

Bør risikoreduksjoner forklares ved hjelp av tall?

AV PEDER A. HALVORSEN

Denne kronikken er skrevet som et bidrag til boken «Skapar vården ohälsa? Allmänmedicinska reflektioner» – red. J Brodersen, B Hovelius, L Hvas. Lund: Studentlitteratur 2009. En dansk versjon av kronikken er tidligere publisert i Dansk Selskab for Almen Medicins medlemsblad Practicus nr 194, januar 2009.



Peder A. Halvorsen. F. 1967 i Øksfjord, Loppa kommune. Cand. med. Univ. i Tromsø 1995. Fastlege ved Nordlys legesenter i Alta, spes. allm.med. 2004. Stip. ved Forskningsenheden for Almen Medicin i Odense 2005–2008. Doktorgraden ble forsvart ved Syddansk Universitet, Odense, des. 2008. Fra aug. 2008 bistilling som forsker ved Nasjonalt senter for distriktsmedisin og Allmennmedisinsk forskningsenhet, Universitetet i Tromsø.

«Tar du ikke piller mot blodtrykket, vil du dø!»

Sitatet er hentet fra en tankevekkende kasuistikk publisert i Tidsskrift for Den norske legeforening (1). Pasienten var en asiatisk flyktning som ble tvangsinnlagt i psykiatrisk avdeling på grunn av påfallende adferd og mistanke om schizofreni. Pasientens identitet var usikker, og nærmeste pårørende var en person som utga seg for å være pasientens bror. Under innleggelsen ble det ved gjentatte undersøkelser påvist et blodtrykk på 230/140 mmHg. Pasienten nektet å la seg utrede eller medisiner under henvisning til at han var frisk. Man overveide en tid å starte blodtrykksbehandling under tvang, men pasientens angivelige bror kom avdelingen til unnsetning. Han brukte verken risikotall eller blodtrykksstall, men ga pasienten valget mellom blodtrykkspiller eller døden, slik sitatet ovenfor antyder.

Valget mellom medisiner og døden er selvsagt fiktivt, både i denne situasjonen og i tilfeller med mindre alvorlig hypertensjon. Ingen blodtrykksmedisiner er vist å kunne forhindre døden; i beste fall forlenger de livet. Spørsmålet om forebyggende medikamenter kan forhindre sykdom er mer komplisert. Hos diabetikere med hypertensjon og hyperlipidemi er det nærliggende å tenke seg at forebyggende medikasjon utsetter senkomplikasjoner snarere enn å forhindre dem, med mindre diabetikeren skulle være uheldig og dø tidlig av noe helt annet. Ved osteoporose virker det derimot plausibelt at bisfosfonater hos enkelte pasienter kan være den avgjørende faktor for at hoften tåler et fall. Frakturen blir i så fall forhindret.

Hvilken informasjon trenger man egentlig for å ta stilling til tilbud om forebyggende medikasjon? I følge forventet

nytteteori (2) består beslutningsgrunnlaget i prinsippet av kun to komponenter: Sannsynligheter for ulike hendelser (hjerteinfarkt, hjerneslag, medikamentbivirkninger etc.) med og uten behandling og verdsetting av de samme hendelsene (hvor alvorlige er konsekvensene av hjerteinfarkt eller et hjerneslag?). Målsetningen er den optimale beslutning, det vil si den som gir størst forventet nytte. Av dette følger at det er større grunn til å intervensere mot sannsynlige utfall enn mot mindre sannsynlige, og mot mer alvorlige hendelser sammenlignet med mindre alvorlige. Er alt annet likt, er det også større grunn til å anvende effektive intervensjoner framfor mindre effektive, og billige intervensjoner framfor dyre. Det er denne tenkningen som ligger til grunn for begreper som kvalitetsjusterte leveår og kostnadseffektivitet.

Sannsynligheter og effektmål er numeriske størrelser. Om man ønsker å fatte beslutninger i samsvar med forventet nytteteori, forutsetter det altså at man er i stand til å ta stilling til tall. Tradisjonelle tallstørrelser for medikamenteffekt er relativ risikoreduksjon (RRR), absolutt risikoreduksjon (ARR) og number needed to treat (NNT). Jo lavere NNT og jo høyere ARR og RRR, jo bedre effekt har medikamentet. Derfor skulle man forvente at når NNT går ned eller RRR går opp, øker andelen som takker ja til et forebyggende medikament. Dette er imidlertid ikke hva vi har observert i empiriske studier (TABELL 1). Tvert imot virker det som om andelen som takker ja er praktisk talt uavhengig av tallstørrelsen på NNT (3) og RRR (4), med mindre NNT blir meget høy eller RRR blir meget lav. Dersom man istedenfor uttrykker medikamenteffekten som utsettelse av

en uønsket hendelse, blir sensitiviteten for tallstørrelsen noe bedre, men ikke imponerende (TABELL 1) (5).

En umiddelbar reaksjon på disse funnene er at folk ikke forstår tall for risikoreduksjon. Det er sikkert noe i det, men er det så enkelt? Det er solide empiriske holdepunkter for at menneskers kognitive apparat ikke er spesielt godt tilpasset de beslutningsprosesser som forventet nytteteori forutsetter. Innen psykologien anvendes begreper som *heuristics and biases* (6) og *kognitive illusjoner* (7) for å beskrive kognitive prosesser vi anvender til å forenkle komplekse beslutninger. Slike mentale «short cuts» tenderer til å gi suboptimale beslutninger i henhold til forventet nytteteori. Til gjengjeld gjør de oss i stand til å ta mange beslutninger forholdsvis raskt i en verden med begrenset tid og med den informasjon vi til enhver tid har til rådighet (8).

Folks insensitivitet for effektstørrelser handler trolig om mer enn manglende tallforståelse, selv om dette nok også bidrar vesentlig (9). I det følgende omtales tre psykologiske mekanismer – *optimism bias*, *availability heuristics* og *affect heuristics* – som alle kan bidra til å forklare at folk er lite tilbøyelige til å vektlegge størrelsen på risiko og risikoreduksjoner i sine beslutninger. Det er gode holdepunkter for at disse mekanismene også kan gjøre seg gjeldende i klinikken (6, 10–12).

Optimism bias

Folk har en iboende tendens til å oppfatte seg som annerledes enn andre. Det er i seg selv ikke så urimelig – som individer er vi annerledes. Begrepet *optimism bias* antyder imidlertid at vi trekker denne annerledesheten lenger enn det er grunnlag for, og da oftest i en bestemt retning. Vi tenderer til å vektlegge det positive mer enn det negative, og denne tendensen gjennomsyrrer språk, persepsjon, hukommelse og vurderinger (7). Kort sagt – mange av oss – trolig flertallet – har en iboende optimist. Studier har for eksempel vist at folk gjennomgående vurderer sine egne og sine

nærmestes egenskaper som bedre enn gjennomsnittet (7). Videre at man oppfatter sin risiko for å bli rammet av trafikkulykker, sykdommer, kriminelle handlinger, uønskede svangerskap etc. som lavere enn gjennomsnittet. Fenomenet går også under navnet *Lake Wobegon-effekten* etter en fiktiv landsby i Minnesota (7). Radiostasjoner i USA og Canada har i over 30 år kringkastet nyheter fra Lake Wobegon, «where all the women are strong, all the men are good looking, and all the children are above average». Befolkningen i Lake Wobegon er for øvrig av norsk opphav.

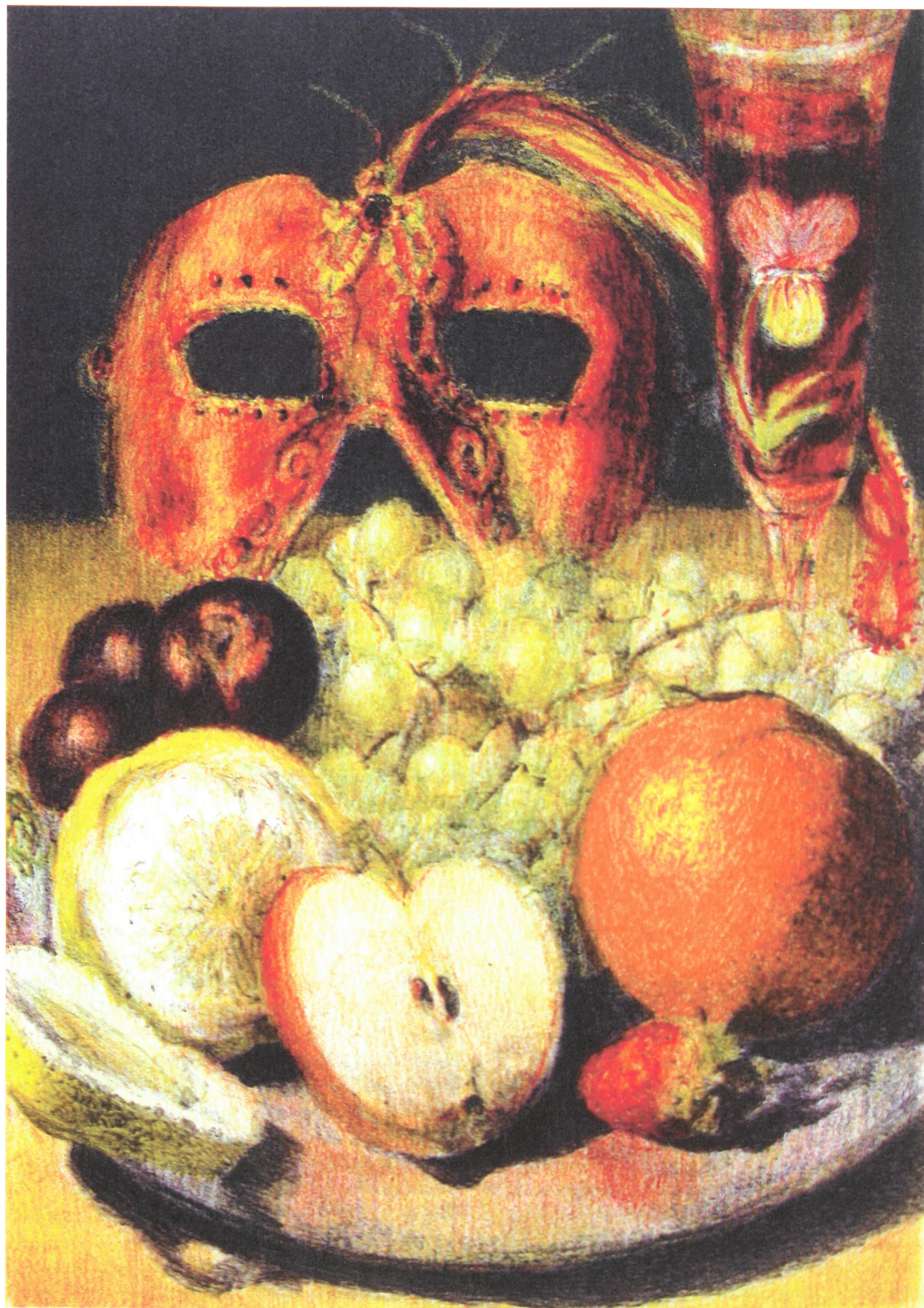
At det finnes pasienter som undervurderer sin risiko for sykdom er velkjent både fra klinisk erfaring og empiriske studier (10). Et annet plausibelt utslag av *optimism bias* er urealistiske forventinger til hva man kan oppnå med forebyggende medikamenter. Studerer vi TABELL 1 er det slående at to tredjedeler ville takket ja til et forebyggende medikament der NNT var 1600. Dette tilsvarer en ARR på 0,000625. Å påvise slike effektstørrelser i kliniske studier krever så store datamaterialer at det knapt kan gjennomføres. Nesten halvparten av de spurte syntes at ti prosent relativ risikoreduksjon virket attraktivt, mens 40 prosent takket ja til *én måneds* gjennomsnittlig utsettelse av hjerteinfarkt. Kan det være slik at den iboende optimisten i oss tenker: «1600 for å forebygge et tilfelle? Synd for de 1599 andre!»? Og at ti prosent og én måneds utsettelse først og fremst gjelder *de andre*?

Leger er også mennesker. Våre mange suboptimale beslutninger fører iblant til underbehandling og oversett alvorlig sykdom («det vanligste er vanligst»). Allmennleger kritiseres også for passivitet i forhold til implementering av retningslinjer for risikodiagnostikk og behandling. Fenomenet – *clinical inertia* (13) – har helt sikkert sammensatte årsaksforhold, herunder kritisk refleksjon i allmennpraksis. Jeg ser ikke bort fra at vår iboende optimist kan være en alliansepartner: «Retningslinjene passer ikke så godt for *mine* pasienter.»

TABELL 1. Legfolks behandlingsvilje som funksjon av størrelsen på risikoreduksjoner beskrevet ved hjelp av numbers needed to treat (NNT), relativ risikoreduksjon (RRR) og utsettelse.

«NNT pasienter må ta medisinen i 5 år for å forebygge ett sykdomstilfelle»		«Medikamentet reduserer risiko for hjerteinfarkt med RRR % hvis det brukes resten av livet»		«Medikamentet utsetter hjerteinfarkt med x mnd/år hvis det brukes resten av livet»	
NNT ³	Andel som takket ja	RRR ⁴	Andel som takket ja	Utsettelse ⁵	Andel som takket ja
50	76 prosent	10	48 prosent	1 måned	39 prosent
100	71 prosent	20	66 prosent	6 måneder	52 prosent
200	70 prosent	30	62 prosent	12 måneder	56 prosent
400	71 prosent	40	57 prosent	2 år	64 prosent
800	68 prosent	50	60 prosent	4 år	67 prosent
1600	67 prosent	60	57 prosent	8 år	73 prosent

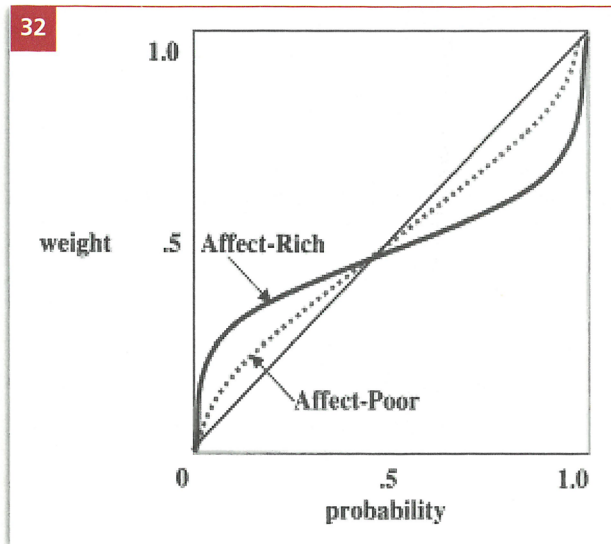
NNT: Number needed to treat RRR: Relativ risikoreduksjon



Availability heuristics

Forestill deg et øyeblikk at din lege meddeler at du har økt risiko for å få hjerteinfarkt. Legen foreslår at du kan bruke et medikament for å forebygge hjertesykdom. Medikamentet må tas daglig, og du må gå til kontroll hos legen to

ganger i året. Bivirkninger er sjeldne og ufarlige. Medikamentet koster 1000 kroner per år. Legen forteller at for hvert tilfelle av hjerteinfarkt som unngås, må 50 pasienter bruke medikamentet i fem år. Ville du velge å bruke et slikt medikament?



FIGUR 1. Hypothetical affect-poor and affect-rich weighting functions (and an identity line). ETTER ROTTENSTREICH OG HSEE 2001¹⁰, GJENGITT MED TILLATELSE.

Slike hypotetiske scenarier ligger til grunn for funnene gjengitt i TABELL 1. Det kan umiddelbart innvendes at de ikke inneholder tilstrekkelig informasjon. Likevel er det komplekst nok. En potensiell pasient blir presentert for en alvorlig sykdom (hjerteinfarkt), legens råd, ordet «forebygge», 1000 kroner per år, daglig medikasjon i fem år, mulige bivirkninger og NNT på 50, faktorer som alle er mer eller mindre relevante for beslutningen som skal fattes. Stilt overfor kompleks informasjon er det en tendens til at vi legger størst vekt på de faktorene som er lettest å evaluere, enten på grunn av gjenkjennelse (*the recognition heuristic*) (8), at faktoren er lett å huske/lett tilgjengelig for bevisstheten (*the availability heuristic*) (6) eller på grunn av kunnskaper, erfaring eller tilgjengelig sammenligningsgrunnlag (*the evaluability hypothesis*) (14). Tilsvarende er det en tendens til å se bort fra informasjon som ikke umiddelbart er lett å evaluere. Dette kan føre til manglende vektlegging av tallstørrelser vi ikke er fortrolige med. En studie viste for eksempel at folks betalingsvilje for et oljesaneringsprogram var uavhengig av hvor mange fugler (2000 – 20 000 – 200 000) programmet kunne redde fra døden (14).

Er NNT på 50 godt eller dårlig? Man kan alltså forstå tallet 50, at så mange pasienter må behandles en viss tid før man observerer et hjerteinfarkt mindre, og at 50 er forskjellig fra for eksempel 100. Om folk flest skal vurdere hvorvidt 50 er bra eller ikke, kan de imidlertid neppe støtte seg på gjenkjennelse, hukommelse, kunnskaper eller erfaring. Den manglende sensitivitet for størrelsen på NNT og RRR som framkommer i TABELL 1 kan godt være et resultat av dette. Folk har trolig lagt større vekt på utsiktene til å forebygge en alvorlig sykdom, kostnader eller legens råd. At sensitiviteten for grad av utsettelse er større (TABELL 1), kunne også passe; folk flest har et naturlig erfaringsgrunnlag for å vurdere hva som er kort og lang tid. Videre er det vist at leger – som *kan* støtte seg på erfaring og kunnskap – er sensitive for størrelsen på NNT (15).

Affect heuristics – the faint whisper of emotion (11)

Følelser spiller en sentral rolle i våre beslutninger. Som overskriften antyder, dreier det seg ikke om de stormende emosjoner, men heller om de subtile, ubevisste. Mye tyder på at følelser utgjør en helt nødvendig støtte for vårt kognitive apparat. Psykologene antar at de erfaringer vi lagrer i underbevisstheten er affektivt ladet. Det affektive apparatet bidrar til at vi i løpet av få sekunder kan ta stilling til om noe vi drar kjensel på er bra eller mindre bra. Dersom tilgangen til det affektive apparatet hemmes på grunn av nevrologiske skader, er det vist at folk kan streve i timevis med trivielle beslutninger (16), for eksempel om man ønsker en timeavtale kl. 10.00 eller kl. 11.00 neste dag. Beslutninger blir ikke nødvendigvis bedre av å tenke nøye over saken.

Å aktivere det affektive apparatet er ikke vanskelig – en pose drops kan være tilstrekkelig. Randomiserte studier har vist at en pose drops påvirker hvordan folk besvarer spørreundersøkelser – deriblant en undersøkelse vedrørende jobbtilfredshet og kreativ problemløsning *blant leger* (12). I en studie basert på skriftlige pasientkasuistikker var leger mer tilbøyelige til å anbefale innleggelse eller henvisning dersom kasuistikken inneholdt en trussel fra pasienten om å gå til avisene eller klage til Fylkeslegen (17).

Selv om vårt affektive apparat er en effektiv veileder i mange av hverdagens valg og beslutninger, viser empiriske studier at det induserer en interessant effekt på vår risikopersepsjon: Overreaksjon på små og store sannsynligheter og manglende sensitivitet for midlere sannsynligheter (18). Som en illustrasjon kan man forsøke å bruke magefølelsen til å vurdere to ulike cancertvaksiner: Tenk deg at din livstidsrisiko for malignt melanom er en prosent. Det er utviklet en vaksine A som kan eliminere denne risikoen fullstendig. Vaksinen koster 5000 kroner og har kun trivielle bivirkninger. Tenk deg så en annen situasjon: Du får vite at pga. en genmutasjon er din livstidsrisiko for malignt melanom 50 prosent. Det er utviklet en vaksine B som reduserer denne risikoen til 49 prosent. Også denne vaksinen koster 5000 kroner og har kun trivielle bivirkninger. Vaksine A og B gir samme ARR (en prosent), men oppleves vaksine B umiddelbart like attraktiv som vaksine A?

Hvis en pose drops eller en hypotetisk trussel om klage til Fylkeslegen i en papirkasuistikk kan indusere affekt, er det rimelig å anta at ord som hjerteinfarkt og malignt melanom også kan det. FIGUR 1 antyder hvordan dette fungerer: Når affekt først er indusert, er tendensen at våre beslutninger ikke blir en lineær funksjon av sannsynligheter. Kurven blir bratt ved små avvik fra det sikre – sannsynligheter på

null og én – og temmelig flat over et bredt intervall av midlere sannsynligheter. Studerer vi tallene i TABELL 1, kan vi ane konturene av en slik kurve. FIGUR 1 antyder også hvorfor vaksine A framstår som mest attraktiv.

Er tall brukbare i en travel klinisk hverdag?

Jeg vet ikke. Flere menneskelige egenskaper og tilbøyeligheter – *optimism bias, availability heuristics* og *affect heuristics* – ser alle ut til å trekke i samme retning: Manglende vektlegging av tallstørrelser når komplekse beslutninger skal fattes. Dette ser ut til å stikke dypere enn manglende forståelse for eller aversjon mot tall som sådan. Sett i lys av forventet nytteteori er dette problematisk av flere grunner. På makronivå kan det føre til overforbruk av intervensjoner som ikke er kostnadseffektive og underforbruk av intervensjoner som er det. På individnivå kan det føre til beslutninger som er lite i samsvar med pasientens verdier. Tilsynelatende foreligger et paradoks: Om man ønsker at pasienter skal fatte beslutninger i samsvar med forventet nytteteori – hvor tallstørrelser har grunnleggende betydning – er det problematisk å forklare risikoreduksjon ved hjelp av tallstørrelser. Inntil videre må man kanskje velge andre strategier, slik den hypertensive flyktningens angivelige bror gjorde. Gigerenzer og medarbeidere viser til at mangel på grunnleggende tallforståelse – *statistical illiteracy* – er utbredt i befolkningen og argumenterer for at opplæring i statistikk og sannsynlighetsregning må starte allerede i barneskolen (19).

Mye tyder på at vårt kognitive og affektive apparat egentlig er godt tilpasset til hverdagslivets mange beslutninger (8). Disse må ofte fattes med begrenset tilgang til informasjon og med dårlig tid til rådighet. Da gjelder det å kunne fatte beslutninger som er tilstrekkelig gode til å leve med. Jakten på den optimale beslutning kan kreve mer tid og ressurser enn vi faktisk har. Dette er for øvrig betingelser som er lett gjenkjennelige for allmennpraktiserende leger. Hverdagen kan beskrives som en kontinuerlig strøm av beslutninger, noen store, mange små og trivielle. Mange beslutninger fattes svært hurtig, nærmest automatisk, basert på intuisjon og mønstergjenkjennelse (6), trolig ved å utnytte sterke sider av vår felles menneskelige kognitive og affektive utrustning. De samme beslutninger kan imidlertid føre til både over- og underforbruk av intervensjoner som for eksempel antibiotika, billeddiagnostiske metoder, screeningundersøkelser, antihypertensiva og statiner. Metoder for å bringe beslutninger mer i samsvar med forventet nytteteori kan ha mye for seg. Det forskes mye på utvikling av retningslinjer og beslutningsstøtte hvor tall for risikoreduksjoner inngår som en sentral komponent. Metodeproblemene er imidlertid mange (20), og det er ikke avklart om dette fører til bedre beslutninger.

Referanser

1. Gulbrandsen P. Somatisk tvangsbehandling. Tidsskr Nor Lægeforen 2001; 121(29):3441.
2. Von Neumann J, Morgenstern O. Theory of games and economic behaviour. Third ed. John Wiley & Sons Inc; 1953.
3. Halvorsen PA, Kristiansen IS. Decisions on drug therapies by numbers needed to treat - A randomized trial. Arch Intern Med 2005; 165(10):1140–1146.
4. Sorensen L, Gyrd-Hansen D, Kristiansen IS, Nexoe J, Nielsen JB. Laypersons' understanding of relative risk reductions: randomised cross-sectional study. BMC Med Inform Decis Mak 2008; 8:31.
5. Dahl R, Gyrd-Hansen D, Kristiansen I, Nexoe J, Bo Nielsen J. Can postponement of an adverse outcome be used to present risk reductions to a lay audience? A population survey. BMC Med Inform Decis Mak 2007; 7(1):8.
6. Elstein AS, Schwarz A. Evidence base of clinical diagnosis: Clinical problem solving and diagnostic decision making: selective review of the cognitive literature. BMJ 2002; 324(7339):729–732.
7. Pohl R. Cognitive Illusions. A Handbook on Fallacies and Biases in Thinking, Judgement and Memory. 1st ed. East Sussex: Psychology Press; 2004.
8. Gigerenzer G, Todd P. Simple heuristics that make us smart. Oxford University Press; 1999.
9. Schwartz LM, Woloshin S, Black WC, Welch HG. The role of numeracy in understanding the benefit of screening mammography. Ann Intern Med 1997; 127(11):966–972.
10. Lloyd AJ. The extent of patients' understanding of the risk of treatments. Qual Health Care 2001; 10 Suppl 1:i14–i18.
11. Slovic P, Finucane ML, Peters E, MacGregor DG. Risk as Analysis and Risk as Feelings: Some Thoughts about Affect, Reason, Risk, and Rationality. Risk Anal 2004; 24(2):311–322.
12. Estrada CA, Isen AM, Young MJ. Positive Affect Improves Creative Problem-Solving and Influences Reported Source of Practice Satisfaction in Physicians. Motiv Emot 1994; 18(4):285–299.
13. Phillips LS, Branch WT, Cook CB, Doyle JP, El Kebbi IM, Gallina DL et al. Clinical inertia. Ann Intern Med 2001; 135(9):825–834.
14. Hsee CK, Loewenstein GF, Blount S, Bazerman MH. Preference reversals between joint and separate evaluations of options: A review and theoretical analysis. Psychol Bull 1999; 125(5):576–590.
15. Halvorsen PA, Kristiansen IS, Aasland OG, Forde OH. Medical doctors' perception of the «number needed to treat» (NNT). A survey of doctors' recommendations for two therapies with different NNT. Scand J Prim Health Care 2003; 21(3):162–166.
16. Loewenstein GF, Weber EU, Hsee CK, Welch N. Risk as feelings. Psychol Bull 2001; 127(2):267–286.
17. Kristiansen IS, Forde OH, Aasland O, Hotvedt R, Johnsen R, Forde R. Threats from patients and their effects on medical decision making: a cross-sectional, randomised trial. Lancet 2001; 357(9264):1258–1261.
18. Rottenstreich Y, Hsee CK. Money, kisses, and electric shocks: on the affective psychology of risk. Psychol Sci 2001; 12(3):185–190.
19. Gigerenzer G, Gaissmaier W, Kurz-Milcke E, Schwartz LM, Woloshin S. Helping Doctors and Patients Make Sense of Health Statistics. Psychological Science in the Public Interest 2007; 8(2):53–96.
20. Nelson WL, Han PKJ, Fagerlin A, Stefanek M, Ubel PA. Rethinking the Objectives of Decision Aids: A Call for Conceptual Clarity. Med Decis Making 2007; 27(5):609–618.

Evt. spørsmål og kommentarer kan rettes til:
e-mail: peder.halvorsen@kraftlaget.no