

Markus Leskinen ja Sture Andersson

Tekoäly vastasyntyneiden tehohoidossa

Sähköisessä muodossa oleva lääketieteellinen tieto lisääntyy jatkuvasti, ja erityisesti tehohoidossa olevien potilaiden hoidosta pystytään integroitujen mittalaitteiden ansiosta muodostamaan potilastietojärjestelmiin yhä kattavampi kuva. Yhdistämällä nämä tiedot terveydenhuollon rekistereihin ja hyödyntämällä viime vuosikymmenellä huomattavasti kehittyneitä suurten tietomassojen analysointia on mahdollista kehittää malleja, jotka pystyvät koneoppimisen avulla ennustamaan komplikaatioita ja arvioimaan potilaan ennustetta. Tällaiset mallit pystyvät analysoimaan valtavaa määrää muuttujia ja niiden ajallisia yhteyksiä. Niiden läpinäkyvyys on kuitenkin usein heikko, ja klinikon on vaikea seurata, miten lopputulokseen on päästy. Kun algoritmeja tarjotaan yhä enemmän ratkaisuksi klinisiin ongelmiin, lääkärin tulee ymmärtää mallien rajoitukset ja toisaalta niiden tarjoamat mahdollisuudet.

Tekoälyllä on tarkoitettu tietojärjestelmiä, jotka kykenevät jäljittelemään ihmisten kykyä ratkaista ongelmia päättelemällä ja soveltaa aikaisemmin hankittua tietoa uusiin olosuhteisiin. Tällaisesta aidosta tekoälystä ollaan vielä kaukana, mutta suppeammilla alueilla tekoälyn sovellukset, erityisesti koneoppiminen, edistyivät 2010-luvulla huomattavasti, ja niitä alettiin soveltaa myös lääketieteessä (1).

Ohjatussa koneoppimisessa tietojärjestelmälle tarjotaan mahdollisimman yksityiskohtaista dataa suuresta määrästä tapauksia, joiden lopputulos tunnetaan. Tästä opetusdatasta algoritmi kehittää itse sääntöjä etsimällä, mitkä tekijät erottavat eri lopputuloksiin päätyviä potilaita tai tilanteita toisistaan. Algoritmille ei siis kerrota etukäteen, millä perusteella jako pitää tehdä, mutta sille kerrotaan lopputulos. Ohjaamattomassa oppimisessa opetusdatasta ei kerrota algoritmille etukäteen mitään.

Kun algoritmi on saatu opetettua ennustamaan lopputulos riittävällä varmuudella, sen toimivuus testataan testidatalla, jota ei ole käytetty opetuksessa mutta jonka lopputulos tunnetaan. Koneoppimisessa voidaan hyödyntää valtavaa määrää muuttujia ja näiden ajallista riippuvuutta toisistaan sekä luoda malleja, joi-

ta olisi ollut mahdotonta kehittää totunnaisilla menetelmillä. Mallien monimutkaisuus tekee kuitenkin niiden läpinäkyvästä arvioinnista vaikeaa, ja klinikon voi olla mahdotonta seurata mallin päättelyä sekä ymmärtää, miten saatuun diagnoosiehdotukseen on päädytty. Tätä on kutsuttu mustan laatikon ongelmaksi (2).

Tekoälyä ja koneoppimista tarjotaan yhä enemmän ratkaisuksi vaihteleviin lääketieteellisiin ongelmiin, mutta käytännössä toimivia sovelluksia on vielä rajallisesti (3). Ollaan siis Gartnerin hypekäyrän nousevalla osalla, jossa odotuksia on enemmän kuin onnistuneita toteutuksia (4). Uudet teknologiat noudattavat usein Roy Amaran nimiin laitettua kehityskulkua, jossa niiden vaikutus lyhyellä aikavälillä yliarvioidaan, mutta teknologian vakiintuessa pitkän aikavälin vaikutus aliarvioidaan. Onkin uskottavaa, että tekoälyn merkitys lääketieteessä lisääntyy tällä vuosikymmenellä merkittävästi.

Tehohoito lääketieteellisen datan tuottajana

Tehohoidossa joudutaan tekemään monista eri lähteistä kerätyn tiedon perusteella reaaliaikai-

sia päätöksiä, joilla voi olla potilaan kannalta vakavia ja pitkäkestoisia vaikutuksia. Tarve kerätä tämä data, esittää se ymmärrettävässä muodossa ja arkistoida myöhempään käyttöön johti integroitujen potilastietojärjestelmien varhaiseen käyttöönottoon teho-osastoilla ja vastasyntyneiden teho-osastoilla (5).

Potilasmonitorit ja muut elintoimintoja seuraavat mittalaitteet pystytään yhä useammin kytkemään suoraan potilastietojärjestelmään, johon saadaan laiteliitäntöjen kautta dataa myös monista hoitolaitteista, kuten hengityskoneista. Näiden automaattisesti kerättävien tietojen lisäksi hoitohenkilökunta kirjaa käsin järjestelmään muiden mittalaitteiden tuottaman tiedon, hoitolaitteiden asetukset, annetut lääkkeet ja ravitsemuksen sekä omat havaintonsa. Lääkärit määräävät järjestelmään lääkkeet ja laboratoriotulokset sekä kirjaavat tehdyt toimenpiteet, sairauskertomuksen ja diagnoosit.

Laboratorio- ja kuvantamistutkimuksien tulokset kerätään usein omiin tietojärjestelmiinsä, mutta niistäkin yhä suurempi osa tuodaan integroituun järjestelmään. Näistä teho-osaston tietojärjestelmään tallennetuista tiedoista on vuosien aikana kertynyt yksityiskohtainen tietokanta laajan potilasjoukon elintoiminnoista, annetuista hoidoista ja osin hoitotuloksistakin.

Kun nämä tiedot yhdistetään terveydenhuollon rekistereistä saataviin tietoihin potilaan myöhemmistä vaiheista, erityisesti diagnoosietoihin ja eloon jäämiseen, syntyy koneoppimiseen sopiva aineisto, jossa on runsaasti biosignaaliainetta ja vahvat selitettävät eli lopputulosuuttajat.

Koneoppiminen vastasyntyneiden teho-osastolla

Datan keräämisen kannalta vastasyntyneiden tehohoito ei oleellisesti poikkea muusta tehohoidosta, ja käytössä on yleensä vastaavat laiteintegraatiot. Koneoppimisen kannalta pikkukeskoset ovat kuitenkin monella tavalla ideaalinen potilasryhmä. Sekoittavia tekijöitä on vähemmän kuin aikuisten hoidossa, ja hoito on voimakkaammin keskitettyä.

Valtaosa pikkukeskosista syntyy Suomessa yliopistosairaaloissa (6). Heitä hoidetaan vasta-

syntyneiden teho-osastoilla, mikä mahdollistaa tiedon kattavan keräämisen. Hoito alkaa syntymästä, ja raskaudenaikaiset tiedot pyritään keräämään järjestelmällisesti. Pikkukeskosien hoitajakset ovat aikuisten tehohoitoon verrattuna pitkiä, joten dataakin kertyy vastaavasti enemmän.

Hoidon kehittämisestä huolimatta pikkukeskosilla esiintyy merkittävää sairastavuutta. Tyypillisiä komplikaatioita ovat sepsis, aivokammionsisäiset verenvuodot, nekrotisoiva enterokoliitti, bronkopulmonaalinen dysplasia, retinopatia ja menehtyminen. Koneoppimiselle on siis selkeitä päätetapahtumia. Vastasyntyneiden teho-osastolla pikkukeskoset ovat pitkien hoitajaksojensa takia kuormituksen kannalta merkittävä potilasryhmä, ja heidän hoitonsa optimointi on siksi tärkeää.

Tehohoitojakson jälkeen pikkukeskoset siirtyvät jatkohoitoon, jossa dataa ei enää kerätä samalla tarkkuudella. Terveiden ja hyvinvoinnin laitos kerää kuitenkin pikkukeskosten merkittävät komplikaatiot valtakunnallisen syntymärekisterin yhteyteen pienten keskosten rekisteriin, josta on mahdollista saada varsin hyvä käsitys tehohoidon jälkeisistä vaiheista.

Koneoppimista on käytetty vastasyntyneiden teho-osastolla esimerkiksi nekrotisoivaan enterokoliittiin liittyvän suolen puhkeaman tunnistamiseen, hoitoaikojen ennustamiseen, kouristusten tunnistamiseen EEG-datasta, sepsiksen tunnistamiseen ja bronkopulmonaalisen dysplasian ennustamiseen (7–15). Olemme itse käyttäneet alle 1 500 g:n painoisia pikkukeskosia mallina tekoälyn hyödyntämisessä.

Pienet keskoset – iso data -hanke

HUS:n vastasyntyneiden teho-osasto otti integroidun Centricity for Critical Care (Clinisoft) -tietojärjestelmän käyttöön vuonna 1999, ja vuosien mittaan järjestelmään on kerätty yksityiskohtaiset tiedot yli 2 500 pikkukeskosten tehohoidosta. Yhdessä Aalto-yliopiston tietotekniikan laitoksen tutkijoiden kanssa olemme tutkineet vuosina 1999–2013 hoidettujen 2 091 pikkukeskosten tehohoitojakson aikana kerättyä tietoa Pienet keskoset – iso data



KUVA. Vastasyntyneiden tehohoito on hyvin laite- ja datakeskeistä. Tärkeimmät datalähteet ovat potilasmonitori ja laiteintegraatiolla varustettu hengityskone.

-massadatahankkeessa. Tarkoituksena on ollut analysoida ensimmäisten viikkojen aikaisten hoitoratkaisujen vaikutusta kehittyviin komplikaatioihin ja myöhempään terveyteen. Datan laajuutta kuvaa se, että potilaista on 31 miljoonaa happikyllästeisyyskirjausta ja 17 miljoonaa verenpaineikirjausta. Tietoja on analysoitu tekniikan tohtori Olli-Pekka Rinta-Kosken väitöskirjassa (16).

Alkuvaiheessa analysoimme dataa kuvailevasti ja tutkimme happikyllästeisyyden vaihtelua virka- ja päivystysaikana sekä käsin ja koneella tehtyjen kirjausten eroja (17). Tuloksissa ilmeni selvä ero samoista happikyllästeisyysmittareista automaattisesti kerätyn datan ja hoitajien kirjaaman datan välillä. Hoitajien kirjauksissa poikkeavan pienet arvot olivat yliedustettuina ja happikyllästeisyyttä oli selvästi pyöristetty lähimpään tasakymmeneen. Normaaliaralueella olevia arvoja taas oli kirjattu selvästi harvemmin kuin koneellisissa kirjauksissa.

Tämä korostaa automaattisesti kerättävän datan merkitystä. Olemme myös osoittaneet, että potilasmonitorilta kerätyn aikasarjadatan pohjalta luodun koneoppimismallin avulla pikkukeskosten menehtymistä tehohoidon aikana voidaan ennustaa vastaavalla tarkkuudella kuin tavanomaisilla SNAPPE-II- ja CRIB-II-riskipistemalleilla ja että näiden yhdistelmällä saadaan parempia tuloksia kuin kummallakaan yksin (18). Yhdistelmämallin tarkkuus oli 0,95, herkkyys 0,68 ja ennustusvoimaa kuvaava AUC-luku 0,94. Tulos on hyvä, mutta sen ja muidenkin tehohoitojakson aikaista kuolleisuutta ennustavien mallien ongelmana on, että usein menehtyminen johtuu tehohoidosta luopumisesta, jolloin voidaan perustellusti kysyä, ennustaako malli enemmän lääkärin päätöksentekoa vai taudin kulkua.

Kolmen ensimmäisen elinvuorokauden aikana kerätty biosignaali data ennustaa myös vähäisemmässä määrin muita tehohoitojakson aikai-

Ydinasiat

- ▶ Valtaosa kaikesta potilaasta kerättävästä terveystiedosta on sähköisessä muodossa, ja sitä pystytään analysoimaan koneoppimista käyttämällä.
- ▶ Koneoppimiseen perustuvien diagnosti- ja ennustemallien tarkkuus paranee jatkuvasti ja lähestyy kliinisesti hyödyllistä tasoa, mutta niiden läpinäkyvyys on heikko.
- ▶ Potilaiden aikaisempaan hoitoon perustuvat mallit voivat sisältää systemaattisia ongelmia, jotka voivat johtaa joidenkin potilasryhmien epäoptimaaliseen hoitoon.
- ▶ Integroitujen potilasjärjestelmien laajentuminen teho-osastojen ulkopuolelle tuo aivan uudenlaisia mahdollisuuksia tiedon keräämiseen sairaaloissa ja perusterveydenhuollossa.

sia komplikaatioita, kuten bronkopulmonaalista dysplasiaa (19). Tämän mallin tarkkuus oli 0,93 mutta herkkyys vain 0,52, AUC-luku oli 0,87. Onkin oletettavaa, että bronkopulmonaalisen dysplasian ennustaminen vaatii pidemmän seuranta-ajan eivätkä kolme ensimmäistä elinvuorokautta vielä ratkaise taudin kehittymistä.

Sepsiksen ennustaminen CHAID-päätöspuumallilla

Samaa tietokantaa hyödynnettiin myös HUS:n ensimmäisessä IBM:n Watson-projektin kanssa kehityksessä algoritmissa, joka tunnistaa pikku-keskosten myöhäisen sepsiksen. Vertasimme veriviljelypositiivisten sepsikseen sairastuneiden lasten biosignaalidataa 24–36 tuntia ennen veriviljelyn ottoa ja opetimme CHAID-päätöspuumallin (chi-square automatic interaction detection) erottamaan sepsispotilaat verrokeista.

Erillisessä testiaineistossa malli tunnisti myöhäisen sepsiksen 24 tuntia ennen veriviljelyn ottoa 90 %:n herkkyydellä ja 89 %:n tark-

kuudella. AUC-luku oli 0,93. Mallissa merkittäviä sepsistä ennustavia tekijöitä olivat pieni happikyllästeisyys, matala happiosapaine sekä pieni happikyllästeisyyden ja sisäänhengityskaasun happiosuuden (FiO₂) suhde, mikä sopii kliiniseen kokemukseen.

Verenpainemuutoksilla ei ainakaan tässä aikaikkunassa ollut merkittävää ennustearvoa. Laboratoriokokeiden perusteella asidoosi, leukopenia ja hyponatremia nousivat mallissa esiin. CRP-pitoisuuden rooli ei ollut merkittävä, mikä voi kuitenkin selittyä osaston kliinisellä käytännöllä, jossa CRP-pitoisuuden suureminen johtaa usein veriviljelyn ottamiseen. Malli oli rajattu 24 tuntia ennen veriviljelyn ottamista kerättyyn dataan.

Algoritmi on tarkoitus validoida riippumattomassa vuosien 2014–2017 sepsiskohortissa. Samalla selvitetään, voiko nykyisen algoritmin korvata arkkitehtuuri- ja lisenssiratkaisujen suhteen kevyemmällä mallilla. Jos menetelmän toimivuus varmistuu, sitä pyritään kokeilemaan kliinisesti ottamalla se tausta-ajoon Apotissa, HUS:n uudessa potilastietojärjestelmässä.

Algoritmi kehitettiin keskiarvoistetun monitoridatan pohjalta, eikä siinä voitu hyödyntää sydämen lyöntitiheyden vaihtelusta kertovaa EKG-raakadataa, jota on menestyksellä käytetty vastasyntyneiden sepsiksen tunnistamiseen (20). HUS:n tietoallas mahdollistaa jatkossa tämäntyyppisen raakadatan analyysin, mutta datan kerääminen on vasta aloitettu ja jää nähtäväksi, paljonko tämä parantaa ennusteen tarkkuutta. Sepsisalgoritmi käyttää jo nykyisin kerättävää mittaustietoa eikä siten vaadi muutoksia potilaiden monitorointiin.

Ongelmat ja riskit

Koneoppiminen mahdollistaa säännönmukaisuuksien löytämisen valtavasta muuttujajoukosta, mutta luo samalla uusia ongelmia. Algoritmit ovat usein erittäin monimutkaisia ja kehittyneemmissä monikerroksellisia hermoverkkoja hyödyntävissä malleissa ihmiselle lähes mahdollottomia seurata. Mitä monimutkaisempaa ja vähemmän läpinäkyvää algoritmia käytetään, sitä varovaisemmin sen antamiin ennusteisiin tai diagnooseihin tulee suhtautua (2).

Teho-osastojen potilastietojärjestelmät luotiin aikanaan ensisijaisesti kliiniseen käyttöön, eikä vanhemmissa järjestelmissä painotettu syntyvän tiedon tutkimuskäyttöä, mikä rajoittaa jossain määrin tiedon hyödyntämistä jälkikäteen. Kaikkea kerättyä dataa ei ole arkistoitu, vaan sitä on saatettu keskiarvoistaa erityisesti laiteliitännöistä ja potilasmonitorilta kerättävän tiedon osalta.

Tietokantoja on myös muokattu kliinisen tarpeen mukaan, vanhoja muuttujia on jätetty pois ja uusia otettu mukaan. Lisäksi muuttuvat kirjaamiskäytännöt ovat saattaneet muuttaa samaankin muuttujaan kirjattavaa tietoa. Tämä vaikeuttaa eri aikoina hoidettujen potilaiden datan vertailua. Vanhemmissa järjestelmissä arkistoitavaa tietoa ei ole järjestelmällisesti validoitu versiopäivitysten yhteydessä, ja kokemuksemme mukaan dataa on menetetty.

Tietojärjestelmään tallennettu tieto ei aina ole luotettavaakaan, vaan sisältää virhekirjauksia, joiden määrä vaihtelee syöttötavan mukaan. Koneoppimisen pohjana käytetty data voi myös olla vinoutunutta, mistä saattaa seurata harhaanjohtavia tai jotain potilasryhmää syrjiviä tuloksia (21). Keskosena syntyneet tytöt selviävät keskimäärin hieman poikia paremmin, samoin yhdysvaltalaisissa aineistoissa afroamerikkalaiset lapset (22). Jos algoritmia käytetään resurssien allokoinnin osana, missä määrin on hyväksyttävää ottaa nämä tekijät huomioon?

Koska mallit pohjautuvat historialliseen dataan, potilaiden hoitoon liittyneet päätökset voivat vaikuttaa mallin antamaan kuvaan selviämismahdollisuuksista. Eri maista aktiivihoidon eri rajojen mukaan kerätty aineistot antaisivat aivan erilaisen kuvan raskausviikolla 23 syntyvän lapsen ennusteesta. Myös synnynäisiin oireyhtymiin liittyvät käytännöt hoidon rajauksista voivat johtaa itseään toteuttaviin ennustemalleihin.

MARKUS LESKINEN, LT, lastentautien erikoislääkäri
HUS, Uusi lastensairaala

STURE ANDERSSON, neonatologian emeritusprofessori, lastentautien erikoislääkäri, tutkija
Lastenkliniikka, Helsingin yliopisto ja HUS

VASTUUTOIMITTAJA
Pekka Lahdenne

Lopuksi

Koneoppimista hyödyntävät mallit ovat riippuvaisia kerätyn datan määrästä ja laadusta. Useissa yliopistosairaaloissa on jo aloitettu järjestelmällisen tiedon kerääminen biologisten näytteiden osalta biopankkeihin sekä muun terveystiedon, potilaista kerättävän monitoridatan ja muun raakadatan osalta tietoaletsiin. Myös kerättävän tiedon laatu muuttuu, kun uudet potilastietojärjestelmät kuten Apotti keräävät yhä suuremman osan tiedosta rakenteisena, jolloin sen hyödynnettävyys koneoppimisessa on vapaata tekstiä parempaa.

Rakenteinen kirjaaminen ja tietoaletat mahdollistavat yksityiskohtaisen potilastiedon keräämisen teho-osastoja laajemmin niin erikoissairaanhoidossa kuin perusterveydenhuollosakin. Jatkossa käytössämme ovat siis selvästi nykyisiä laajemmat tietoaaineistot.

Myös kykymme hyödyntää kerättävää tietoa paranee. Koneoppimisen menetelmät kehittyvät jatkuvasti, ja niitä opitaan soveltamaan paremmin lääketieteelliseen tietoon, vaikkakin hitaammin kuin välillä on odotettu. Uusiin potilastietojärjestelmiin kuten Apottiin on myös usein rakennettu valmius hyödyntää algoritmeja ja esittää niiden tulokset lääkärille saman käyttöliittymän kautta kuin muukin potilasta koskeva tieto.

Algoritmit voivat auttaa lääkäriä huomamaan harvinaisen diagnoosin mahdollisuuden tai tunnistamaan poikkeaman potilaan hoidossa. Ne eivät kuitenkaan korvaa tavanomaista lääkärin tutkimusta, analyttistä harkintaa ja päätöksentekoa, vaan täydentävät näitä. Tekoäly toimii parhaimmillaan lääkärin apuvälineenä, tukiälynä, diagnostiikassa ja hoidon suunnittelussa, mutta lopullinen hoitovastuu säilyy lääkärillä. Ainakin toistaiseksi ihmistä hoitaa ihminen, ei kone. ■

SIDONNAISUUDET

Markus Leskinen: Korvaukset koulutus- ja kongressikuluista (Chiesi), hankkeet (salassapitosopimukset tietojärjestelmähankeista IBM:n ja GE:n kanssa)

Sture Andersson: Asiantuntijapalkkio (Chiesi)

KIRJALLISUUTTA

1. Gutierrez G. Artificial Intelligence in the Intensive Care Unit. *Crit Care* 2020;24:101–9.
2. Price WN. Big data and black-box medical algorithms. *Sci Transl Med* 2018;10:eaa05333–16.
3. Sanchez-Pinto LN, Luo Y, Churpek MM. Big data and data science in critical care. *Chest* 2018;154:1239–48.
4. Hype cycle research methodology. Gartner 2020. <https://gartner.com/en/research/methodologies/gartner-hype-cycle>.
5. Clemmer TP. Computers in the ICU: where we started and where we are now. *J Crit Care* 2004;19:201–7.
6. Rautava L, Lehtonen L, Peltola M, et al. The effect of birth in secondary- or tertiary-level hospitals in Finland on mortality in very preterm infants: a birth-register study. *PEDIATRICS* 2007;119:e257–63.
7. Irlés C, González-Pérez G, Carrera Muiños S, ym. Estimation of neonatal intestinal perforation associated with necrotizing enterocolitis by machine learning reveals new key factors. *Int J Environ Res Public Health* 2018;15:2509.
8. Temple MW, Lehmann CU, Fabbri D. Predicting discharge dates from the NICU using progress note data. *PEDIATRICS* 2015;136:e395–405.
9. Zwanenburg A, Andriessen P, Jellema RK, ym. Using trend templates in a neonatal seizure algorithm improves detection of short seizures in a foetal ovine model. *Physiol Meas* 2015;36:369–84.
10. Temko A, Thomas E, Marnane W, ym. EEG-based neonatal seizure detection with support vector machines. *Clin Neurophysiol* 2011;122:464–73.
11. Mani S, Ozdas A, Aliferis C, ym. Medical decision support using machine learning for early detection of late-onset neonatal sepsis. *J Am Med Inform Assoc* 2014;21:326–36.
12. Gur I, Markel G, Nave Y, ym. A mathematical algorithm for detection of late-onset sepsis in very-low birth weight infants: a preliminary diagnostic test evaluation. *Indian Pediatr* 2014;51:647–50.
13. Wang K, Bhandari V, Chepustanova S, ym. Which biomarkers reveal neonatal sepsis? *PLoS ONE* 2013;8:e82700.
14. Masino AJ, Harris MC, Forsyth D, ym. Machine learning models for early sepsis recognition in the neonatal intensive care unit using readily available electronic health record data. *PLoS ONE* 2019;14:e0212665.
15. Ochab M, Wajs W. Expert system supporting an early prediction of the bronchopulmonary dysplasia. *Comput Biol Med* 2016;69:236–44.
16. Rinta-Koski OP. Machine learning in neonatal intensive care. Väitöskirja. Aalto-yliopisto 2018. <http://urn.fi/URN:ISBN:978-952-60-8210-3>.
17. Rinta-Koski O-P, Hollmén J, Leskinen M, ym. Variation in oxygen saturation measurements in very low birth weight infants. Kirjassa: Makedon F, toim. Kofu: Pervasive Technologies Related to Assisitive Environments (PETRA) 2015, s.29.
18. Rinta-Koski O-P, Särkkä S, Hollmén J, ym. Gaussian process classification for prediction of in-hospital mortality among preterm infants. *Neurocomputing* 2018;298:134–41.
19. Rinta-Koski O-P, Särkkä S, Hollmén J, ym. Prediction of major complications affecting very low birth weight infants. 2017 IEEE Life Sciences Conference (LSC), 13. -15.12.2017 Sydney, Australia. DOI:10.1109/LSC.2017.8268174.
20. Moorman JR, Lake DE, Griffin MP. Heart rate characteristics monitoring for neonatal sepsis. *IEEE Trans Biomed Eng* 2006;53:126–32.
21. Char DS, Shah NH, Magnus D. Implementing machine learning in health care - addressing ethical challenges. *N Engl J Med* 2018;378:981–3.
22. Profit J, Gould JB, Bennett M, ym. Racial/ethnic disparity in nicu quality of care delivery. *Pediatrics* 2017;140:e20170918.

SUMMARY

Artificial intelligence and machine learning in neonatal intensive care

Electronic medical data are accumulating at an increasing pace, especially in intensive care where integrated patient information systems and constant monitoring of vital signs produce detailed datasets. This data can be analysed by machine learning algorithms to predict mortality and common complications. These models can incorporate numerous variables and utilize time series data, simultaneously producing good prognostic power. However, the complexity of the models may obscure how results were achieved, which has been called the black box problem. As machine learning models are increasingly offered as solutions to clinical problems, understanding the limitations and possibilities offered by AI is important for doctors.