

Big Data Fishing

Tusinder og atter tusinder af patienter lægges hvert år i fuld narkose på landets hospitaler og privatklinikker. Datamængden vi genererer fra hver patient synes uendelig. Det *må* naturligvis være Big Data, og det *må* være oplagt at benytte de nye statistiske modeller for Big Data, som Google, USA's spiontjeneste NSA, MIT, m.fl. bruger og konstant udvikler på. *Virkeligheden* for de data, vi genererer perioperativt, er imidlertid en anden. Langt den største datamængde er af ren administrativ karakter, grovmasket og løst indtastet.

For at vi skal kunne snakke om Big Data og kunne applicere avancerede statistiske algoritmer, skal en række kriterier være opfyldt. Man skal overordnet forholde sig til de fire V'er: Volume, Variety, Velocity og Veracity. Volumen skal være stor, rigtig stor. Hvis du kan have data genereret på en dag på en ny hjemmecomputer, er det som en tommelfingerregel ikke Big Data. Databaserne bygges ikke op med rækker og kolonner, men med en "NoSQL"-arkitektur. Variationen skal være stor, det vil sige data skal komme fra mange forskellige kilder og inkludere mange formater og koncepter. Hastigheden på indsamlingen af data skal være meget høj og så tæt på real-time som muligt.

Det fjerde V - Veracity - gælder for alle typer forskningsdata, og det får lov til at stå alene her, så vigtigt er det. Det refererer til sandhedsværdien af data. Fejlbehæftede data vil være fejlbehæftede, uanset hvilken statistisk model vi applicerer.

Hvordan er perspektiverne for brug af Big Data i perioperativ sammenhæng? Bruger vi hjertefrekvens som eksempel på en variabel i en to timers anæstesi, så vil en registrering hvert 5. min generere 94 bytes data, men en waveform sampling på et 5-afledt EKG vil generere 37 MB. Hvis vi kobler waveform sampling for alle de fysiologiske variable, som vi registrerer med eksempelvis prospektiv indsamlet genotype på en given patient og i sidste ende en stor kohorte af patienter, så begynder vi at træde ind i en ny æra af anæstesiologisk forskning og personaliseret behandling.

Læs mere her:

Levin MA et al. "Data, Big Data, and Metadata in Anesthesiology", *Anesthesia & Analgesia* 2015, 121;6;1661-1667. PMID: 26579664.

[LINK](#) til artiklen.

Martin Kryspin Sørensen
Emneredaktør