeller

TLV skattar tre olika typer av modeller, från enkel till komplicerad, för att testa hur man kan balansera träffsäkerhet mot transparens:

1. Tiotandvårdsmodellen: Denna logistiska regression använder endast en riskindikator – om patienten är över eller under 67 år gammal vid riskbedömningen.
2. Logistisk modell med behandlingshistorik: Denna logistiska regression använder alla riskindikatorer som beskrevs i avsnitt 3.1.3. Modellen är mer avancerad och kan fånga upp komplexa samband mellan olika riskindikatorer, samtidigt som den är mer transparent än djupinlärningsmodellen.
3. Djupinlärningsmodell med behandlingshistorik: Denna djupinlärningsmodell använder alla riskindikatorer som beskrevs i avsnitt 3.1.3. Modellen är mer komplex och det kan vara svårare att förstå exakt hur den kommer fram till sina bedömningar. Samtidigt har den tack vare komplexiteten möjligheten att ge en stor förbättring i träffsäkerhet. Modellen beskrivs i mer detalj i bilaga 2.

Alla modeller har i sin tur byggts för att förutsäga de tre utfallsmåtten:

- 1) Tandförlust
- 2) Parodontit
- 3) Tandförlust eller parodontit.

Nedan utvärderar vi prediktionsmodeller för utfall 1 och 2 i detalj, följt av en summering för alla tre.

3.3 Hur modellerna utvärderas

Målet med utvärderingarna är att se till att modellerna ger tillförlitliga bedömningar både på populations- och individnivå. På populationsnivå handlar det om att bedöma hur stor andel av en grupp som kommer att drabbas av tandsjukdomar. Till exempel skulle en modell som förutspår att alla patienter har sju procents risk att förlora en tand under nästa år vara korrekt på populationsnivå, eftersom det är den genomsnittliga risken. Men en sådan modell vore oanvändbar

på individnivå eftersom den inte skiljer mellan patienter med hög och låg risk. På individnivå behöver modellen i stället kunna ta hänsyn till att risken varierar mellan patienter beroende på faktorer som tidigare tandsjukdomar, allmänhälsa, levnadsvanor och munhygien.

Prediktionsmodellen beräknar sannolikheten för att patienten ska förlora tänder eller behöva omfattande tandlossningsbehandling under det kommande året. Denna sannolikhet uttrycks som en procentsats mellan 0 och 100, där exempelvis 4 procent innebär att patienten har en relativt låg risk för tandförlust, medan 70 procent indikerar en betydligt högre risk. 70 procent risk betyder också att i en grupp av 100 patienter med denna risknivå förväntas 70 patienter förlora en eller fler tänder.

Modellen använder patientens tidigare tandvårdshistorik till riskindikatorerna i förutsägelsen. Genom att analysera miljontals tidigare tandvårdsbesök kan modellen identifiera vilka faktorer och mönster som historiskt sett lett till tandförlust eller tandlossning, och därmed skatta framtida risker för varje enskild patient. Dessa individuella riskbedömningar kan sedan användas för att dela in patienter i olika riskgrupper, till exempel låg-, medel- och högriskgrupper. Till skillnad från manuella bedömningar är modellens beräkningar lika för alla patienter med samma sjukdomshistorik, ålder, kön med mera, vilket gör att alla patienter bedöms objektivt, systematiskt och efter samma kriterier.

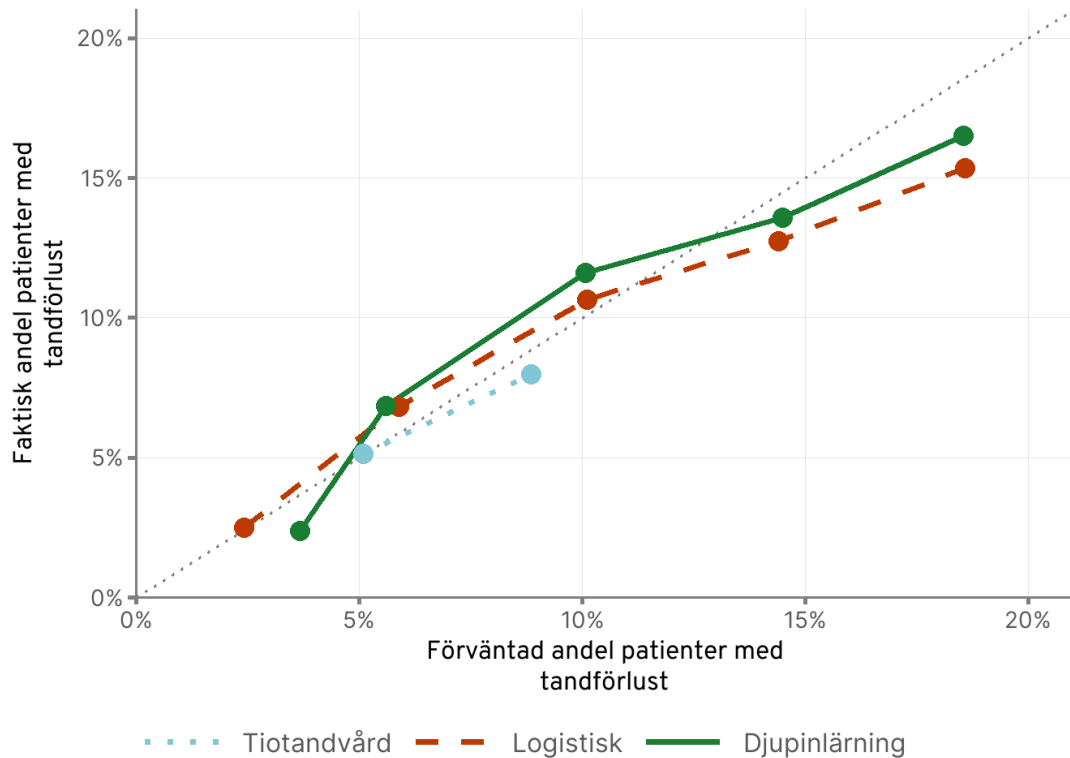
Det finns flera sätt att utvärdera modeller och testa hur bra de är på att skilja på patienter med hög och låg risk för tandförlust eller parodontit. Tre metoder har valts som på olika sätt beskriver modellens förmåga att korrekt skilja patienterna efter risk.

3.3.1 Metod 1: Kalibreringskurva

För att utvärdera modellernas träffsäkerhet använder vi en så kallad kalibreringsanalys. Denna metod undersöker om modellens beräknade risknivåer överensstämmer med faktiska utfall. Till exempel, bland alla patienter där modellen beräknar en 20-procentig risk för tandförlust, bör omkring 20 procent faktiskt förlora tänder under det kommande året. På samma sätt bör cirka tre procent av patienterna i gruppen med tre procents beräknad risk drabbas av tandförlust.

Genom att rita en kalibreringskurva, där modellens beräknade risker jämförs med observerade utfall, kan vi utvärdera denna överensstämmelse för alla risknivåer. En väl kalibrerad modell följer den diagonala streckade linjen i diagrammet, vilket betyder att de beräknade riskerna är rimliga över hela skalan. Kalibreringskurvan visar hur väl förutsägelserna överensstämmer med de faktiska utfallen. Den streckade diagonalen representerar en perfekt kalibrering där de förutsagda sannolikheterna exakt matchar de observerade frekvenserna. På x-axeln ser vi det av modellen förväntade antalet förlorade tänder för den gruppen patienter, och på y-axeln ser vi det faktiska antalet förlorade tänder under uppföljningsåret.

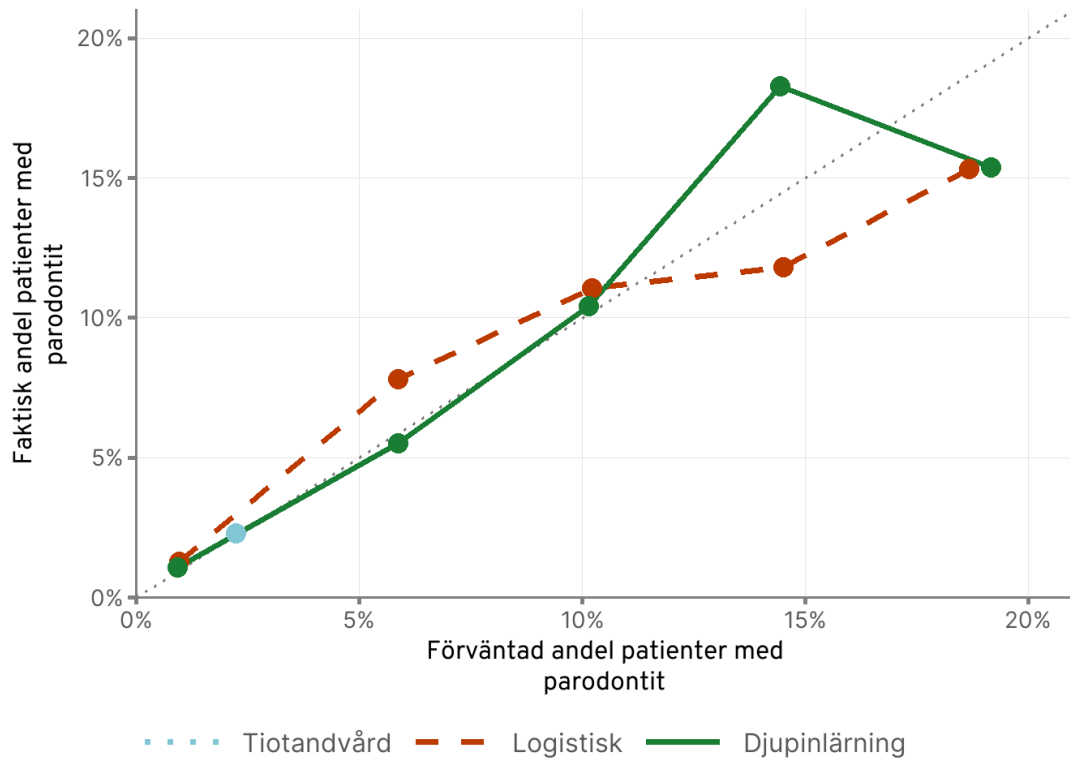
Figur 2 Kalibreringskurva för tandförlust.



Not: Figuren visar hur väl olika modellers förutsägelser stämmer överens med faktiska utfall. Den streckade diagonallinjen representerar perfekt kalibrering där förutsägelse och utfall är identiska. En modell som ligger under den streckade linjen överskattar risken, medan en modell över linjen underskattar risken. Axlarna sträcker sig till 20 procent eftersom endast fem procent av patienterna fick en risk högre än så.

När modellen förutsäger tandförlust uppvisar alla tre modeller en relativt god kalibrering, med grupperna strax över eller under den streckade linjen i figur 2. Tiotandvårdsmodellen har enbart två grupper eftersom patienterna bara kan få en av två riskpoäng (en för patienter under 67 år och en för 67+). Alla modeller är därmed bra på att förutspå antalet patienter med tandförlust bland olika patientgrupper.

Figur 3 Kalibreringskurva för parodontit.



Not: Figuren visar hur väl olika modellers förutsägelser stämmer överens med faktiska utfall. Den streckade diagonallinjen representerar perfekt kalibrering där förutsägelse och utfall är identiska. En modell som ligger under den streckade linjen överskattar risken, medan en modell över linjen underskattar risken. Axlarna sträcker sig till 20 procent eftersom endast två procent av patienterna fick en risk högre än så.

Vid förutsägelse av parodontit i figur 3 uppvisar alla tre modeller också en relativt god kalibrering. Djupinlärningsmodellen följer den streckade linjen nära och uppvisar mycket god kalibrering upp till runt 10 procents risk (vilket är 95 procent av alla patienter). Den logistiska modellen underskattar risken något för patienter med lägre risk och överskattar risken för högriskpatienter.

3.3.2 Metod 2: ROC-kurva

Med riskbedömningsmodellen måste två viktiga mål vägas mot varandra. Det ena är att fånga upp så många patienter som möjligt som faktiskt riskerar att förlora tänder. Det andra är att undvika att klassificera för många friska patienter som högriskpatienter. Överskattning av antalet sjuka ger ineffektiv användning av vårdresurser och onödiga kostnader för både patienter och samhället.

Om vi exempelvis sätter gränsen för högriskgruppen vid tio procents risk för tandförlust eller parodontit kommer modellen att fånga upp de flesta patienter som riskerar att förlora tänder. Samtidigt kommer många patienter som inte skulle ha förlorat några tänder att felaktigt placeras i högriskgruppen. Om gränsen i stället höjs till 20 procents risk kommer färre friska patienter att klassas fel, men då riskerar vi att missa patienter som faktiskt behöver förebyggande vård. Det innebär

att valet av gränsvärde måste baseras på en noggrann analys av konsekvenserna, där både patientsäkerhet och resurseffektivitet vägs in. Hur gränsen ska sättas är i grunden ett politiskt beslut som behöver tas om modellen ska kopplas till tandvårdstersättning.

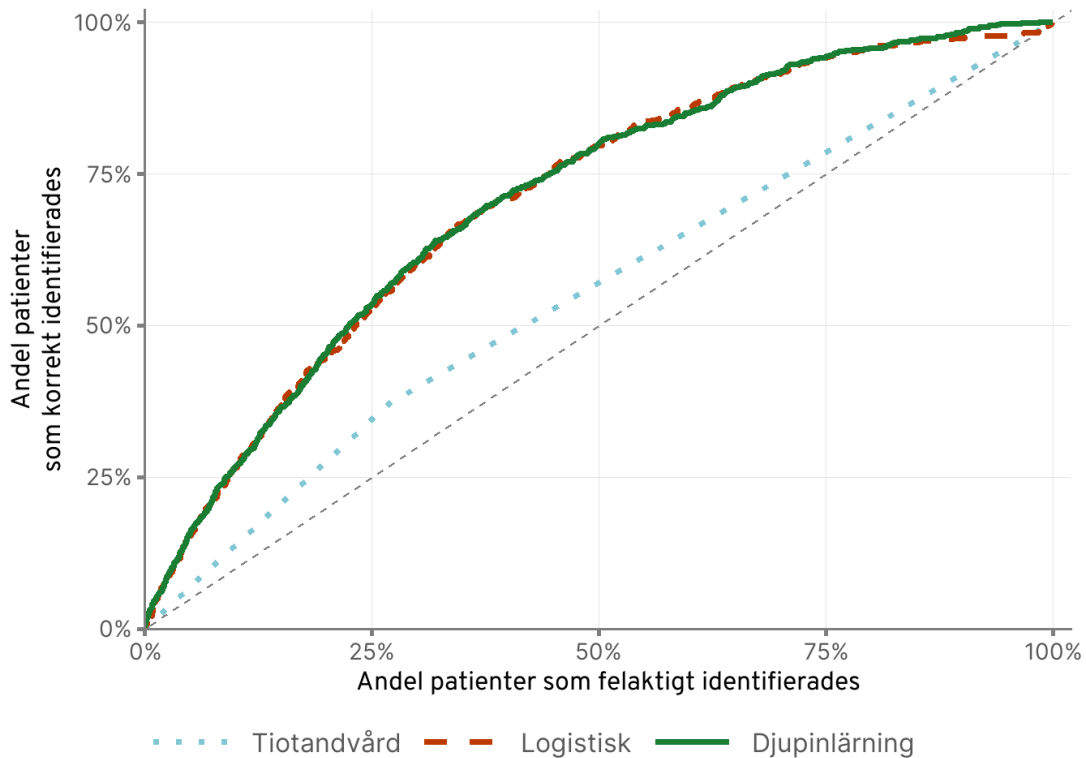
Kalibreringskurvan i förra avsnittet visar inte hur väl modellen rangordnar patienter efter risk, vilket är avgörande för att kunna placera patienter i rätt riskgrupp. En modell kan vara väl kalibrerad men ändå ge samma risknivå till patienter med väsentligt olika vårdbehov. Tiotandvårdsmodellen är ett exempel på detta. Vi har därför valt att använda en annan vanlig metod för att utvärdera modellernas förmåga att korrekt klassificera patienter - ROC-kurvan (Receiver Operating Characteristic).

När man använder en prediktionsmodell för att bedöma risken för tandförlust eller parodontit, är det viktigt att förstå hur väl modellen kan skilja mellan patienter med hög och låg risk. ROC-kurvan är en metod för att göra detta. Den visar hur bra modellen är på att identifiera rätt patienter genom att jämföra hur ofta den förutspår rätt och fel vid olika gränsvärden. Till exempel, om vi tittar på en specifik punkt (X %, Y %) på kurvan, betyder det att när vi sätter gränsen så att Y procent av patienterna som förlorade tänder eller fick parodontit identifieras som högrisk, kommer vi samtidigt att felaktigt klassificera X procent av de patienter som inte riskerar att förlora en tand eller få parodontit.

ROC-kurvan illustrerar tydligt avvägningen mellan att fånga upp verkliga högriskpatienter och att felaktigt klassificera friska patienter som högrisk. Kurvan visar att andelen patienter som korrekt identifieras inte kan höjas utan att andelen patienter som felaktigt identifieras samtidigt ökar. Den streckade linjen skulle varit resultatet om modellen slumpmässigt hade gissat patienternas utfall. En perfekt modell skulle nå punkten (X = 0 procent, Y = 100 procent), vilket skulle innebära att vi identifierar alla högriskpatienter utan några felklassificeringar. Det är dock så gott som omöjligt att uppnå ett perfekt resultat.

Arean under ROC-kurvan (ROC AUC) sammanfattar modellens prestanda över alla möjliga gränsvärden i ett enda mått mellan 0 och 100 procent. Ett ROC AUC-värde på 70 procent betyder att om man slumpmässigt väljer ut en patient som kommer att förlora tänder och en patient som inte kommer att förlora tänder, kommer modellen i 70 procent av fallen att ge en högre riskvärdering till patienten som faktiskt förlorar tänder. En perfekt modell skulle ha ett ROC AUC-värde på 100 procent, medan en modell som gissar slumpmässigt skulle ha ett värde på 50 procent.

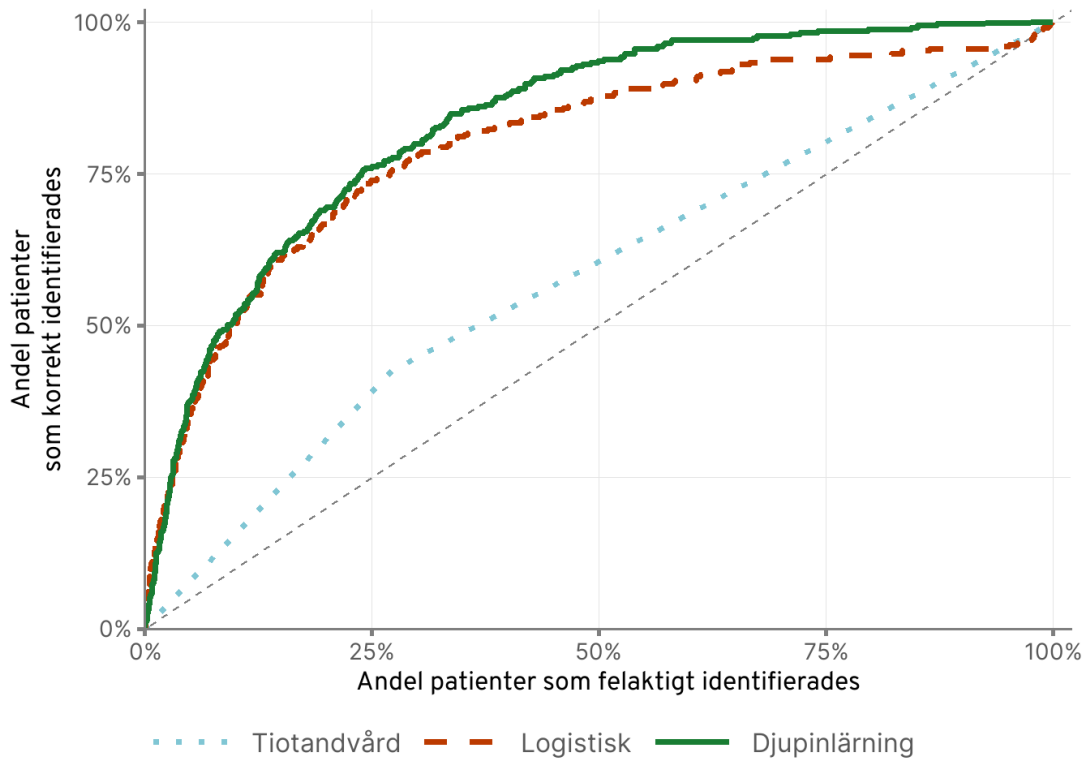
Figur 4 ROC-kurva för tandförlust.



Not: Figuren visar hur bra modellen är på att identifiera patienter som riskerar att förlora tänder. En perfekt modell skulle ha en kurva som går genom övre vänstra hörnet, medan en modell som gissar slumpmässigt skulle följa diagonallinjen.

När vi jämför kurvorna kan vi alltså tolka de modeller vars kurvor som når längre upp i det vänstra hörnet som bättre. I figur 4 som visar ROC-kurvor för tandförlust presterar tiotandvårdsmodellen sämst med en ROC AUC på 55 procent, vilket bör tolkas som att denna modell är bara lite bättre på att hitta högriskpatienter än slumpen. Den logistiska modellen och djupinlärningsmodellen presterar betydligt bättre med ROC AUC med 71 procent för båda.

Figur 5 ROC-kurva för parodontit.



Not: Figuren visar hur bra modellen är på att identifiera patienter som riskerar att få parodontit. En perfekt modell skulle ha en kurva som går genom övre vänstra hörnet, medan en modell som gissar slumpmässigt skulle följa diagonallinjen.

När modellerna förutsäger parodontit presterar de ännu bättre vilket visas i figur 5. Tiotandvårdsmodellen har fortfarande lägst ROC AUC på 58 procent och den logistiska modellen får ett bra resultat på 80 procent. Djupinlärningsmodellen var allra bäst på att sortera patienter efter risk för parodontit med en ROC AUC på 84 procent. Kurvorna för både djupinlärningsmodellen och den logistiska modellen passerar nära punkten $X = 25$ och $Y = 75$, vilket betyder att när modellerna korrekt identifierar 75 procent av patienterna som får parodontit identifierar den samtidigt 25 procent av de friska patienterna felaktigt.

3.3.3 Metod 3: Realiserad relativrisk

För att bedöma hur bra modellerna är på att skilja mellan patienter med hög och låg risk för tandförlust, kan vi titta på resultaten i de mest extrema riskgrupperna. Patienterna delas in i tio lika stora grupper baserat på den risk som varje modell beräknar. Sedan jämförs den grupp som har lägst risk med den som har högst risk. Om modellen fungerar bra, bör det finnas en stor skillnad i antalet patienter som förlorar tänder mellan dessa grupper, en hög realiserad relativrisk.

Till exempel visar vår djupinlärningsmodell att för varje patient som förlorade en tand i tiondelen med lägst risk (de som ligger under den tionde percentilen), förlorade 15 patienter en tand i tiondelen med högst risk (de som ligger över den

nittionde percentilen). Det förhållande ger oss ett enkelt och tydligt sätt att mäta modellens förmåga att hitta de patienter som mest behöver förebyggande tandvård.

Denna skillnad mellan de två ytterlighetsgrupperna visar hur modellen effektivt kan identifiera patienter med mycket olika risknivåer. Det är viktigt för att kunna fördela tandvårdsresurser på ett effektivt sätt.

3.3.4 Jämförelse av modellerna

Tabell 1 Sammanfattning av utfallsmått för tre modeller och tre utfall.

Modell	Utfall	ROC AUC (%)	Realiserad relativrisk
Tiotandvårdsmodell	Tandförlust	55	1
	Parodontit	58	2
	Tandförlust eller parodontit	56	2
Logistisk modell	Tandförlust	71	9
	Parodontit	80	11
	Tandförlust eller parodontit	72	13
Djupinlärningsmodell	Tandförlust	71	15
	Parodontit	84	171
	Tandförlust eller parodontit	70	17

Tabell 1 summerar utvärderingarna från de tre prediktionsmodellerna som utvärderar risken för tre olika utfallsmått relaterade till tandhälsa: tandförlust, parodontit och tandförlust eller parodontit. Utvärderingen av de tre prediktionsmodellerna visar att den logistiska modellen och djupinlärningsmodellen presterar betydligt bättre än tiotandvårdsmodellen. Detta framgår särskilt tydligt för parodontit, där både den logistiska modellen och djupinlärningsmodellen uppnår ROC AUC-värden på 80% respektive 84%. Den realiserade relativrisken, som jämför utfallet mellan grupperna med högst och lägst risk, är också markant högre för dessa modeller. Särskilt anmärkningsvärt är djupinlärningsmodellens mycket höga relativrisk för parodontit, där antalet patienter med parodontit var 171 gånger högre i gruppen med högst risk i jämförelse med gruppen med lägst risk.

4 Balansen mellan manuella och automatiserade bedömningar

Det nuvarande tandvårdsstödet är tillitsbaserad och bygger på att behandlaren har kunskapen nog att välja rätt tillstånd och åtgärd. Det är inte självklart att samma arbetssätt kan appliceras på riskbedömningar. Socialstyrelsen och TLV har under arbetets gång diskuterat en modell som bygger på att behandlaren själv avgör patientens riskgrupp. TLV föreslår att automatisera patientens riskgruppering så mycket som möjligt, eftersom en modell med enbart manuell riskbedömning ökar sannolikheten för att målet om enhetliga bedömningar och mer jämlik tandvård inte uppnås.

Båda formerna av beslutsregler finns redan i tandvårdsstödet. Ett exempel på automatiskt avgörande är det allmänna tandvårdsbidraget (ATB), där Försäkringskassans datorsystem bestämmer om patienten ska få ett enkelt eller dubbelt ATB. Ett exempel på manuellt avgörande är vilken åtgärdskod som används. Om tandläkaren gör en fyllning på en framtand eller hörntand kan hen välja mellan tre olika åtgärder, där valet påverkar tandvårdsstödet storlek. Det enda sättet att kontrollera om rätt åtgärd har använts är genom Försäkringskassans efterhandskontroller.

En betydande nackdel med manuella riskbedömningar är att det blir en mycket hög samhällskostnad för att effektivt verifiera patienternas risknivå för att förhindra felaktig tillämpning. En annan viktig nackdel med manuella riskbedömningar är att den administrativa bördan på behandlarna troligtvis också skulle öka kraftigt.

4.1 Svagheter med långtgående manuell riskbedömning

Det finns många olika faktorer som kan påverka risken för ohälsa i munnen. Dessa faktorer bidrar med olika grader av risk och måste bedömas och vägas samman. Det skulle vara en omöjlig uppgift för varje behandlare att göra de här bedömningarna manuellt på ett tidseffektivt sätt och med ett likvärdigt resultat över hela landet.

En annan viktig aspekt att beakta är risken för att välfärdsbrott i tandvården kan komma att öka. TLV har visat att mer subvention leder till högre vårdgivarpriser i tandvården och att det i sin tur påverkar vilken behandling patienten får.¹³ Behandlarens ekonomiska incitament kan leda till prioriteringar som inte alltid är för patienternas bästa. En svaghet med en modell med manuell riskbedömning är

¹³ TLV, *Empirical studies on reference pricing*, dnr 04076/2024, 2024., TLV, *Examining Heterogeneity in Healthcare Provider Responses to Reference Pricing Through a Synthetic Control Study*, dnr 4763/2024, 2024., TLV, *Effekten av 2008 års tandvårdsreform på vårdgivarpriser*, dnr 04409/2014, 2015.

att den ger behandlaren mycket stort utrymme att bestämma hur mycket stöd patienten ska få, vilket öppnar upp för överutnyttjande och försäkringsbedrägerier.

Manuell bedömning riskerar också öka administrationen för behandlarna. Jämfört med en modell som bygger på individanpassade automatiska beslut skulle en modell med övervägande manuella bedömningar innebära att behandlaren måste ägna betydligt mer tid åt undersökningen än idag.

Den exakta samhällskostnaden med en manuell bedömning går inte att skatta idag, men vi har gjort en tidig beräkning baserat på rimliga antaganden. Anta att patienternas riskbedömning är giltig mellan ett till två år, beroende på risknivå. Efter det behövs det en ny basundersökning med tillhörande riskbedömningar. Fyra miljoner patienter besöker tandvården varje år och det görs tre miljoner basundersökningar. TLV beräknar i referenspriserna att en basundersökning (åtgärd 101) tar i genomsnitt 20 minuter av tandläkarens tid, inklusive administration. Om tandläkaren också ska dokumentera allt som krävs för patientens riskbedömning kan basundersökningen ta 10–25 minuter ytterligare. Samhällskostnaden för enbart basundersökning åtgärd 101 skulle därför öka med upp till 80 %. Om vi även inkluderar undersökningsåtgärderna för tandhygienister, 111 och 112, skulle den ökade administrationen kosta samhället 1,2–3,1 miljarder kronor varje år, vilket motsvarar 6–16 % av omsättningen av all vuxentandvård inom tandvårdsstödet. Om vi utgår från dagens förhållande skulle cirka 1/3 av dessa kostnadsökningar drabba statens finanser och resterande blir en ökad kostnad för patienterna. Med hänsyn till de nuvarande ofta höga kostnaderna för patienterna inom tandvården skulle detta vara en oönskad utveckling. Vår beräkning tar inte i beaktande de samhällsekonomiska kostnaderna för utökad tillsyn och en ökad välfärdsbrottslighet. En annan aspekt av samhällskostnaden är att tandläkarnas tid och resurser tas upp av utökad administration, vilket hindrar dem från att fokusera på mer prioriterad tandvård. Varje minut som läggs på dokumentation är tid som inte kan läggas på att behandla patienterna, allt till patienternas nackdel.

Försäkringskassan skulle ha ytterst begränsade möjligheter att automatiskt kontrollera de riskbedömningar som görs manuellt. Det skulle kräva omfattande manuella efterhandskontroller för att fungera i praktiken. För närvarande utför Försäkringskassan manuella kontroller framför allt på de patientärenden som bedöms ha högst risk för felaktiga utbetalningar. Försäkringskassans har idag enbart kapacitet att utreda mellan 4000 och 5000 sådana patientärenden per år. Utökade manuella kontroller skulle kräva att antalet utredare ökas. För att minska risken för överutnyttjande och försäkringsbedrägerier skulle staten således behöva investera väsentligt mer i efterhandskontroller.

Det finns en påtaglig risk för att statens kostnader för förmånerna kommer att öka, givet de omständigheter som beskrivits. Om subventionssystemet utformas så att det är mer lönsamt att behandla patienter med hög risk finns det risk att fler patienter hamnar i denna grupp än förväntat. Det kan leda till en okontrollerad utgiftsökning, liknande den som skedde efter införandet av 65 plus-systemet för över 20 år sedan. Har riskbedömningen betydelse för om fast eller fri prissättning ska tillämpas kan det bli ett annat utfall. Behandlare kan börja bedöma att

patienterna har lägre risk än vad som är odontologiskt motiverat så att tandvården kan utföras inom det generella högkostnadsskyddet där fri prissättning gäller. Patienten skulle då inte få det ekonomiska stöd hen har rätt till.

5 Framtiden: från prototyp till produktion

5.1 Förbättring av modellen som kan göras på kort sikt

I detta avsnitt presenteras möjliga sätt att på kort tid göra prediktionsmodellerna mer träffsäkra, det vill säga att de med hög precision ska kunna skilja mellan patienter med stora respektive små tandvårdsbehov.

5.1.1 Förbättrad tidsindelning

I det pågående uppdraget delades riskindikatorerna i två tidsperioder: "Under pandemin" (mars 2020-mars 2021) och övriga år. Exempelvis innebär det att ett större kariesangrepp som lagades sju år före uppföljningen antas ha samma effekt på risken att förlora en tand som en fyllning gjord ett år före uppföljningen. Det antagande är inte helt korrekt ur ett odontologiskt perspektiv, men vi gjorde det för att minska antalet variabler i modellen. Färre variabler gör att modellen kan skattas fortare. För prediktion av kariessjukdom kan det exempelvis vara viktigt att inkludera tid i modellen, eftersom de senare årens kariesaktivitet kan ha större betydelse för utfallet än de första årens.

I nästa version av modellen planerar vi att dela upp data i mer än två tidsperioder för att bättre förstå sambandet mellan utfallet och åtgärder som gjorts nyligen jämfört med längre tillbaka. Vi kommer också att undersöka om träffsäkerheten ökar om vi förlänger tidsperioden för riskindikatorerna till mer än sju år, eftersom vi har upp till 16 års data från det statliga tandvårdsstödet att använda.

5.1.2 Modellera ordningen av åtgärder

Ett annat sätt att hantera tiden är att modellera i vilken ordning patienten fick sina åtgärder. Till exempel kan två patienter ha fått samma åtgärder men i olika ordning, vilket kan indikera olika riskprofiler. Vi kan använda moderna djupinlärningsmodeller¹⁴, som de som används för textklassificering, för att modellera ordningen av åtgärder.

Dessa modeller kan identifiera mönster i vilka åtgärder som gjorts och i vilken ordning de rapporterats, precis som de kan skilja mellan positiva och negativa recensioner baserat på ordens ordning. Våra åtgärds-koder kan ses som ord som beskriver vilken behandling patienten fått, vilket gör dem lämpliga för denna typ av modellering.

¹⁴ Exempelvis en- eller flerdimensionella "convoluted neural networks".

En nackdel med dessa modeller är att det kan bli mer omständligt att förklara hur modellen kommit fram till ett visst resultat på ett sätt som uppfyller förvaltningslagens krav, eftersom man inte bara behöver förklara hur patientens åtgärder påverkade utfallet men också i vilken ordning de rapporterades (se avsnitt 5.3).

5.1.3 Alternativa utfallsvariabler

Vi har märkt att modeller som använder "förlorad tand" som utfallsvariabel är mindre träffsäkra än de som använder åtgärder för behandling av parodontit. Det kan bero på att patienter som förlorar en tand inte är en homogen grupp. Olika riskindikatorer kan vara avgörande för varför en patient får en rotfyllning jämfört med varför en annan väljer att dra ut en tand. Därför kan det vara bra att skapa fler modeller med nya utfallsvariabler som är bättre anpassade för specifika odontologiska sjukdomar.

Vi har inte heller tagit hänsyn till tandens position i munnen, trots att var tanden sitter påverkar patientens förmåga att tugga, tala och le. Man kan inte utesluta att patienter kommer i högre utsträckning vilja rädda en tand som syns än en långt bak i munnen. En möjlig lösning är att skapa olika utfallsvariabler för tandförlust beroende på tandens position.

En annan möjlighet är att riskgruppera patienter baserat på hur många behandlingar de fått. Just nu är alla våra utfallsvariabler binära, vilket innebär att vi likställer en patient som fått en rotfyllning under uppföljningsperioden med en som fått fem. Patienter som behöver flera rotfyllningar har sannolikt ett större tandvårdsbehov än de som bara behöver en. Därför måste vi bättre skilja på omfattande och mindre omfattande behandlingar.

Om våra utfallsmått även tar hänsyn till antal behandlingar patienten fick kan vi också modellera mer komplexa beslutsprocesser än idag. Vilken vård patienten får beror på två beslut: att söka vård och att acceptera eller avvisa behandlaren förslag. Hantering av det första beslutet – att söka eller inte söka vård – kan förbättras i vår framtida analys, det finns inspiration att ta från den vetenskapliga litteraturen.¹⁵

5.1.4 Data från fler patienter och fler år

Vi har använt historiska data från cirka sju år och 60 000 patienter. Eftersom tandvårdsstödet infördes 2008 kan vi använda data från över 8 miljoner patienter och 16 år.

Ett större dataset kan hjälpa oss att hitta även svaga samband mellan riskindikatorer och utfallsmått och förbättra träffsäkerheten i våra prediktioner.

¹⁵ Se exempelvis Feng, C. X., A comparison of zero-inflated and hurdle models for modeling zero-inflated count data. *Journal of Statistical Distributions and Applications*, Volume 8, Article number: 8, 2021. <https://doi.org/10.1186/s40488-021-00121-4>.

5.1.5 Inkludera fler variabler i analysen

Omgruppering av åtgärder kan vara ett sätt att skapa nya riskindikatorer. Vi har testat att gruppera åtgärder efter typ av vård, men man kan också använda maskininlärning för att gruppera åtgärder efter hur de ofta kombineras i praktiken.¹⁶ Modeller som explicit tar hänsyn till hur kombinationen av åtgärder påverkar resultatet ser vi också som värda att utforska.¹⁷

Modellen kan modifieras för att bättre ta hänsyn till att tandvårdskonsumtionen varierar mellan olika åldersgrupper men är stabil över tid för samma patienter. Relevanta riskindikatorer kan därmed identifieras för att bättre särskilja ålders- och årskullseffekter.

Det finns också andra datakällor att lägga till, som nationella läkemedelslista. Studier har visat samband mellan läkemedelsanvändning och tandhälsa, till exempel att vissa läkemedel kan orsaka muntorrhet och blödningar i tandköttet.¹⁸

5.2 Förbättring över längre tid som kräver lagändring

Vi beskriver här vilken information Försäkringskassan redan samlar in i sin ärendehantering, vilken information de skulle kunna samla in.

5.2.1 Vilken information har Försäkringskassan idag?

För att riskbedömningsmodellen ska kunna styra hur mycket stöd patienten får inom högkostnadsskyddet för tandvård måste den integreras i Försäkringskassans IT-system. Vi har undersökt vilken information Försäkringskassan redan hanterar.

Försäkringskassan använder information för sina socialförsäkringsrapporter och annan analysverksamhet. De samlar också in information för att hantera de socialförsäkringssystem de ansvarar för.

Information som är relevant för riskbedömningsmodellens träffsäkerhet och som redan hanteras inom det statliga tandvårdsstödet är:

- Information om tillstånd, åtgärder, information om patienten (ålder, kön, ID-nummer etc.) och patientens tandhälsa
- Klinik- och företagsinformation

¹⁶ Exempelvis hur röntgenåtgärder ofta kombineras med akutundersökningar.

¹⁷ Ett exempel från djupinlärningen är "deep and cross networks", där modellen inkluderar information om hur exempelvis effekten av att patienten inte fått förebyggande tandvård samverkar med effekten av att patienten bor i ett glesbygdsområde. Se Wang, R., Fu, G., Fu, B. et al., Deep & Cross Network for Ad Click Predictions, 2017. <https://arxiv.org/abs/1708.05123>.

¹⁸ Esposito, M., Kalinowski, J., Mikhaeil, M., The Effects of Recreational and Pharmaceutical Substance Use on Oral Microbiomes and Health. *Bacteria*, no. 3, 2024. <https://doi.org/10.3390/bacteria3030015>.

Försäkringskassan hanterar också information för att administrera andra socialförsäkringssystem:

- Typ av ekonomiskt stöd (sjukpenning, arbetslöshetsersättning etc.)
- Anställning, avtalad lön, anställningsform och anställningsperiod
- Sjukdomsdiagnos inom olika sjukförsäkringsformer, sjukdomsgrupp, svårigheter och varaktighet för diagnosen
- Civilstånd
- Familjemedlemmar
- Personbegränsningar, såsom aktivitetsbegränsning, funktionsnedsättning och hjälpbehov
- Adresser

Försäkringskassans analytiker får också information från andra myndigheter, såsom:

- Utbildning
- Födelseland
- Befolkningstäthet där patienten bor
- Tandläkartäthet där patienten bor

I nästa steg bör man undersöka hur man kan inkludera dessa uppgifter och de juridiska förutsättningarna för att använda dem.

5.2.2 Utforska de praktiska förutsättningarna för sammanvägning av munhälsoprofilen

I den här rapporten har vi förklarat hur man skapar statistiska modeller för att förutsäga en patients framtida risk för försämrad munhälsa. Dessa modeller beräknar sannolikheten för att en person kan förlora tänder eller behöva omfattande behandling för parodontit. Dessutom har vi beskrivit hur vi i framtiden kan utöka prognosmodellerna med andra diagnoser som patienten har, om patienten tar läkemedel som påverkar munhälsan, samt andra faktorer som kan påverka risken.

Den framtida risken för försämrad munhälsa är dock bara en del av den övergripande munhälsoprofilen. Lika viktigt är det att ta hänsyn till patientens tidigare tandvård (inklusive behandling utanför stödet), nuvarande munhälsa och vårdbehov. Alla aspekter av munhälsoprofilen måste således kunna sammanfattas i en gemensam beslutsregel när patientens munhälsoprofil ska bli en del av ett behovsstyrt tandvårdsstöd.

En munhälsoprofil kommer att leda till en mer rättvis fördelning av tandvårdsresurser och en stärkt kontroll över det statliga tandvårdsstödet. Fördelningen mellan olika riskgrupper av patienterna kommer att påverkas av den budget som avsätts till högkostnadsskyddet. Emellertid kan vi på TLV redan nu börja beskriva olika sätt som beslutsreglerna kan utformas på. Detta planerar vi att fortsätta med i nästa rapport, förutsatt att vi får ett fortsatt uppdrag.

5.3 Hur man kan förklara ett automatiskt beslut för patienten

TLV föreslår att subventioner ska baseras på patientens risk för framtida tandsjukdomar. Vi föreslår att denna risk ska beräknas med hjälp av en statistisk prediktionsmodell. Enligt 32 § förvaltningslagen (2017:900) måste myndighetsbeslut förklara “vilka omständigheter som varit avgörande för myndighetens ställningstagande”. I detta avsnitt förklaras hur man kan uppfylla detta krav när man använder en komplex prediktionsmodell som beslutsverktyg.

Modellerna vi presenterar i denna rapport har tränats på ett omfattande dataunderlag som innehåller många hundra, ibland tusen riskindikatorer. Även om prediktionsmodellerna är matematiskt relativt enkla, gör antalet parametrar dem för komplexa för att man ska kunna förstå hur alla delar av modellen hänger ihop. Denna komplexitet innebär en särskild utmaning när vi ska förklara för patienten hur modellen fungerar, vilket lagen kräver.

5.3.1 Vad är marginaleffekter

Genom att beräkna så kallade “marginaleffekter” av de viktigaste förklaringsvariablerna kan vi förklara vilka faktorer som påverkar en patients riskprofil, utan att behöva gå in på varje matematisk detalj i modellen.

Marginaleffekten visar hur mycket patientens riskprofil förändras när en specifik variabel ändras lite, medan alla andra variabler hålls oförändrade.

Låt oss säga att en patient har fått ett riskpoäng på 79, och gränsen för att hamna i den gruppen med högst risk är 70. Genomsnittspatienten har ett riskpoäng på 50.

Om vi tittar på variabeln “antal rotfyllningar under pandemin”, kan vi se att denna patient fick en rotfyllning medan genomsnittspatienten fick 0,1 rotfyllningar. Vi kallar denna variabel för \mathbf{x} och alla andra förklarande variabler för \mathbf{Z} .

Riskbedömningsmodellen beskrivs med formeln $\mathbf{P}(\mathbf{x}, \mathbf{Z})$, vilket betyder att $\mathbf{P}(\mathbf{x}=1, \mathbf{Z})=79$ för den aktuella patienten. Marginaleffekten är skillnaden i riskpoäng om vi sänker \mathbf{x} från 1 till genomsnittet 0,1.

Marginaleffekten av \mathbf{x} är: $\mathbf{P}(\mathbf{x}=1, \mathbf{Z}) - \mathbf{P}(\mathbf{x}=0,1, \mathbf{Z})$.

Om marginaleffekten är positiv betyder det att patienten fick ett högre riskbetyg än genomsnittet eftersom hen fick fler rotfyllningar under pandemin än genomsnittet.

Det viktigaste att komma ihåg är att marginaleffekten visar **hur mycket** vi förväntar oss att utfallet ska förändras när vi justerar en viss variabel, medan alla andra variabler hålls konstanta. För att ge en samlad bild av “vilka omständigheter som varit avgörande för myndighetens ställningstagande” behöver vi beräkna marginaleffekten för alla variabler. Med flera hundra eller till och med tusen variabler är detta inte rimligt. Det är dock sällan mer än ett fåtal variabler som har stor marginaleffekt, och det räcker med att presentera dem.

Begreppen “stor” och “liten” är subjektiva, och vi behöver ett standardiserat sätt att omvandla siffror till begrepp. Vi kommer därför att skala om alla variabler till en så kallad “standardnormalskala”. Det betyder att 0 för en variabel representerar genomsnittet, 1 betyder att patientens värde för denna variabel är 1 standardavvikelse högre än genomsnittet, och så vidare. Om variabeln är normalfördelad kan vi förvänta oss att cirka 68 procent av patienterna ligger inom -1 till 1 standardavvikelse och 95 procent inom -2 till 2 standardavvikelser.

Tabell 2 Exempel på hur man kan uttrycka avvikelser från genomsnittet i ord

Avvikelse uttryckt i antal standardavvikelser	Avvikelse uttryckt i ord
0 till 1	Måttlig avvikelse. Lite högre/mindre.
1 till 2	Stor avvikelse. Högre/mindre.
2 till 3	Mycket stor avvikelse. Mycket högre/mindre.
>3	Extra stor avvikelse. Extra mycket högre/mindre.

Utifrån vår tabell ovan och vårt tidigare exempel skulle marginaleffekten kunna uttryckas så här:

*“Du tilldelades gruppen ”högriskpatienter” för att ditt riskbetyg var 80 av 100. Ditt riskbetyg indikerar **högre** risk än genomsnittspatienten. En viktig förklaring var att du fick **en** rotfyllning under pandemin, vilket är **mycket högre** än genomsnittspatienten. En annan viktig förklaring var att ...”*

5.4 Interoperabilitet och produktionssättning av prediktionsmodellen

Vi presenterar här ett förslag på hur den tekniska plattformen för riskbedömningsmodellen kan se ut. Riskbedömningsmodellen måste vara interoperabel, det vill säga att den ska kunna integreras och användas i Försäkringskassans IT-miljö eftersom den ska integreras med deras ärendehanteringssystem för det statliga tandvårdsstödet på ett säkert sätt. Riskbedömningsmodellen måste också utformas så att den i framtiden även kan integreras i andra ärendehanteringssystem, såsom regionernas system för barn- och ungdomstandvård.

Riskbedömningsmodellen kommer att behöva uppdateras varje år när TLV ändrar reglerna för det statliga tandvårdsstödet. Det är viktigt att uppdateringen av riskbedömningsmodellen kan ske separat från utvecklingen av Försäkringskassans ärendehanteringssystem.

En förutsättning för kommunikation mellan olika datorsystem är att de talar samma språk. En modern branschstandard för detta är ett REST-API¹⁹, vilket gör det möjligt för en applikation att skicka förfrågningar till en server och få svar i ett standardformat.

5.4.1 Praktiska förutsättningar för att införa riskbedömningsmodellen

Vi kan idag inte säga när riskbedömningsmodellen kan tas i bruk. Det är dock viktigt att redan nu tänka på hur modellen ska användas i praktiken för att undvika problem med integration i befintliga IT-system. Vi kan inte veta exakt hur införandet ska ske förrän vi har det slutgiltiga författningsförslagen, men vi kan göra några rimliga antaganden.

Den första versionen av modellen kommer troligtvis endast att gälla de som har rätt till statligt tandvårdsstöd, vilket innebär att barn- och ungdomstandvård inte ingår.²⁰ Initialt behöver man därför bara fokusera på hur Försäkringskassans system ska kunna kommunicera med riskbedömningsmodellen.

Riskbedömningsmodellen måste bygga på en statistisk prediktionsmodell som tilldelar riskgrupper baserat på sannolikheten för att uppnå en viss tandhälsostatus den kommande perioden. Denna information kommer att baseras på data från socialförsäkringssystemen hos Försäkringskassan. När en behandlare skapar ett nytt ärende i journalsystemet eller inhämtar patientens riskgruppsinformation från Försäkringskassan, sker tilldelning av patientens riskgrupp.

Om patientens riskgrupp påverkar vilket tandvårdsstöd de får, har behandlaren ett starkt ekonomiskt intresse av att försöka placera patienten i den riskgrupp som ger mest intäkt. För att minimera risken för mänskliga fel och säkerställa en rättvis och effektiv process är det av största vikt att beslutet är automatiserat och inte manuellt framtaget av behandlaren.

5.4.2 Utveckling och interoperabilitet

De som utvecklar Försäkringskassans ärendehanteringssystem kan inte vara samma personer som utvecklar riskbedömningsmodellen. Modellen behöver utvecklas av statistiker eller analytiker med programmeringskunskap i till exempel R eller Python, medan Försäkringskassan använder systemutvecklare med erfarenhet av plattformen för deras eget ärendehanteringssystem. Försäkringskassan planerar också att byta ut det nuvarande ärendehanteringssystemet till ett nytt tidigast 2027. Denna skillnad i kompetenskrav innebär att de som tar fram prediktionsmodellen

¹⁹ Ett REST-API är ett sätt för olika program och system att prata med varandra över ett nätverk på ett enkelt och standardiserat sätt. REST-API är populärt för att det är enkelt att använda, och det är byggt för att fungera med vanliga webbprotokoll, som HTTP, vilket gör det lätt att integrera i olika miljöer och applikationer.

²⁰ SOU 2024:70 nämner att modellen ska inkludera både friskfaktorer och riskfaktorer, och att den ska kunna anpassas efter olika patientgrupper, inklusive barn och sköra äldre, samt ta hänsyn till sociala förhållanden och livsstil. Dock finns det ingen specifik information om huruvida barntandvård kommer att ingå i den första versionen av modellen.

inte kommer att vara samma personer som kopplar ihop modellen med Försäkringskassans system.

Det är därför viktigt att den tekniska plattformen för riskbedömningsmodellen utformas så att överlämnanden mellan modellutvecklarna och modellintegrerarna blir smidiga och förutsägbara. Vidare måste riskbedömningsmodellen och ärendehanteringssystemet vara tekniskt åtskilda för att klara av tidsramarna för uppdateringar och förvaltning (se avsnitt 5.4.4).

5.4.3 Tekniska krav på plattformen för interoperabilitet

Utifrån de praktiska begränsningar vi har diskuterat kan de tekniska kraven på plattformen formuleras. När tandläkaren trycker på en knapp i journalsystemet skickas en fråga om patientens riskgrupp till Försäkringskassans system, som i sin tur skickar avidentifierade data till riskbedömningsmodellen. Modellen svarar snabbt med den förväntade riskgruppen, och denna information vidarebefordras till journalsystemet.

För att ärendehanteringssystemet ska kunna kommunicera med riskbedömningsmodellen föreslår vi att modellen publicerar ett väl definierat applikationsgränssnitt (API) som ärendehanteringssystemet kan använda. Ett REST-API över HTTPS är den dominerande standarden för moderna API:er och är därför ett bra val.

5.4.4 Uppdateringsprocess för riskbedömningsmodellen

Nedan följer ett exempel på hur uppdateringsschemat för modellen kan vara utformat:

- I maj beslutar TLV om nya regler i patienternas skydd mot höga tandvårdskostnader. De som förvaltar riskbedömningsmodellen ser över hur regeländringarna påverkar modellen, och om detaljerna i kommunikationsprotokollet mellan ärendehanteringssystemet och riskbedömningsmodellens REST-API behöver korrigeras.
- Mellan maj och september tar förvaltarna fram en preliminär uppdaterad version av koden och paketerar den som en docker-applikation (ett modernt sätt att underlätta distributionen av serverprogramvara). De uppdaterar också testkoden för att säkerställa att modellen fungerar med Försäkringskassans system.
- I november och december testar Försäkringskassan den nya koden i sin staging-miljö. Eventuella fel korrigeras av modellutvecklarna på begäran från Försäkringskassan.
- Den 15 januari sätts den nya koden i produktion samtidigt som de nya reglerna i tandvårdsstödet träder i kraft.

5.5 Kontinuerlig anpassning efter det nya tandvårdsstödet

I det här avsnittet går vi igenom hur vi ska hantera det faktum att riskbedömningsmodellen bygger på vilken tandvård patienten fått samtidigt som vilken tandvård patienten får kommer påverkas av riskbedömningsmodellen.

Att låta en maskininlärningsmodell avgöra mängden tandvårdsstöd har många fördelar jämfört med att låta behandlaren eller patienten bestämma. Men det finns en viktig nackdel som kan påverka modellens träffsäkerhet på sikt. Om riskbedömningsmodellen påverkar hur mycket tandvårdsstöd en patient får, kan det leda till att de som får mer stöd också efterfrågar mer tandvård.

Om patienter som plötsligt hamnar i en högre riskgrupp får fler åtgärder debiterade, kan det gamla sambandet mellan patientens risk och mängden vård inte längre vara giltigt. Vissa patienter kan få mer vård än tidigare, vilket kan få den gamla modellen att tro att dessa patienter är sjukare än de faktiskt är. Dessutom kan de nya åtgärderna förbättra patientens framtida tandhälsa, vilket innebär att skillnaden mellan patientens faktiska tandhälsa och den förutsagda inte längre är densamma som när modellen skapades. Med andra ord försämras träffsäkerheten i modellen när stödet införs.

Om det nya tandvårdsstödet påverkar modellens träffsäkerhet måste vi justera för ökad tandvårdskonsumtion för att modellen ska vara hållbar på lång sikt. Vi behöver uppskatta hur mycket vård patienten skulle ha fått utan det nya tandvårdsstödet och hur patientens tandhälsa skulle ha varit utan detta stöd.

För att mäta den kausala effekten av sjukvårdsinsatser ses experiment, där patienter slumpmässigt tilldelas olika nivåer av tandvårdsstöd som de mest tillförlitliga. Tyvärr är sådana experiment extremt sällsynta och det är osannolikt att ett sådant experiment kommer genomföras i Sverige för att ta reda på vilken nytta tandvårdsstödet har på folkhälsan. Rena observationsstudier kommer heller inte kunna svara på frågan vilken effekt tandvårdsstödet har. Men man kan ändå uppskatta den kausala effekten eftersom reglerna i tandvårdsstödet kommer skapa ett så kallat "naturligt experiment" som går att analysera med en "diskontinuitetsanalys".

Naturliga experiment är vetenskapliga analyser av situationer där vilken behandling en patient fick styrdes av faktorer utan behandlaren och patientens kontroll. Sådana "kvasiexperimentella" analyser är populära inom vetenskapen för de fall där kontrollerade experiment inte är möjliga.

Diskontinuitetsanalys är en statistisk metod som används för att utvärdera effekten av en insats när deltagandet bestäms av ett gränsvärde. Om individer får en behandling bara om de ligger över ett visst värde på en variabel, kan man jämföra dem som precis ligger över gränsen med dem som precis ligger under. Eftersom

dessa individer är väldigt lika förutom behandlingen, kan skillnader i resultat tillskrivas själva insatsen.

Om vårt förslag på en riskbedömning baserat på en maskininlärningsmodell godkänns, kommer villkoren för diskontinuitetsanalysen vara uppfyllda. Låt oss anta att det införs tre riskgrupper: grön, gul och röd. Gränserna för den gula och den röda gruppen kommer bestämmas av patientens förväntade framtida risk. Om gränsen för att hamna i den röda gruppen är 70 %, kan vi jämföra patienter som ligger mellan 65 och 70 % med de som ligger mellan 70 och 75 %. Dessa patienter kommer vara i stort sett lika, med undantaget att en grupp får mer tandvårdsstöd.

Metoden för diskontinuitetsanalys introducerades av Donald Thistlethwaite och Donald Campbell år 1960.²¹ De använde den för att studera effekten av meritbaserade stipendier på studenternas framtida prestationer. Sedan dess har metoden använts brett i forskningen för att analysera situationer där andra experiment inte är möjliga.

²¹ Thistlethwaite, D.L., Campbell, D.T., Regression-discontinuity analysis: An alternative to the ex post facto experiment. *Journal of Educational psychology*, 1960

Projektorganisation

Projektgruppen har bestått av:

Elena Borsci, projektledare och odontologisk utredare

Peter Karlsson, senior odontologisk utredare

Jonas Samuelsson, analytiker

Mikael Moutakis, senior analytiker

Lars Sjödin, odontologisk sakkunnig

Carl Björvang, analytiker

Achilleas Kitsoulis, odontologisk utredare

Catherine Bäckvall, jurist

Gun-Britt Lundin, enhetschef

Johan Salomonsson, analytiker

Christoffer Karlsson, analytiker

Experter:

Marie Sand, tandhygienist, Sveriges Tandhygienistförening

Gunnar Johanssen, docent, specialisttandläkare i parodontologi, Institutionen för odontologi, Karolinska Institutet

Mattias Michelin, specialisttandläkare i parodontologi

Michael Ingre, statistiker, Institutionen för medicin, Karolinska Institutet

Saikat Chatterjee, docent inom teknisk informationsvetenskap, skolan för elektroteknik och datavetenskap, Kungliga tekniska högskolan (KTH)

Referenser

Ástvaldsdóttir, Álfheiður m fl, ÅRSRAPPORT 2023 *Svenskt Kvalitetsregister för Karies och Parodontit*, SKaPa, s. 44.

Esposito, M., Kalinowski, J., Mikhaeil, M., The Effects of Recreational and Pharmaceutical Substance Use on Oral Microbiomes and Health. *Bacteria*, no. 3, 2024. <https://doi.org/10.3390/bacteria3030015>.

Feng, C. X., A comparison of zero-inflated and hurdle models for modeling zero-inflated count data. *Journal of Statistical Distributions and Applications*, Volume 8, Article number: 8, 2021. <https://doi.org/10.1186/s40488-021-00121-4>.

Förvaltningslag (2017:900) 32 §.

Inspektionen för socialförsäkringen, *remissvar SOU 2024:70 Tiotandvård - ett förstärkt högkostnadsskydd för tandvård*, S2024/01892, s. 2 och 7.

Prop. 2007/08:49 *Statligt tandvårdsstöd*, s. 100-101.

Socialdepartementet, *Uppdrag att utveckla en nationell modell för riskbedömning inom tandvården*, S2024/01260 (delvis).

Socialstyrelsen, *Förutsättningar för en nationell modell för riskbedömning inom tandvården*, 2024, s. 72.

Socialstyrelsen, *Statistik om tandhälsa*, 2023.

Socialstyrelsen, *Tandhälsoregistret*, 2024, <https://www.socialstyrelsen.se/statistik-och-data/register/tandhalsoregistret/> [hämtad 2025-01-23].

SOU 2024:70 *Tiotandvård – ett förstärkt högkostnadsskydd för tandvård*, 2024, s. 25.

SOU 2025:12 *AI-kommissionens Färdplan för Sverige*, 2024.

Statens beredning för medicinsk och social utvärdering (SBU), SBU. *Prediktionsmodeller för karies och parodontit*, 2024.

Thistlethwaite, D.L., Campbell, D.T., Regression-discontinuity analysis: An alternative to the ex post facto experiment. *Journal of Educational psychology*, 1960.

TLV, *Effekten av 2008 års tandvårdsreform på vårdgivarpriser*, dnr 04409/2014, 2015.

TLV, *Empirical studies on reference pricing*, dnr 04076/2024, 2024.

TLV, *Examining Heterogeneity in Healthcare Provider Responses to Reference Pricing Through a Synthetic Control Study*, dnr 4763/2024, 2024.

TLV, *Riskbedömning i tandvården – En nationell modell baserad på historisk tandvårdskonsumtion*, dnr 00763/2024, 2024.

Wang, R., Fu, G., Fu, B. et al., *Deep & Cross Network for Ad Click Predictions*, 2017, <https://arxiv.org/abs/1708.05123>.

Bilaga 1. Termer och begrepp

Adam-algoritmen – En optimeringsalgoritm för djupinlärning som justerar inlärningshastigheten under träningen för att förbättra hur snabbt modellen går mot en träffsäker lösning.

Aktiveringsfunktion – En aktiveringsfunktion i djupinlärning är en matematisk funktion som appliceras på en neurons summerade indata för att avgöra vilken signal (utdata) som ska skickas vidare, och dess egenskaper gör det möjligt för neurala nätverk att lära sig komplexa mönster i data. Se ”*neurala nätverk*”.

Aktiveringsfunktionen ReLU – En aktiveringsfunktion som används i neurala nätverk, känd för sin enkelhet och effektivitet, där output är noll för negativa input och linjär för positiva input.

Alternativa utfallsvariabler – Olika mått eller händelser som används för att mäta resultatet av en behandling eller intervention, exempelvis tandförlust eller behandling av parodontit.

Artificiell intelligens (AI) – Teknik som gör det möjligt för datorer att utföra uppgifter som normalt kräver mänsklig intelligens, såsom inlärning, problemlösning och beslutsfattande.

Automatiserad riskbedömning – En process där en modell automatiskt bedömer en patients risknivå baserat på data, vilket kan minska behovet av manuella bedömningar.

Automatiserat beslutsstöd – System som använder teknik, ofta AI, för att hjälpa vårdgivare att fatta beslut baserat på analyser av stora datamängder.

Binär korsentropi (Binary Cross-Entropy) – En förlustfunktion som används för att träna neurala nätverk i klassificeringsproblem där målet är att avgöra mellan två möjliga utfall.

Binära utfallsvariabler – Variabler som har två möjliga utfall, exempelvis ja/nej eller 0/1, och används för att klassificera data i två kategorier.

Convolutad neural network – Ett convolutional neural network (CNN) är en speciell typ av neuralt nätverk som främst används för att analysera visuella data genom att lära sig lokala mönster och egenskaper.

Diskontinuitetsanalys – En statistisk metod för att utvärdera effekten av en insats när deltagandet bestäms av ett gränsvärde, genom att jämföra individer nära det gränsvärde.

Djupinlärning (Deep Learning) – En gren inom artificiell intelligens som använder neurala nätverk med flera lager för att analysera komplexa mönster i stora datamängder.

Docker – Docker är en plattform för att paketera applikationer med deras beroenden, vilket gör det enklare att distribuera och köra dem konsistent i olika miljöer.

Förlustfunktion (Loss Function) – En matematisk funktion som används i maskininlärning och optimeringsproblem för att mäta hur fel en modells förutsägelse är jämfört med de faktiska utfallen. Modellen tränas för att minimera värdet på förlustfunktionen, vilket innebär att dess förutsägelser blir mer korrekta.

Försäkringskassan – Den svenska myndighet som administrerar och betalar ut det statliga tandvårdsstödet.

Förvaltningslagen – Svensk lag som reglerar hur myndighetsbeslut ska fattas och förklaras, inklusive krav på att förklara avgörande omständigheter för beslut.

HTTPS – Ett protokoll för att skicka data över internet. Det används för säkra anslutningar, till exempel vid inloggningar och betalningar online.

Hyperparametrar – Inställningar i en maskininlärningsmodell som bestämmer dess prestanda. Exempel inkluderar antalet dolda lager och inlärningstakt.

Högkostnadsskydd – Ett system som begränsar de kostnader en patient betalar för tandvård under ett år, för att ge ekonomiskt skydd mot höga vårdkostnader.

ICD-10 – Den internationella statistiska klassifikationen av sjukdomar och relaterade hälsoproblem, version 10, som används för att koda diagnoser i vården.

Inlärningstakt – En hyperparameter som bestämmer steglängden vid uppdatering av vikterna i en modell under träningen, vilket påverkar hur snabbt eller långsamt en modell lär sig.

Interoperabilitet – Förmågan hos olika IT-system och databaser att kommunicera och samarbeta effektivt, vilket är avgörande för att automatiserade riskbedömningsmodeller ska kunna användas i vården.

Kalibreringskurva – En grafisk representation som visar hur väl en modellens förutsägelser överensstämmer med faktiska utfall.

Kausal effekt – En kausal effekt innebär att en förändring i en faktor direkt orsakar en förändring i en annan faktor. Det handlar om att visa ett orsakssamband snarare än bara ett statistiskt samband.

Keras – Ett open-source bibliotek för djupinlärning som används för att bygga och träna neurala nätverk, känt för sin användarvänlighet och flexibilitet.

Klassificeringsproblem (Classification Problem) – En typ av maskininlärningsproblem där målet är att tilldela data till en eller flera fördefinierade kategorier (klasser).

L2-straff (L2 Regularization) – En metod som används för att minska överanpassning genom att lägga till en straffterm i modellens kostnadsfunktion.

Logistisk regression – En statistisk metod för att modellera och analysera data där utfallet är binärt, exempelvis förekomst eller frånvaro av en sjukdom.

Marginaleffekter – Ett mått på hur mycket en förändring i en specifik variabel påverkar utfallet i en statistisk modell, medan andra variabler hålls konstanta.

Maskininläring (Machine Learning) – En gren av artificiell intelligens där datorprogram kan identifiera mönster och fatta beslut baserat på stora datamängder utan att vara explicit programmerade.

Munhälsoprofil – En bedömning av en patients munhälsa som tar hänsyn till både aktuella och tidigare tandvårdsbehov för att förutsäga framtida risker för munsjukdomar.

Naturliga experiment – Naturliga experiment uppstår när händelser i verkliga livet slumpmässigt eller oavsiktligt skapar förutsättningar som efterliknar ett kontrollerat experiment. Forskare kan då observera effekterna utan att själva ingripa.

Neurala nätverk – Neurala nätverk är en beräkningsmodell inspirerad av den mänskliga hjärnan, där många enkla, sammankopplade enheter (neuroner) samarbetar för att identifiera och lära sig mönster i data.

Odontologiskt perspektiv – Ett synsätt inom tandvården som fokuserar på tändernas och munhälsans tillstånd och behandlingar.

Parodontal behandling – Tandvårdsbehandling som syftar till att stoppa eller bromsa utvecklingen av tandlossningssjukdomar (parodontit).

Parodontit – En allvarlig tandköttssjukdom som orsakar inflammation i vävnaderna kring tänderna och kan leda till ben- och tandförlust.

Peri-implantit – En allvarlig tandköttssjukdom som påverkar vävnader runt ett tandimplantat och kan leda till ben- och implantatförlust.

PR-kurva (Precision-Recall Curve) – En grafisk representation som används för att utvärdera prestandan hos en klassificeringsmodell, särskilt när klasserna är obalanserade, genom att visa förhållandet mellan precision och recall.

Prediktion av kariessjukdom – Förutsägelse av sannolikheten för att en patient kommer att utveckla karies baserat på tidigare data och riskfaktorer.

Prediktionsmodell – Ett statistiskt verktyg som använder information om tidigare händelser för att förutse framtida utfall. Inom tandvård används dessa modeller för att beräkna risken för framtida tandsjukdomar.

Python (programmeringsspråket) – Python är ett mångsidigt, hög nivå-språk som är känt för sin läsbarhet och stora standardbibliotek. Det används inom bland annat webbutveckling, dataanalys, maskininlärning och automatisering.

R (programmeringsspråket) – R är ett språk som främst används för statistisk analys, datahantering och visualisering. Det erbjuder många inbyggda funktioner och bibliotek för avancerad databehandling och analys.

Realiserad relativrisk – Ett mått på hur mycket större risken är i en högriskgrupp jämfört med en lågriskgrupp, baserat på faktiska data och utfall.

Referenspriser för tandvård – Det pris som anger vad ett normalpris för en tandvårdsåtgärd kan vara. Referenspriset är det pris som ersättningen i högkostnadsskyddet för tandvård beräknas utifrån.

REST-API – Ett REST-API är ett sätt att strukturera utbytet av information mellan två program över internet. Det gör det enkelt för olika system att samarbeta och dela data.

Riskbedömning – En process för att identifiera och analysera potentiella risker som kan påverka en individs hälsa, i det fall inom tandvården.

Riskbedömningsmodell – En metod eller algoritm som används för att bedöma risken för framtida tandproblem hos en patient baserat på historiska data och andra riskfaktorer.

Riskindikatorer – Faktorer eller variabler som används för att identifiera och bedöma risken för ett specifikt utfall, såsom tandförlust.

ROC-kurva (Receiver Operating Characteristic Curve) – En ROC-kurva används för att utvärdera hur väl en prediktionsmodell skiljer mellan olika utfall. Den visar sambandet mellan den sanna positiva andelen (True Positive Rate, TPR) och den falska positiva andelen (False Positive Rate, FPR) för olika gränsvärden i klassificeringen.

Sekventiella neurala nätverk – En typ av neuralt nätverk där lager av noder är ordnade i en sekvens, där varje lager tar emot input från det föregående lagret och skickar output till nästa lager.

Sigmoidfunktion – En aktiveringsfunktion som används i neurala nätverk, särskilt i utgångslager för binära klassificeringsproblem, som komprimerar input till ett värde mellan 0 och 1.

Staging-miljö – En staging-miljö är en testmiljö som ligger nära produktionsmiljön där nya funktioner och ändringar kan testas innan de släpps till riktiga användare. Den minskar risken för att felaktiga uppdateringar påverkar den faktiska driftmiljön.

Statligt tandvårdsstöd – Ekonomiskt stöd från staten för att hjälpa patienter med kostnader för tandvård, ofta med fokus på att skydda mot höga kostnader.

Straffterm (Regularization Term) – En extra komponent som läggs till i förlustfunktionen för att förhindra överanpassning (overfitting) i en maskininlärningsmodell. Strafftermer tvingar modellen att ha enklare parametrar eller vikter, vilket hjälper till att förbättra dess generalisering på nya data.

Subventionssystem – De regler och kriterier som styr hur mycket ekonomiskt stöd patienter kan få för tandvård genom det statliga tandvårdsstödet.

Tandhälsoregistret – Ett nationellt register som används för att följa utvecklingen av tandhälsa och tandvård i Sverige, samt för forskning och kvalitetsuppföljning. Administreras av Socialstyrelsen.

Tandlöshet – Ett tillstånd där en person saknar en eller flera tänder, vilket kan påverka bettfunktion och estetik.

Tandvårds- och läkemedelsförmånsverket (TLV) – En statlig myndighet som beslutar om vilka läkemedel och förbrukningsartiklar som ska ingå i läkemedelsförmånerna, samt vilken tandvård som ska omfattas av högkostnadsskyddet i det statliga tandvårdsstödet.

Tillståndsbaserad riskbedömning – En metod där en tandläkare manuellt klassificerar en patients risknivå baserat på specifika tillstånd, snarare än en automatiserad modell.

Träningsdata och testdata – Två uppsättningar data där träningsdata används för att utveckla en modell och testdata används för att utvärdera modellens prestanda. Se ”*överanpassning*”.

Utfallsmått – Det resultat eller den händelse som en prediktionsmodell försöker förutsäga. Inom tandvård kan det vara exempelvis tandförlust eller parodontit.

Åtgärds-koder – Koder som beskriver specifika behandlingar eller ingrepp som en patient har fått inom tandvården.

Överanpassning (Overfitting) – När en prediktionsmodell anpassar sig för mycket till de data som används vid träningen, vilket leder till sämre förmåga att förutsäga framtida utfall eftersom modellen fångar slumpmässiga variationer i träningsdata istället för de verkliga mönstren.

Bilaga 2: Hur vi byggde våra djupinlärningsmodeller

Utgångspunkten för denna analys var 14 olika dataset med förklaringsvariabler och tre olika resultatvariabler, men där alla dataset innehöll samma tidsperiod och information om patientens tandhälsa, ålder, kön och civilstånd. Dataseten innehöll också uppgifter om vilken vård patienten hade fått. Det som skilde sig åt var hur vården mättes: antingen i antal åtgärder inom tandvårdsstödet eller i minuter per åtgärd. Vi undersökte också om åtgärderna skulle grupperas i olika åtgärdsgrupper eller inte, och om de skulle delas in beroende på vilket tillstånd de utfördes under. Dessutom grupperades patientens hemkommun antingen efter kommun eller region.

Vi använde tre olika binära resultatvariabler, tidigare beskrivna: *parod*, *rotex* och *parodrotex*. Alla modeller var sekventiella neurala nätverk med 1-3 dolda lager. Vi byggde modellerna med hjälp av verktyget Keras. För de dolda lagren användes aktiveringsfunktionen ReLU, och för resultatlagret användes sigmoid. Optimeringsalgoritmen vi valde var Adam, och förlustfunktionen var binär korsentropi.

Eftersom vår resultatvariabel var mycket obalanserad, justerades vikten för observationerna i minoritetsgruppen i omvänd proportion till deras frekvens så att de fick större betydelse i beräkningarna.

Vi använde Keras autotuning för att hitta de bästa inställningarna, eller hyperparametrarna, för varje resultatvariabel och dataset. De hyperparametrar vi justerade var antalet dolda lager (1-3), antalet noder i det första dolda lagret (mellan 50 och 75 % av antalet inputnoder), antalet noder i det andra dolda lagret (mellan 10 och 40 % av antalet inputnoder), och antalet noder i det tredje dolda lagret (mellan 5 och 10 % av antalet inputnoder). Vi justerade också l2-straffpoängen, som varierade mellan 0,001 och 0,005, och inlärningstakten, som kunde vara 1, 2, 3 eller 4 %. Vi undersökte också om utdata från de dolda lagren skulle normaliseras eller inte.

Totalt testades 20 olika varianter av varje modell för varje dataset och resultatvariabel. Varje modell beräknades under 12 epoker. Varje variant beräknades tre gånger för att minska effekten av slumpmässiga variationer i modellberäkningarna. Sammanlagt beräknade vi 2520 modeller, vilket tog ungefär 2,5 dygn att köra på en vanlig bärbar dator. Den modell som hade det lägsta prediktionsfelet valdes ut. Prediktionsfelet beräknades på ett valideringsdataset som innehöll andra patienter och en annan tidsperiod än träningsdatasetet.

Bilaga 3: Definition av utfallsmått

Utfallsmått för patienter med icke-trivial parodontit, peri-implantit eller andra sjukdomar i vävnaderna kring tänder eller implantat:

För att klassas som ”parod”, ”peri-implantit” eller ”annat”-patient måste antingen **villkor A**, eller **villkor B** eller **villkor C** nedan gälla. Om mer än ett villkor gäller klassas man fortfarande till denna grupp.

Villkor A: Ha fått minst TRE av någon av följande åtgärder under ett års tid, alla kopplade till tillstånd 3043 (parodontit), 3044 (peri-implantit) eller 3046 (andra sjukdomar):

341: Behandling av parodontal sjukdom eller peri-implantit, mindre omfattande

342: Behandling av parodontal sjukdom eller peri-implantit, omfattande

343: Behandling av parodontal sjukdom eller peri-implantit, särskilt tidskrävande

Villkor B: Ha fått minst EN av någon av följande åtgärder under ett års tid kopplat till tillstånd 3043 (parodontit):

401: Tanduttagning, en tand

402: Tanduttagning, när separation eller friläggning krävs, en tand

403: Tanduttagning, enkel

404: Kirurgiskt avlägsnande av en eller flera tänder eller annan vävnad i samma kvadrant

407: Övrig kirurgi eller plastik

409: Kirurgiskt avlägsnande av tand i annan kvadrant i samband med parodontalkirurgiska åtgärder, tilläggsåtgärd

410: Tanduttagning, ytterligare när flera tanduttagningar utförs under samma dag eller i samband med lambåoperation, tilläggsåtgärd

446: Rekonstruktiv behandling med membran (GTR) eller emaljmatrixprotein, tilläggsåtgärd

447: Rekonstruktiv behandling med benersättningsmaterial, tilläggsåtgärd

451: Parodontalkirurgi i en kvadrant eller inom tandposition 3–3

452: Parodontalkirurgi i flera kvadranter eller omfattande parodontalkirurgi i en kvadrant eller inom tandposition 3–3

Villkor C: Minst en av följande åtgärder kopplat till tillstånd 3044 (periimplantit):

429: Kirurgiskt avlägsnande av implantat

435: Avlägsnande av ett implantat

436: Avlägsnande av ett implantat, enkel

453: Kirurgisk behandling av periimplantit och i förekommande fall tänder med parodontit i en kvadrant eller inom tandposition 3–3

454: Kirurgisk behandling av periimplantit och i förekommande fall tänder med parodontit i flera kvadranter eller omfattande kirurgisk behandling i en kvadrant eller inom tandposition 3–3

Utfallsmått för patienter med tandförlust som inte beror på parodontit eller sjukdomar eller tillstånd i vävnaderna kring tänder eller implantat

För att klassas som patient med tandförlust som inte beror på parodontit måste antingen **villkor A** eller **villkor B** nedan gälla. Om mer än ett tillstånd gäller klassas man fortfarande till denna grupp.

Villkor A: Minst en av följande åtgärder i alla tandpositioner förutom de åtgärder som gjordes på tänder i position 8 eller 9:

501: Rensning och rotfyllning, en rotkanal

502: Rensning och rotfyllning, två rotkanaler

503: Rensning och rotfyllning, tre rotkanaler

504: Rensning och rotfyllning, fyra eller fler rotkanaler

Villkor B: Minst en av följande åtgärder förutom för de åtgärder som gjordes på tänder i position 8 eller 9 ELLER för åtgärder där tillståndet var 3043, 3044 eller 3046:

401: Tanduttagning, en tand

402: Tanduttagning, när separation eller friläggning krävs, en tand

403: Tanduttagning, enkel

404: Kirurgiskt avlägsnande av en eller flera tänder eller annan vävnad i samma kvadrant

405: Omfattande dentoalveolär kirurgi

407: Övrig kirurgi eller plastik

410: Tanduttagning, ytterligare när flera tanduttagningar utförs under samma dag eller i samband med lambåoperation, tilläggsåtgärd

Utfallsmåttens förekomst under uppföljningsperioden

Tabell 3 Antal och andel patienter från urvalet testdata som fick tandförlust, parodontit eller båda under uppföljningsperioden september 2023 till augusti 2024.

Utfall	Fick utfall	Fick ej utfall
Tandförlust	893 (6 %)	14 164 (94 %)
Parodontit	343 (2 %)	14 713 (98 %)
Tandförlust eller parodontit	1 164 (8 %)	13 894 (92 %)

Not: Fördelningen var likartad i träningsdatan.

Bilaga 4: Ytterligare modellutvärdering

PR-kurva

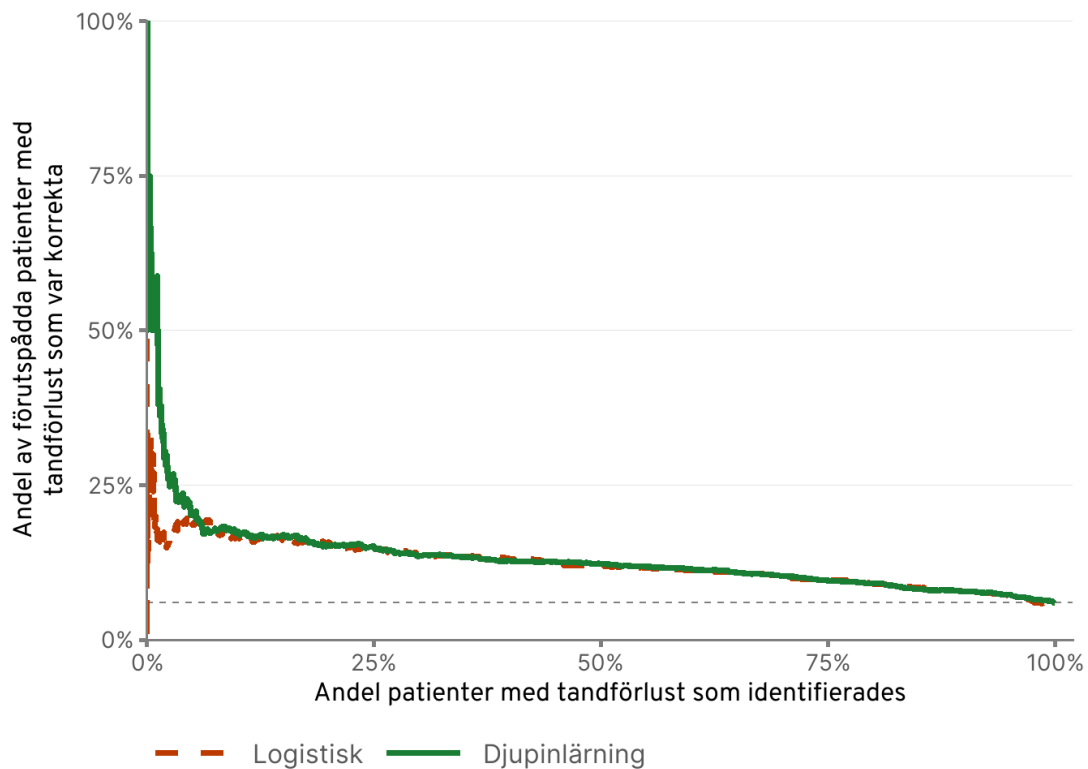
En annan metod för att utvärdera prediktionsmodeller är en PR-kurva (precision-recall). PR-kurvan ger oss särskilt värdefull information eftersom den är mer känslig för hur väl modellen presterar just för högriskgruppen. Den hjälper oss att förstå hur effektiv modellen är på att både hitta de patienter som verkligen är i riskzonen och att undvika att felaktigt klassificera patienter som högrisk när de inte är det. Genom att komplettera utvärderingen med denna metod får vi en mer heltäckande bild av modellens prestanda och kan säkerställa att den är tillräckligt träffsäker.

När vi tittar på en punkt på PR-kurvan, till exempel (X %, Y %), betyder det att modellen lyckas identifiera X procent av alla patienter som riskerar att förlora tänder eller utveckla parodontit. Samtidigt visar Y procent hur många av de patienter som modellen har klassificerat som högrisk som faktiskt förlorade en tand eller fick parodontit.

Den horisontella streckade linjen visar genomsnittsriskerna bland alla patienter i populationen. En perfekt modell skulle nå grafens övre högra hörn, vilket skulle innebära att vi både identifierar alla högriskpatienter och att alla vi klassar som högrisk verkligen är det.

I PR-kurvan presenteras inte tiotandvårdsmodellen då en modell med endast två sätt att beskriva patienternas bakgrund (över och under 67 år) inte kan generera en meningsfull PR-kurva eftersom den endast ger två möjliga sannolikhetsvärden. För att kunna rita en PR-kurva behöver modellen kunna rangordna patienter kontinuerligt efter risk, vilket kräver fler riskindikatorer.

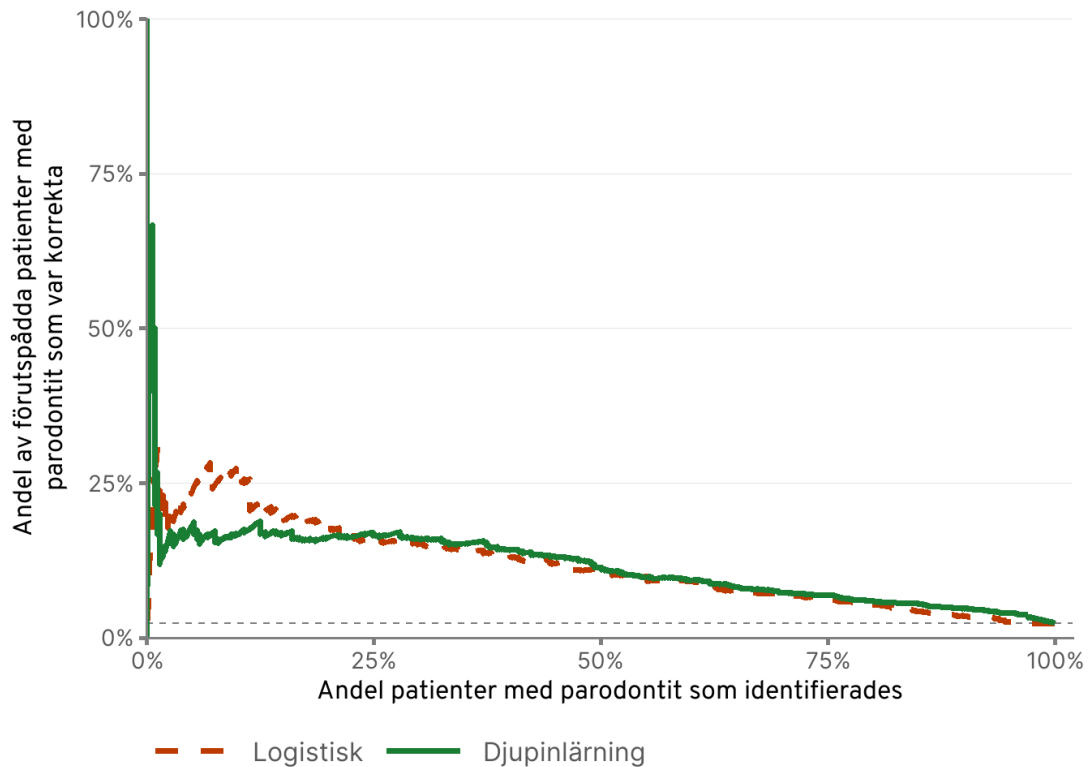
Figur 6 PR-kurva för tandförlust.



Not: Figuren visar förhållandet mellan modellens precision (andel korrekta identifieringar bland alla predikterade fall) och recall (andel identifierade fall av alla verkliga fall). En perfekt modell skulle ha en kurva som går genom övre högra hörnet, medan en modell som gissar slumpmässigt skulle följa den streckade linjen som visar andelen tandförluster i befolkningen.

PR-kurvan för tandförlust visar att både den logistiska modellen och djupinlärningsmodellen presterar likartat. När modellerna identifierar cirka 25 procent av patienterna som kommer att förlora tänder, är ungefär 20 procent av förutsägelseorna om tandförlust korrekta. Den streckade linjen visar den genomsnittliga risken i populationen, omkring sex procent.

Figur 7 PR-kurva för parodontit.



Not: Figuren visar förhållandet mellan modellens precision (andel korrekta identifieringar bland alla predikterade fall) och recall (andel identifierade fall av alla verkliga fall). En perfekt modell skulle ha en kurva som går genom övre högra hörnet, medan en modell som gissar slumpmässigt skulle följa den streckade linjen som visar andelen tandförluster i befolkningen.

PR-kurvan för parodontit visar också att den logistiska modellen och djupinlärningsmodellen presterar likartat. Den logistiska modellen visar något bättre precision för de patienter som har högst risk, vilket syns i kurvans början där den röda linjen ligger över den gröna. Den streckade linjen visar den genomsnittliga risken i populationen, omkring två procent.