

Statistik og lægemidler

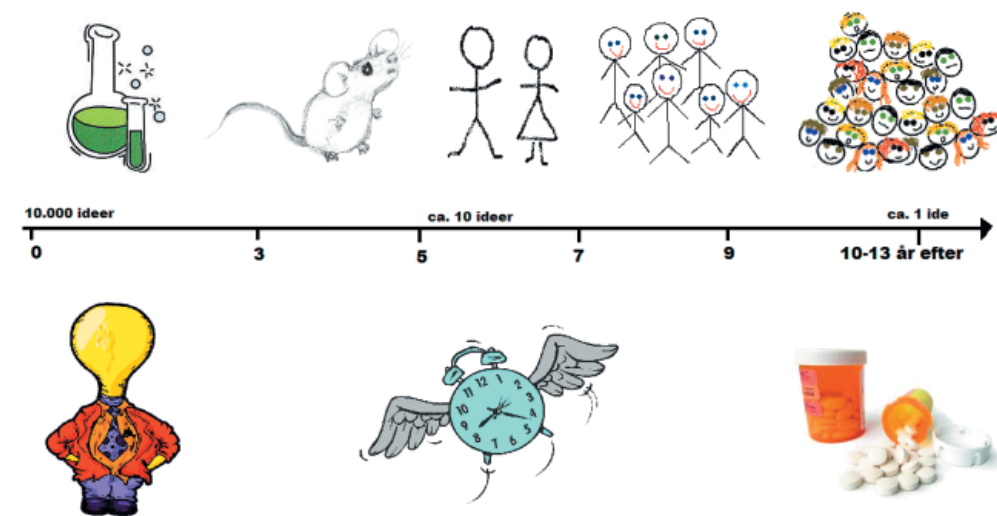
Statistiske analyser er vores sikkerhed for, at medicin virker som den skal og ikke har alt for alvorlige bivirkninger. Derfor spiller statistik en væsentlig rolle inden for lægemiddeludvikling. For at undgå snyd er der strenge regler for, hvordan medicinske forsøg skal udføres.

Af Trine Saugstrup

■ Hvor fremmedartet statistik end kan forekomme for den almindelige borger, er statistik som redskab afgørende for udviklingen af produkter, der er yderst nærværende. Det gælder ikke mindst for udviklingen af lægemidler. Som forbrugere af medicin har vi alle brug for en sikkerhed for, at den medicin, vi tager, virker efter hensigten – ligesom vi er interesserede i, at der hele tiden bliver udviklet ny medicin. Statistik er et af forskernes centrale værktøjer i udviklingen af lægemidler. Statistiske analyser er i sidste ende forbrugernes garanti for, at virkningen af et lægemiddel ikke bare er en trossag.

Medicin vs. placebo

Udviklingen af et lægemiddel er en lang og omstændelig proces, som det fremgår af artiklen om *Den lange vej mod sikre lægemidler* i dette nummer. Ud af 10.000 ideer, der fostres af forskere i laboratoriet, er det kun ca. 10, der ender med at blive omsat til et stof, som faktisk bliver testet på mennesker. Og ud af disse 10 stoffer er det kun ca.



Udviklingen af lægemiddel er en lang proces, der starter med en ide i forskerens hoved og går over forsøg med cellekulturer og forsøgsdyr, forsøg på få personer og derefter mange personer for til sidst at ende som færdigt lægemiddel, der kan købes på apoteket.

et enkelt, der i sidste ende bliver til et lægemiddel, som kan købes på apoteket. Hele denne proces tager typisk omkring 10-13 år.

Statistikken spiller en vigtig rolle i denne proces, særligt når det skal undersøges, om medicinen nu også virker på mennesker som den skal.

For at undersøge dette sammenlignes effekten af medicinen ofte med effekten af “sny-

demedicin” kaldet placebo. Et placebo er et stof, der bevisligt ikke har nogen medicinsk virkning, som for eksempel en kalkpille. Inden man kan sammenligne effekten af medicinen, bestemmer man sig for, hvilke kriterier, der har betydning for om medicinen virker – det kan for eksempel være ændringer i blodtryk, antallet af astma-

anfald, patienten har efter at være begyndt at tage medicinen eller ændringer på en godkendt effektskala, der fortæller noget om, hvordan patienten har det. Hvilke kriterier man vælger afhænger selvfølgelig af sygdommen. Effektskalaer bliver typisk brugt inden for psykiatrien og patienterne sammenlignes typisk efter en form for samlet score eller ordnede kategorier, så

som "forværret tilstand", "ingen ændring" eller "forbedret tilstand". Kriterierne kaldes i kliniske forsøg for *effektmål*.

Testpersonerne i forsøget modtager typisk enten placebo eller det lægemiddel, man vil undersøge, i en tilfældig opdeling. Man måler herefter i en fastlagt tidsperiode på hvert af de forudbestemte effektmål. Resultaterne af alle målingerne samles ind i et stort datasæt.

Statistiske hypoteser

Statistikerens job er nu at undersøge, om de patienter, som har modtaget medicinen, har det væsentligt bedre, end de patienter, der har modtaget placebo. Statistikeren opstiller typisk en hypotese, der siger, at medicinen har samme – eller dårligere – effekt på sygdommen end placebo med hensyn til de valgte effektmål. Et positivt resultat er altså – set fra lægemiddeludviklerens side – hvis man kan forkaste hypotesen og dermed vise, at medicinen har en bedre effekt på sygdommen end placebo.

Der findes ikke blot en enkelt statistisk metode til at teste den opstillede hypotese, som statistikeren benytter sig af hver gang. Derimod findes der mange forskellige metoder, der afhænger af hvordan forsøget er designet og hvilken type effektmål, der betragtes.

Der er næsten uanede muligheder for, hvordan et forsøg kan designes. Forsøget kan for eksempel indeholde flere forskellige doser af lægemidlet, placebo eller andre lægemidler, der allerede er på markedet. Patienterne kan skifte mellem de forskellige behandlinger i løbet af forsøget. Forsøget kan løbe over en længere periode, hvor patienten modtager medicinen mange gange og effektmålene bliver målt gentagne gange på samme patient. Eller forsøget kan bestå af en enkelt behandling og en enkelt måling af effektmålene for hver patient osv.

Forskellige typer af effektmål giver forskellige typer af data, som har indflydelse på, hvilken statistisk metode der skal benyttes. Blodtryk er f.eks. såkaldt



Milepæle mod sikre lægemidler

Før i tiden var ingen generelle retningslinjer for, hvordan medicinske forsøg skulle udføres, hvilket kunne have fatale følger. I 1937 var der således et tragisk tilfælde i USA, hvor et sulfapræparat dræbte mere end 100 børn. Det fik den daværende præsident Franklin D. Roosevelt til at grundlægge sundhedsmyndigheden *Food and Drug Administration* (FDA), der populært sagt skulle holde medicinal- og fødevarerbranchen i ørerne.

I 1964 udkom den første deklaration (Helsinki-deklaration) – skrevet af læger for læger.

I 1977 indførte FDA som de første retningslinjer for *Good Clinical Practice* (GCP). Disse bygger på Helsinki-deklarationens principper, men er mere detaljerede.

I 1980'erne fulgte flere europæiske lande med tilsvarende retningslinjer.

I 1995 blev *European Medicines Agency* (EMA) grundlagt som et europæisk modstykke til FDA.

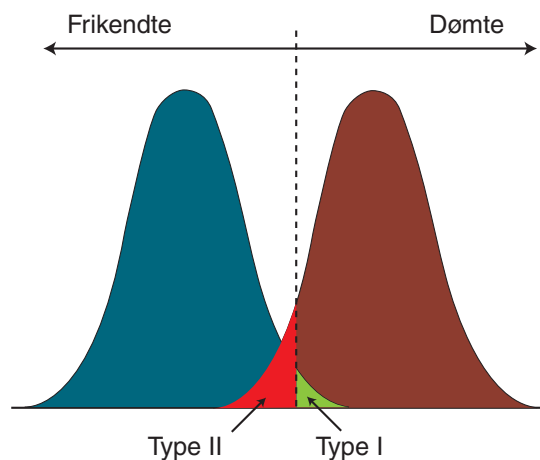
Og i 2004 blev det et lovmæssigt krav, at alle medicinalvirksomheder i Danmark skal følge GCP.

Typer af fejl

Inden for lægemiddeludvikling betegner en type I fejl, at man forkaster en sand hypotese, mens en type II fejl betyder, at man accepterer en falsk hypotese. Begge dele ønsker man selvfølgelig at undgå.

For bedre at kunne forstå sammenhængen mellem type I fejl og type II fejl, kan man illustrere det med et retssystem, som kan være nemmere at forholde sig til. I retssystemet kan man også begå to typer af fejl, en type I fejl er når en uskyldig bliver dømt og en type II fejl er når en skyldig bliver frikendt. Også her ønsker man at eliminere antallet af begge typer af fejl. Dette er ikke så enkelt, da udfaldet af en retssag afhænger af flere ting, så som bevismateriale, de mistænkte fremtræden og de mistænkte advokater. De to kurver på figuren repræsenterer hhv. de uskyldige og de skyldige mistænkte. Betragtes først de uskyldige, repræsenterer venstre side af kurven mistænkte med et vandtæt alibi, hvorimod de uskyldige til højre side er de uskyldige mistænkte, hvor alt bevismateriale peger i deres retning. De skyldige i venstre side er de "smarte" kriminelle med gode advokater og de skyldige til højre er de "dumme"

	Uskyldig	Skyldig
Dømt	Type I fejl	Korrekt
Frikendt	Korrekt	Type II fejl



kriminelle med de dårlige advokater. Figuren viser også sammenhængen mellem de to typer af fejl. For ønsker man at sænke antallet af uskyldige, der bliver dømt (type I fejl) skal skillelinjen mellem

frikendte og dømte rykkes yderligere til højre, hvilket resulterer i at flere skyldige bliver frikendt, hvilket er en forøgelse af type II fejl og omvendt. Kunsten er altså at finde den rette balance.

Når det går galt

Et af lægemiddelhistoriens mest berømte skrækkempler er *thalidomid*, der i perioden 1957-64 blev givet til gravide kvinder mod kvalme. Midlet forårsagede svære misdannelser på tusindvis af børn. Årsagen var, at det kemiske stof findes i to kemisk set identiske, men rumligt forskellige, udgaver (et fænomen, der kaldes kiralitet). Den ene variant havde den ønskede virkning, mens den anden variant havde den uheldige bivirkning. En moderne krølle på historien om *thalidomid* er i øvrigt, at man for nylig med succes er begyndt at bruge stoffet mod den sjældne kræftsygdom myelomatose.

På trods af skærpede krav til forsøg, sker det alligevel hvert år, at medicin må trækkes tilbage fra markedet pga. uheldige bivirkninger. Eksempler fra nyere tid er:



Efalizumab (Raptiva) – et middel mod psoriasis, der i juni 2009 er trukket tilbage, da det har vist sig at øge risikoen for en dødelig hjernesygdom forårsaget af virusinfektion.

Rimonabant – et sultnedsættende middel mod overvægt, der blev trukket tilbage i 2008, fordi det øgede risikoen for svær depression og selvmord. I denne forbindelse er bumsemidlet *Accutane* (som i Danmark sælges under navnet *Roaccutan*) under mistanke for de samme bivirkninger. Sagen har fået stor bevågenhed i USA, fordi det amerikanske kongresmedlem Bart Stupak har involveret sig kraftigt i sagen. Stupaks søn begik selvmord tilbage i år 2000 efter en bumsekur, og det er siden kommet frem, at der allerede dengang var mistanke mod midlet.

Vioxx – et middel mod gig, der blev trukket tilbage i 2004, da det viste sig at kunne medføre blodprop- og hjerteanfald. Fra midlet kom på markedet i 1999 havde mere end 80 millioner mennesker på verdensplan taget pillerne. En lang række retssager er blevet rejst mod firmaet Merck & Co. i kølvandet på historien – bl.a. af 108 patienter i Danmark.

Om forfatteren



Trine Saugstrup er cand. scient. i statistik og ansat hos Novo Nordisk som forsker.
trinesaugstrup@gmail.com

Artiklen er oprindeligt skrevet til et symposium for studerende, der er blevet tildelt et scholarstipendium af virksomhederne Novo Nordisk og Novozymes.

kontinuerte data, da blodtryk måles som en reel værdi på talskalaen. Antallet af anfald er derimod diskrete data, da anfald kun kan tælles sammen som heltal, og en effektskala kan være kategoriske data (dvs. f.eks. "forbedret tilstand").

Statistiske fejl

Når den statistiske metode er valgt og hypoteserne bliver testet, er der selvfølgelig en fare for, at man træffer en forkert beslutning. En forkert beslutning kan enten være, at man forkaster en sand hypotese (hvilket kaldes en type I fejl), eller at man accepterer en forkert hypotese (en type II fejl).

Man kan udregne en såkaldt p-værdi, som angiver sandsynligheden for, under hypotesen, at få et resultat, der er mindst lige så ekstrem, som det fundne resultat. Jo mindre p-værdien er, jo mere usandsynligt er det, at se et resultat, der er mere ekstremt end det fundne, hvilket gør det mere usandsynligt, at hypotesen er sand, og dermed stiger sandsynligheden for at afvise hypotesen. Inden forsøget går i gang, bestemmer man sig for et *signifikansniveau*. Denne stør-

relse angiver den grænse, hvor man siger, at nu vil man ikke længere tro på hypotesen, men vælger i stedet alternativet. Den mest brugte værdi er 0,05.

Det er dog ikke så simpelt at bestemme signifikansniveauet, når man har flere effektmål, hvilket ofte er tilfældet i kliniske forsøg. Det kan måske være, at man måler på blodtryk, blodprocent og lungkapacitet samtidig og alle tre effektmål er med til at bestemme, om medicinen har en positiv effekt. Man kan nemlig ikke blot teste hvert effektmål ved 5 % niveau, da den samlede sandsynlighed for at man forkaster en sand hypotese (type I fejl) i så fald bliver større end 5 %.

Det er derfor meget vigtigt for statistikerne at vælge den helt rette metode til at bestemme det niveau, der resulterer i at antallet af type I fejl samlet set holdes under det valgte signifikansniveau, mens antallet af type II fejl holdes på et minimum i den givne situation.

Svært at snyde

Der er som tidligere nævnt mange måder at gøre tingene på, når man er statistiker. Så det

er jo nærliggende at spørge sig selv om, det så ikke er muligt for statistikerne at snyde og finde en metode, der giver positive resultater, så medicinalvirksomheden kommer til at tjene penge og dermed sikrer statistikerens løn i fremtiden.

Det er det heldigvis ikke. I dag skal forsøgsdesignet og alle analyser leve op til standarden for *Good Clinical Practice* (GCP), som er en international etisk og videnskabelig kvalitetsstandard for forsøg, hvor mennesker deltager samt godkendes af alle involverede sundhedsmyndigheder. De to største er *Food and Drug Administration* (FDA) i USA og *European Medicines Agency* (EMA) i Europa. FDA har omkring 9.000 ansatte, mens EMA huser omkring 500 ansatte og har et tæt netværk af yderligere 3.500 eksperter. Blandt de mange ansatte og eksperter i FDA, EMA og andre myndigheder er en stor del statistikere. Og de holder et vågent øje med alle de statistiske analyser, der udføres af deres kolleger, der arbejder rundt om i de mange medicinalvirksomheder i verden. ■

Videre læsning.

To danske lærebøger i statistik:

Erik M Bøye:
Deskriptiv statistik.

Erling B Andersen, Gorm Gabrielsen, Nils Kousgaard & Anders Milhøj:
Statistik for civiløkonomer.