



# Brennpunkt: Legekoden

## Metoderapport, Data-SKUP

Bjørn Olav Nordahl, Maria Hasselgård, Espen Andersen

Marie Gjerver, Ole Jørgen Kolstadbråten

NRK, september 2018

# Innhold

|          |                                                     |           |
|----------|-----------------------------------------------------|-----------|
| <b>1</b> | <b>Innledning</b>                                   | <b>1</b>  |
| 1.1      | 9000 leger, 90 mill. regninger, 15 mrd. kroner...   | 1         |
| 1.2      | Tillit gir penger i kassa . . . . .                 | 2         |
| <b>2</b> | <b>Lukten av misbruk</b>                            | <b>3</b>  |
| 2.1      | En tørr og triviell paragraf . . . . .              | 3         |
| 2.2      | Sykt mye informasjon . . . . .                      | 3         |
| 2.3      | Å sladde 90 millioner regninger . . . . .           | 4         |
| <b>3</b> | <b>Analysen</b>                                     | <b>6</b>  |
| 3.1      | SQL og statistisk metode . . . . .                  | 6         |
| 3.2      | Fomlefasen . . . . .                                | 8         |
| 3.3      | Helsesjekk av helsedataene . . . . .                | 8         |
| 3.4      | Hjemmesnekret rapportverktøy . . . . .              | 9         |
| 3.5      | Max Otto peker ut den grådige fjerdedelen . . . . . | 10        |
| 3.6      | 100 år tilbake i tid . . . . .                      | 12        |
| 3.7      | Kollektivt hakeslipp . . . . .                      | 14        |
| 3.8      | Tall til folket . . . . .                           | 16        |
| <b>4</b> | <b>Stein til byrden?</b>                            | <b>17</b> |
| <b>5</b> | <b>Andre kilder</b>                                 | <b>17</b> |
| 5.1      | Koder for det meste . . . . .                       | 17        |
| 5.2      | Mangelfulle kontroller . . . . .                    | 17        |
| 5.3      | Innsyn i et skjebnefelleskap . . . . .              | 18        |
| 5.4      | Ikke grovt bedrageri? . . . . .                     | 18        |
| 5.5      | Deilige regnskaper . . . . .                        | 19        |
| 5.6      | Helsetilsynet leverer — mer i vente . . . . .       | 19        |
| <b>6</b> | <b>Tilbake til kilden</b>                           | <b>20</b> |
| <b>7</b> | <b>En passende cv</b>                               | <b>20</b> |
| 7.1      | Sjekk din lege — eller ikke? . . . . .              | 20        |
| 7.2      | Veien videre . . . . .                              | 20        |

# 1 Innledning

Det startet med et tips fra Drammen. En bioingeniør som tidligere hadde jobbet ved patologisk avdeling ved Drammen sykehus mistenkte kolleger for å signere på prøver de ikke hadde sett på for å utløse mer penger fra det offentlige. Vi undersøkte tipset, fulgte pengene og havnet hos Helfo.

## 1.1 9000 leger, 90 mill. regninger, 15 mrd. kroner...

Pengene som ble sendt til sykehuset i Drammen, kom fra det byråkratene tørt nok kalte "en ytre etat". Men Helfo viste seg å være langt mer sentral enn navnet kunne gi inntrykk av. Institusjonen lignet mer en bank innen Helse-Norge. Den betalte årlig ut 35 milliarder kroner til ulike helsetjenester som hadde avtale med det offentlige. Ikke minst til fastlegene. Helfo hadde 22 filialer som strakk seg fra Kirkenes i nord, via Verdalen i midten og ned til Sola i sørvest. Det undret oss at etaten stort sett hadde fått ligge i fred for grundig innsyn fra journalister.

Ikke til forkleinelse for en avdeling i Drammen. Viktig nok, her var det snakk om forsømte kreftprøver, kodetriksing som måtte undersøkes og avsløres. Det gjorde vi gjennom tradisjonelle gravemetoder som dokumentinnhenting, kildesamtaler, innsyn i kvitteringer ved avdelingen osv, noe som resulterte i viktige publiseringer i prosjektet "Legekoden". Men Helfo finansierte lite av sykehusets virksomhet. Fastlegene derimot var den største gruppen. Her var det snakk om fem milliarder i året. Like viktig: Alle innbyggere i Norge har en fastlege, de enten kan velge eller få utdelt. Pasientene kan bytte om de ikke er fornøyd. Dette ble derfor bare starten på en reise fra Drammen og inn på hvert enkelt fastlegekontor i Norge, der de "store" pengene befant seg. Helfo, "helsebanken" med den offentlige pengebingen, måtte undersøkes.

De tre siste årene hadde altså fastlegene forvaltet 15 milliarder kroner på vegne av storsamfunnet. De var vist stor tillit av Staten fordi pengene ble utbetalt direkte fra Helfo basert på regninger de sendte inn på hver pasient. Kontrollen var minimal sammenlignet med det vi så på andre samfunnsområder.

Vi søkte om innsyn i samtlige regninger til 9 700 leger i Norge over tre år og fikk data fra 90 millioner regninger i fanget. For å håndtere dette bygget vi en database og ved hjelp av statistisk metode som gikk side om side med databasespørringer og tradisjonelt kildearbeid. Her ble linjene trukket tilbake til et av de første eksemplene på datastøttet journalistikk, der nettopp forskningsmetode sto svært sentralt, jf side 6

I løpet av arbeidet produserte vi en datarapport på rundt 150 sider som tok for seg totalbildet, så vel som enkeltlegers refusjonspraksis. Rapporten var et resultat av egne skript, programmer og SQL-spørringer, som er nærmere beskrevet på side 9.

Vi implementerte egne versjoner av kjente algoritmer for utregning av ulike statistiske mål (se for eksempel Lorentz-kurven på side 11 og Gini-koeffisienten på side 12).

I tillegg til metodisk innsynsarbeid i ulike etater, direktorater og tilsyn, kobling av enorme datamengder, samt dialog med materialet, kunne vi sette fastlegenes refusjonsordning under debatt og blant annet avdekke:

- Fastlegers unødvendige undersøkelser og overbehandling av pasientene.
- Hvordan enkelte fastleger misbrukte tilliten overfor storsamfunnet ved å misbruke takssystemet til å skaffe seg millionlønninger. Vi kunne vise til hvem de var og hvordan de hadde misbrukt systemet, blant annet ved å sende inn krav om refusjon for undersøkelser som ikke hadde funnet sted.
- Hvordan Helfo-milliardene var skjevt fordelt mellom fastlegene.
- Mangelen på kontroll av 15 milliarder fastlege-kroner
- Hvordan fastleger og NAV-klienter ble behandlet ulikt i forhold til bestemmelsen for "grovt bedrageri" i straffeloven

- Hvordan alternativet med fast lønn av legevaksleger førte til mer fokus på pasienten og mindre fokus på regningen.

Prosjektet ble presentert som en 52 minutter lang dokumentar på NRK Brennpunkt i mai 2018. I samme periode presenterte vi nyhetssaker på radio, tv og nett.

Politikere ropte på flere anmeldelser av fastleger som jukset, andre tok til orde for å endre lønningssystemet med økt bruk av fastlønn.

Forskere tok også kontakt med ønske om å drive forskning på datamaterialet vi satt på. Noen hadde allerede søkt forskningsmidler, blant annet for å se om finansieringsordningen stod i veien for riktig behandling av pasientene, når oppgjørsformen kunne gi incentiv til å behandle mer enn nødvendig.

Interessant var også forskningsmiljøets utspill om også Helfo må bruke såkalte “big data” for å bedre kontrollene sine, slik som Brennpunkt-journalistene hadde gjort.<sup>1</sup> Helfos svar på manglende kontroll var at Helfo ikke kunne være “tilstede under konsultasjonen”.

– Men det var jo nettopp det de var! poengterte én professor.<sup>2</sup>

Med data fra timebestilling, journalføring, reseptutskrivning, regninger og så videre, kunne Helfo enkelt ha brukt metodene vi hadde brukt i prosjektet til å bedre sine kontroller.

Vi får se om “big data” også kan løse Helfos utfordringer framover. Helseministeren har bevilget nær 150 millioner kroner til å bedre digitale løsninger og bedre kontrollene. Helsedirektoratet utvikler nå en egen ordning så hver pasient kan sjekke regningen fastlegen sender på dem.

Dette hadde ikke vært mulig uten datastøttet journalistikk. Gjennom databasen vi har bygget med hver enkelt fastlege regninger, blant annet supplert med tilsynssaker på hver enkelt lege, har vi har bygget et verktøy vi vil bruke videre. I skrivende stund jobber vi videre med saker som vil bli publisert framover.

## 1.2 Tillit gir penger i kassa

Prislappen på legens arbeid bestemmes av et takstsystem der alt har sin unike kode. Hver kode utløser et fast kronebeløp som legen får. Vi bestemte oss for å følge hver eneste av disse 15 milliarder kronene fra statskassa og ned i lomma på fastlegen, på jakt etter forskjeller og misforhold.

Helsesystemet i Norge er basert på tillit. Pasienten kommer til legen sin og synes det er flott å bli sjekka overalt, at legen bruker god tid. Men dette kan legen spekulere i og tilliten kan misbrukes. En stor del av inntekten til fastlegen kommer i form av penger fra folketrygden som utbetales fra Helfo.

Utbetalingene gjøres på grunnlag av regninger som legen sender inn på hver pasient. Siden fastlegene i Norge jobber på stykkprisbasert lønn, er disse regningene svært detaljerte. De skal speile den behandlingen legen har gitt, og opplysningene som føres inn i pasientjournalen skal samsvare med refusjonskravene.

For eksempel er takstkoden **2cd** “*tillegg for konsultasjoner som varer over 20 minutter*”, og gir 201 kroner i refusjon. Kode **615** er “*samtaleterapi*”, som gir 170 kroner i refusjon. “Prøvetaking” har takstkoden **701a** og gir 55 kroner i kassa. Og så videre<sup>3</sup>.

I treårsperioden 2015-2017 var 169 ulike takstkoder i bruk da 14 876 732 508 kroner (eller 14,9 mrd.) ble delt ut til fastlegene.

---

<sup>1</sup>Begrepet “big data” er hypet i businessmiljøene, og blir ofte fremholdt som noe nytt og revolusjonerende. I virkeligheten er “big data” ikke annet enn hva vi har holdt på med hele tiden, nemlig elektronisk dataanalyse av større informasjonsmengder for å finne mønstre og tendenser.

<sup>2</sup>Kai A.Olsen professor i informatikk, Høgskolen i Molde, UIB, MET

<sup>3</sup>Tall ifølge “Normaltariff for fastleger og legevakt 2018-2019” (normaltariffen.legeforeningen.no)



## 2 Lukten av misbruk

Inne hos Brennpunkt var Maria og Bjørn Olav i gang med å undersøke takstjuks blant patologene ved Drammen Sykehus. Samtidig kjente de på en tiltagende trang til også å se på fastlegenes omgang med takstsystemet. Maria hadde fått los på to takster spesielt. Den ene var for såkalt *samtaleterapi* og den andre for *konsultasjoner over 20 minutter*.

Her luktet det mulighet for misbruk lang vei og nå trengte hun tilgang til data. Det krevde et overbevisende innsynskrav.

### 2.1 En tørr og triviell paragraf

I naboredaksjonen, *Digital Historiefortelling*, satt Espen med hodet langt nede i et helvete av et databaseprosjekt som han trengte en pause fra. Han følte seg for så vidt sikker på at deler av kommunalforvaltningen også kunne trenge en pause fra ham, så da Maria kom bort for å lufte et mulig fastlege-samarbeid var han luttre øre.

Hun snakket i koder:

- 2cd og 615. Går det an å få utlevert en liste over hvor mange ganger fastlegene har brukt disse takstkodene?
- Sikkert. Men kan vi ikke bare be om å få ut alt sammen?
- Blir det ikke veldig mye informasjon å lete gjennom?
- Det er hva vi har databaser til.
- Tror du vi får det?
- Vi bruker § 9.

Denne tørre og tilsynelatende trivielle paragrafen hadde nesten umerkelig funnet veien inn i lovverket for snart ti år siden. Den er et ektefødt barn av EUs viderebruksdirektiv,<sup>4</sup> men har kanskje ikke vekket den helt store entusiasmen blant journalister flest:

*Alle kan krevje innsyn i ei samanstilling av opplysningar som er elektronisk lagra i databasane til organet dersom samanstillinga kan gjerast med enkle framgangsmåtar.*<sup>5</sup>

### 2.2 Sykt mye informasjon

Utviklerne bak HELFOs fagsystem for administrasjon av utbetalinger må ha vært spesielt stolte av navnet de hadde gitt databasesystemet sitt. *“Kontroll og utbetaling av helserefusjoner”* fremstår riktignok ikke som noen merkevaresuksess, men akronymet KUHR ble til gjengjeld et ordspill av de bedre.

KUHR inneholder sykt (!) mye informasjon. Folkehelse bruker data derfra til å overvåke smittsomme sykdommer. Helsedirektoratet bruker systemet når de planlegger budsjetter, SSB lager nasjonal helsestatistikk og forskere bruker informasjon fra KUHR i ulike typer studier.

Og journalister..? De hadde vel egentlig aldri spurt om innsyn. Fram til nå.

Men så var det alle disse pasientene. De ville antakelig ikke likt at alskens liktær, herpesutbrudd, åpne sår og snodige føflekker ble med i databasedumpen til NRKs undersøkende redaksjon.

Her trengtes rett og slett litt tankevirksomhet rundt hvordan unngå å møte en vegg av taushetsplikt.

---

<sup>4</sup>EU-direktiv 2003/98/EF om viderebruk av offentlig informasjon (eur-lex.europa.eu).

<sup>5</sup>Offentleglova § 9 (lovdata.no)

**From:** [Espen Andersen](#)  
**To:** [post@helfo.no](mailto:post@helfo.no)  
**Bcc:** [Maria Hasselgård](#); [Ståle Hansen](#)  
**Subject:** Begjæring om innsyn i refusjonskrav fra fastlegene  
**Date:** torsdag 1. februar 2018 12:59:00

---

### **Begjæring om innsyn i refusjonskrav fra fastlegene**

Med hjemne i offentleglova §§ 3 og 9 ber vi om en sammenstilling fra HELFOs fagsystem for administrasjon og utbetaling av refusjonskrav.

Vi ønsker en oversikt over alle utbetalinger som er knyttet til refusjonskrav etter konsultasjon hos fastleger. For hver utbetaling ønsker vi følgende informasjon:

- Legens navn
- Unik ID på legen (fødselsnr, autorisasjonsnummer, organisasjonsnummer eller lignende)
- Legekontorets navn
- Unik ID for legekontoret (organisasjonsnummer)
- Tidspunkt for refusjon
- Takstkode
- Regningsnummer
- Refusjonsbeløp

Vi ønsker informasjon for årene 2015, 2016 og 2017.

Sammenstillingen bes utarbeidet på et passende maskinlesbart format som er egnet for gjenbruk (CSV-fil eller lignende). Vi går gjerne i dialog om hva som er praktisk mulig i forhold til eksisterende system og databaselayout.

Vennlig hilsen  
Espen Andersen,  
NRK Dokumentar og samfunn

Telefon: 911 58 223

**Figur 1:** Den første henvendelsen til HELFO, sendt 1. februar 2018. Vi er konkrete på hva vi ønsker av data, men åpner for videre dialog.

## **2.3 Å sladde 90 millioner regninger**

KUHR baserer seg på en såkalt *relasjonsdatabase*. Detaljert informasjon om konsultasjoner, resepter kirurgiske inngrep og mye annet er splittet opp i enkeltfelter og deretter organisert i rader og kolonner over flere forskjellige tabeller.

Databaser er hva vi kaller en *selvbeskrivende samling informasjon*. I praksis betyr det at en stakkars saksbehandler ikke selv ender opp hos fastlegen etter å ha forsøkt å tusje over taushetsbelagt informasjon i 90 millioner legeregninger. Ordet *selvbeskrivende* betyr nemlig at datasystemet allerede vet hvor for eksempel fødselsdatoene finnes. Saksbehandleren kan derfor sladde regningene i én operasjon, bare ved å instruere systemet om å utelate pasientens

```

2023687;Gjelsvik;Bjørn Espen;Tanum Legekontor;983479316;H1;6;6;6;912
2023687;Gjelsvik;Bjørn Espen;Tanum Legekontor;983479316;L1;205;417;417;7506
2023687;Gjelsvik;Bjørn Espen;Tanum Legekontor;983479316;L35d;4;4;4;5296
2023687;Gjelsvik;Bjørn Espen;Tanum Legekontor;983479316;L36d;3;3;3;2001
2023687;Gjelsvik;Bjørn Espen;Tanum Legekontor;983479316;L4;42;64;64;4751
2023687;Gjelsvik;Bjørn Espen;Tanum Legekontor;983479316;L40;23;27;27;20134
2023687;Gjelsvik;Bjørn Espen;Tanum Legekontor;983479316;L8;10;12;12;1764
2023687;Gjelsvik;Bjørn Espen;Tanum Legekontor;983479316;1ad;93;163;163;5651
2023687;Gjelsvik;Bjørn Espen;Tanum Legekontor;983479316;1bd;414;679;679;42777
2023687;Gjelsvik;Bjørn Espen;Tanum Legekontor;983479316;1e;2;2;2;68
2023687;Gjelsvik;Bjørn Espen;Tanum Legekontor;983479316;1f;43;114;114;10659
2023687;Gjelsvik;Bjørn Espen;Tanum Legekontor;983479316;1h;92;115;115;1690
2023687;Gjelsvik;Bjørn Espen;Tanum Legekontor;983479316;1i;381;577;577;31735
2023687;Gjelsvik;Bjørn Espen;Tanum Legekontor;983479316;10a;38;60;60;0
2023687;Gjelsvik;Bjørn Espen;Tanum Legekontor;983479316;10b;33;36;36;0
2023687;Gjelsvik;Bjørn Espen;Tanum Legekontor;983479316;10c;5;5;6;0

```

**Figur 2:** Utdrag fra en av filene vi fikk fra Helsedirektoratet. Hver linje består av 10 felter, adskilt med semikolon. Det sjette feltet er takstkoden.

fødselsnummer fra databasepostene.

Det gjør at journalister i teorien kan få ut detaljerte data fra noen av de mest sensitive datasamlingene som finnes uten at denne innsynsretten nødvendigvis kommer i konflikt med den lovpålagte taushetsplikten som forvaltningen må overholde.

Svaret kom mindre enn én uke etter at vi gjorde den første henvendelsen. Filen åpnet døren til et forvaltningssystem som inntil nå stort sett hadde fått være i fred for journalister.

Det hele var oppsummert over 1 521 921 linjer med strukturerte data.

Alt som skjedde på landets fastlegekontorer var registrert her, kodet opp og ledsaget av kronebeløp knyttet til enhver lege på hvert fastlegekontor (figur 2).

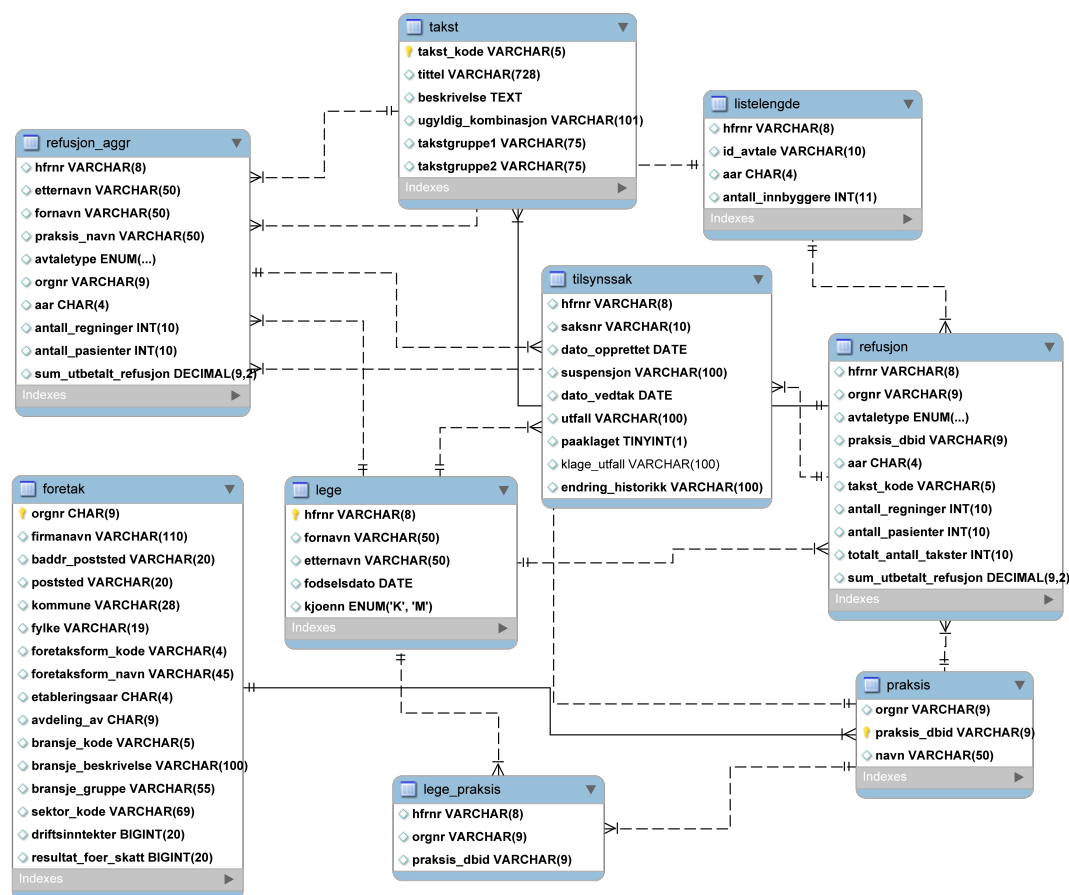
Hver linje beskrev én takstkode for én lege ved ett legesenter, og inneholdt følgende informasjon:

- Legens autorisasjonsnummer (HPR)
- Legens etternavn
- Legens fornavn
- Navn på praksis/legesenter
- Legens organisasjonsnummer
- Aktuell takstkode
- Antall pasienter det er krevd refusjon for etter denne takstkoden
- Antall regninger der denne takstkoden er brukt
- Antall ganger den aktuelle takstkoden er brukt
- Summert refusjonsbeløp knyttet til aktuelle takstkode

På side 6 og utover beskriver vi hvordan vi brukte en oppgradert utgave av databasemotoren MySQL for å organisere og analysere dataene.

### 3 Analysen

Hvordan får man oversikt, når data om 90 millioner fastlege-regninger har dumpet ned i innboksen? Vi hadde informasjon om hvordan 15 milliarder offentlige kroner fordeler seg på 9 700 leger i Norge over tre år. Her fantes bestefars inngrodde tånegl, lillebrors armbrudd og fars røykeavvenningsprat med legen. Men hvordan skulle de brukes?



**Figur 3:** Databaselayoten etter import av data fra ulike kilder i tillegg til Helfo-dataene. Her har vi bygd ut med datasett fra foretaksregisteret, innbyggerantall fra fastlegelistene, tilsynssaker fra helsedirektoratet, takstoversiktene med beskrivelser og kategorier samt data som ble skrapet ned fra webtjenesten til autorisasjonsregisteret for helsepersonell.

Brennpunkt-filmen “Legekoden”, som ble sendt første gang 2. mai 2018, er blant annet basert på funn gjort i denne databasen.<sup>6</sup>

#### 3.1 SQL og statistisk metode

Begrepet *datastøttet journalistikk*, eller CAR,<sup>7</sup> har med tiden blitt brukt på en rekke vidt forskjellige metoder og prosjekter; det finnes ingen vedtatt definisjon av akkurat hva begrepet

<sup>6</sup>tv.nrk.no/serie/brennpunkt/MDDP11000718/02-05-2018

<sup>7</sup>Computer Assisted Reporting

dekker.

Likevel kan det være nyttig å huske at den første boka om et datastøttet, journalistisk prosjekt kom så tidlig som i 1969. Da skrev journalisten Philip Meyer “Precision Journalism”.<sup>8</sup> Meyers bok var i realiteten en lærebok i metode og statistikk. Datamaskinen<sup>9</sup> var kun instrumentet som gjorde analysen mulig.

Meyer la vekt på at journalisten måtte ha fullt fokus på statistiske mål, feilmarginer og metodiske utfordringer. Hvorvidt datamaskinen mellomlagrer data i oppvarmede kvikksølvbad<sup>10</sup> eller på en moderne RAM-brikke, var i og for seg irrelevant for ham. For Meyer er datastøttet journalistikk (CAR) det samme som bruk av vitenskapelig metode innen journalistikk.<sup>11</sup> Nettopp derfor greide Meyer å skrive en bok om datastøttet journalistikk som fortsatt er relevant, selv etter 50 år.

Arbeidet med “Legekoden” aktualiserer nettopp hvor viktig det er at analysearbeidet er forankret i en forsvarlig bruk av statistiske mål og algoritmer.

En statistiker vil nok betrakte våre utregninger som nokså enkle og grunnleggende, men for oss var disse kjøringene avgjørende nettopp for å kunne underbygge de konkrete historiene enhver fjernsynsfortelling så sårpt trenger.

Vi kunne saktens funnet mange gode enkelthistorier også uten tilgang på data fra 90 millioner legeregninger. Men de ville blitt oppfattet som anekdotiske, og journalistikken ville manglet dybde og troverdighet.

Med slike datasett blir det viktig å ha flere tanker i hodet på én gang. På den ene siden er vi på jakt etter de gode historiene og enkelttekstemplene. På den annen side må vi finne tendensene og hovedlinjene. Og det ene må henge sammen med det andre.

Med en SQL-database kunne vi forfølge begge disse sporene parallelt. Men det stiller krav til journalisten. Som Philip Meyer skriver: “In short, a journalist has to be a database manager, a data processor, and a data analyst”.<sup>12</sup>

Når vi dytter researchen vår inn i en relasjonsdatabase<sup>13</sup> (også kalt en SQL-database), gjør vi i virkeligheten ikke annet enn å tilføre en *ekstrem struktur* til informasjonen vi har fått. Databasemotoren rører ikke innhold og betydning, men gjør den foreliggende informasjonen så uendelig mye mer tilgjengelig ved å strukturere dataene for å kunne gjøre presise, meningsfylte sammenstillinger.

Vi bruker spørrespråket SQL for å manipulere data i en slik database.<sup>14</sup> En SQL-spørring kan lagres for fremtidig bruk. Den kan repeteres mot nye og oppdaterte data, og den kan finjusteres, versjoneres og settes sammen med andre spørringer til store, komplekse rapporter.

Det er nettopp denne egenskapen ved SQL som blir viktig når virkelig store datamengder skal analyseres. Journalistikken blir gradvis til i en runddans mellom datakilder, endrede problemstillinger, alternative kilder og annen dokumentasjon. I arbeidet med “Legekoden” startet vi med et sett SQL-spørringer som ble skrevet om, justert og til og med forkastet i møtet med alternative kilder og nye erkjennelser.

---

<sup>8</sup>ISBN: 978-0742510883

<sup>9</sup>Meyer brukte en 7090 stormaskin, en av IBMs første transistorbaserte datamaskiner, som klarte rundt 24 000 regneoperasjoner i sekundet.

<sup>10</sup>Noen av de aller tidligste datamaskinene mellomlagret data som lydbølger i beholdere med oppvarmet kvikksølv (en.wikipedia.org/wiki/Delay\_line\_memory).

<sup>11</sup>www.theguardian.com/commentisfree/2011/dec/09/riot-theory-relative-detroit-england

<sup>12</sup>The New Precision Journalism, Philip Meyer, ISBN 978-0253206640.

<sup>13</sup>en.wikipedia.org/wiki/Relational\_database

<sup>14</sup>SQL står for “Structured Query Language”, og har eksistert siden 70-tallet.

### 3.2 Fomlefasen

I journalistisk sammenheng, snakker vi av og til om å “intervjue dataene”, ut fra ideen om at et datasett skjuler journalistiske poenger som man kan forløse ved nærmest å gå i dialog med materialet gjennom SQL-spørringer.

Det er en ganske god beskrivelse — men bare hvis du tenker deg intervjuobjektet som en trassig fireåring som svarer i enstavelser.

Vi kunne nok også ha beskrevet prosessen som at vi intuitivt forstod hvilke spørringer vi skulle foreta for å finne frem til misforholdene, og at det var journalistenes briljante intuisjon som førte oss fra den ene innertieren etter den andre. Men det hadde vært en tilsnikelse.

Når vi kaller dette avsnittet for “fomlefasen”, så er det i erkjennelse av at en fomlefase alltid vil være der — og nettopp dette ble et viktig læringspunkt for oss: Vi ble etter hvert svært mye bedre på å intervjue dataene.

Etter å ha kartlagt hvilke leger som kunne være interessante å se nærmere på, kunne vi gå videre til postjournalene til Helfo. Her kunne vi så be om innsyn i alle vedtak som var blitt gjort på de som toppet statistikken. Når vi så fant ut at det fantes vedtak om tilbakebetaling på flere av disse begynte det virkelig å bli interessant.

### 3.3 Helsesjekk av helsedataene

Det er fort gjort å snuble ned i riktig mørke fallgruver, spesielt når man bruker aggregerte datasett som utgangspunkt for videre kjøring.

Det er for eksempel viktig å være klar over hvilke verdier som *ikke* kan summeres; slik som for eksempel feltet *antall\_pasienter*, som angir hvor mange unike pasienter hver lege har levert regning for. Pasienter som har vært hos flere forskjellige leger i løpet av året ville ved en uforsiktig summering fra vår side ha blitt talt flere ganger, og tallet ville blitt fullstendig feil.

Vi brukte mye tid på å finne igjen våre egne tall hos andre kilder, gjerne i offisielle rapporter og statistikker. Hadde tallene våre skilt seg vesentlig fra andres utregninger, måtte vi ha søkt oss til bunns i årsaken.

```
mysql> SELECT
->   aar,
->   COUNT( * ) AS antall_linjer,
->   SUM( totalt_antall_takster ) AS antall_takster,
->   SUM( sum_utbetalt_refusjon ) AS utbetalt_refusjon,
->   COUNT( DISTINCT hfrnr ) AS antall_leger
-> FROM
->   refusjon
-> WHERE
->   hfrnr IS NOT NULL
-> GROUP BY
->   aar;
```

| aar  | antall_linjer | antall_takster | utbetalt_refusjon | antall_leger |
|------|---------------|----------------|-------------------|--------------|
| 2015 | 474429        | 67990964       | 4728995088.50     | 7145         |
| 2016 | 467680        | 69293207       | 4936591918.50     | 7192         |
| 2017 | 570815        | 69866417       | 5048552604.60     | 7318         |

3 rows in set (23.93 sec)

**Figur 4:** En rask spørring gir oss noen tall som kan sammenlignes med offisielle kilder. Vi ser blant annet at det samlede refusjonsbeløpet årlig ligger på rundt 5 mrd. kroner.

Senere ba vi Helfo og Helsedirektoratet om å få enda et uttrekk med antall pasienter per lege, kronebeløp og antall regninger. Fra dette siste datasettet, kunne vi få et bredere oversiktsbilde over materialet.

Dette stemmer i grove trekk overens med HELFOs egne måltall.

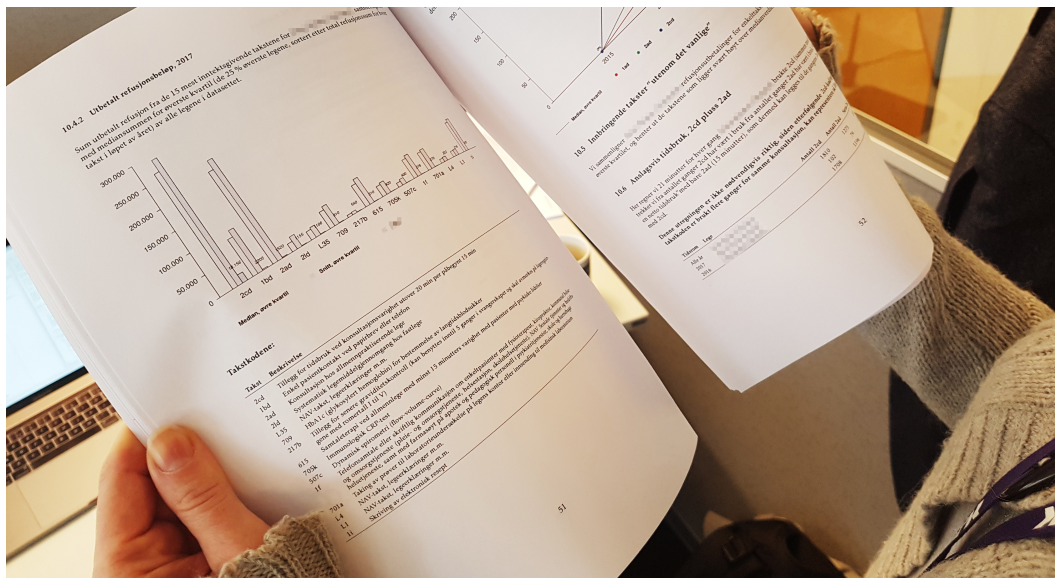
De første spørringene dreide seg som sagt om å få oversikt. De fleste databasemotorene har innebygd støtte for enkle statistiske funksjoner, slik som gjennomsnitt og standardavvik.

Vi har “pimpet opp” MySQL med flere programutvidelser, blant annet MySQL Infusion UDF av Robert Eisele,<sup>15</sup> som også gir oss funksjoner som lettere finner for eksempel persentiler — herunder også median — og typetall.

Vi splittet opp (normaliserte) dataene og fordelte dem på flere tabeller (se figur 3).

Senere utvidet vi databasen med flere kilder og flere tabeller (se figur 3 og også avsnittet om andre kilder, side 17), men dette var altså startpunktet for undersøkelsen.

### 3.4 Hjemmesnekret rapportverktøy



**Figur 5:** Et hjemmelaget skript konverterer SQL-spørringene våre til en leselig rapport med grafikk og tabeller.

Etter hvert som kontrollkjøringene og den pågående dialogen med helsemyndighetene fikk oss til å føle oss sikre på at vi hadde forstått dataene riktig, kunne vi begynne å drille oss ned i datasettene.

SQL-spørringene ble lange og mange. Vi endte opp med en SQL-fil på over 9 000 linjer med spørringer.

Noen av kjøringene ønsket vi å visualisere. Programvaren *Ploticus*<sup>16</sup> kom oss til unnsetning.

Vi gjorde både statistiske kjøringene, så vel som mer målrettede undersøkelser. SQL-spørringene ble konvertert til en leselig PDF-fil ved hjelp av et hjemmesnekrede program som vi ga navnet *mkres*,<sup>17</sup> utviklet for intern bruk.

Rapporten dannet grunnlaget for den videre undersøkelsen, og den vokste stadig, etter hvert som vi la inn nye kjøringene. Rapportdokumentet endte til slutt med å være på nærmere 150 sider.

<sup>15</sup>[github.com/infusion/udf\\_infusion](https://github.com/infusion/udf_infusion)

<sup>16</sup>[ploticus.sourceforge.net](https://ploticus.sourceforge.net)

<sup>17</sup>[github.com/espena/mkres](https://github.com/espena/mkres)



### 3.5 Max Otto peker ut den grådige fjerdedelen

Det var neppe mistanke mot legen sin som gjorde at økonomen Max Otto Lorentz utviklet den såkalte *Lorentz-kurven* i 1905.<sup>18</sup> Men han var ute etter det samme som oss: Å finne en måte å visualisere skjevheter i pengefordeling.

Siden vi ønsket å se på hvordan pengene faktisk er fordelt blant legene, grep vi fatt i hans formel (figur 6).

$$L(y) = \frac{\int_0^y x dF(x)}{\mu}$$

**Figur 6:** Formelen for Lorentz-kurven ser komplisert ut, men er egentlig enkel: Sorter legene stigende etter mottatt refusjonsbeløp. Start til venstre på X-aksen. For hver lege regner du ut *sammenlagt* refusjonsbeløp fra de foregående legene pluss gjeldende lege og deler på hva summen ville vært hvis alle hadde fått like mye (det vil si gjennomsnittet til nå, angitt med  $\mu$  i formelen). Resultatet for hver lege plottes inn på Y-aksen slik at vi får en stigende kurve.

Lorentz-kurven viser oss altså hvor mye den relle pengeutdelingen fra Helfo avviker fra en teoretisk (egalitær) situasjon hvor alle får like mye. Vi designet et SQL-skript som regnet ut punktene langs kurven for oss. Skriptet kjørte vi mot MySQL-basen med Helfo-dataene.

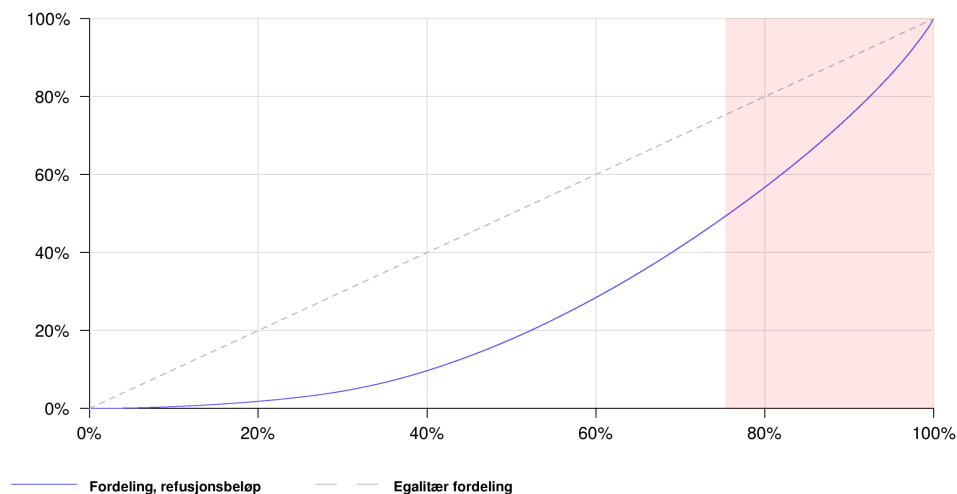
```
1 SET @rank = 0;
2 SET @refusjon = 0;
3
4 SELECT SUM( sum_utbetalt_refusjon ) INTO @max_refusjon FROM refusjon
5 WHERE aar = '2017' AND hfrnr IS NOT NULL AND avtaletype = 'Fastlege';
6
7 SELECT COUNT( DISTINCT hfrnr ) INTO @n_leger FROM refusjon
8 WHERE aar = '2017' AND hfrnr IS NOT NULL AND avtaletype = 'Fastlege';
9
10 SELECT
11 ( @rank := @rank + 1 ) / @n_leger AS 'Antall_leger',
12 ( @refusjon := @refusjon + a.sum_utbetalt_refusjon ) / @max_refusjon
13 AS 'Fordeling|_refusjonsbeloep',
14 ( @rank / @n_leger ) AS 'Egalitaer_fordeling'
15 FROM
16 (
17 SELECT
18 hfrnr ,
19 SUM( sum_utbetalt_refusjon ) AS sum_utbetalt_refusjon
20 FROM
21 refusjon
22 WHERE
23 aar = '2017'
24 AND
25 hfrnr IS NOT NULL
26 AND
27 avtaletype = 'Fastlege'
28 GROUP BY
29 hfrnr
30 ORDER BY
31 SUM( sum_utbetalt_refusjon ) ASC
32 ) a
33 ORDER BY
34 a.sum_utbetalt_refusjon ASC;
```

SQL-skriptet tok rundt 30 sekunder å kjøre og returnerte et datasett på 7318 rader. Ved hjelp av vårt egenutviklede verktøy (se side 9), kjørte vi så dette resultatet videre inn

---

<sup>18</sup>[en.wikipedia.org/wiki/Lorenz\\_curve](https://en.wikipedia.org/wiki/Lorenz_curve)

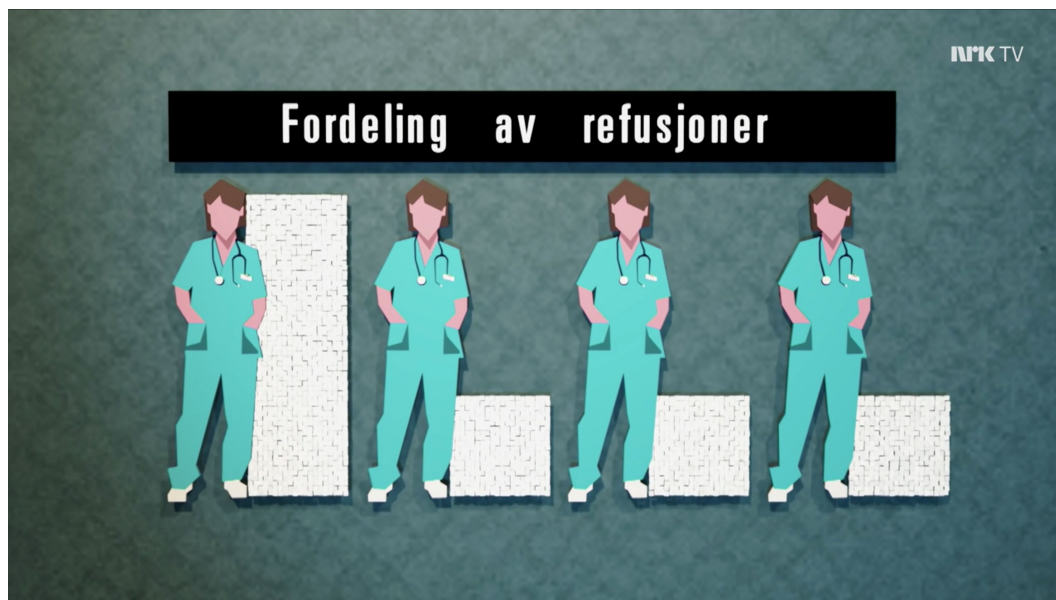
i Ploticus, som spyttet ut dette diagrammet:



**Figur 7:** Resultatet av SQL-spørringen. Lorentz-kurven i blått angir hvordan 5 mrd. skattekroner ble fordelt i 2017. Den viser prosentvis andel leger langs X-aksen, og kumulativ prosentandel av pengene langs Y-aksen. Her ser vi at en fjerdedel av legene (rødt felt) fikk omkring halvparten av pengene. Den stiplede linjen viser hvordan kurven ville sett ut hvis alle legene hadde fått like mye.

Lorentz-kurven kan leses slik: 75% av legene (X-aksen) fikk 50% av pengene (Y-aksen). Dermed må de siste 25% legene (rødt felt) ha fått resten av pengene. Slik kom vi fram til et av hovedpremissene i “Legekoden”:

*En fjerdedel av fastlegene stikker av med halvparten av pengene fra Helfo.*



**Figur 8:** Lorentzkurven, slik den til slutt ble på TV. Fjernsynsgrafikeren ønsker her å formidle at en fjerdedel av legene tar med seg halvparten av pengene fra Helfo.

### 3.6 100 år tilbake i tid

Legenes taksthefte inviterer til grenseløs kreativitet hva gjelder mulighet til inntjening. Samtidig gjelder det å gå under radaren til Helfos kontrollavdeling. Man må med andre ord styre unna de mest (mis)brukte takstene.

Fordi fokus for kontrollregimet ofte ligger på takster som utgjør store utbetalinger, og hvor det er en kjent risiko for misbruk, har noen enkeltleger funnet frem til obskure takster som er verre å avdekke, og som de bruker hele tiden.

Vi ville finne takstkodene *noen få leger anvendte mye mer enn alle de andre*.

Den naive tilnærmingen til denne utfordringen ville være å lage en liste over alle takstkodene, og så sortere etter hvor mange leger som faktisk har brukt dem. Det hadde imidlertid ikke gitt et godt svar på spørsmålet.

Takstene vi leter etter kan nemlig godt ha vært i bruk av svært mange leger. Vi er imidlertid ute etter de takstene et lite mindretall bruker *svært mye mer* enn kollegene.

Det finnes en hundre år gammelt formel vi i dag kan bruke for å hjelpe oss med dette. En statistisk utregning som opprinnelig var tiltenkt helt andre ting enn å avsløre særegen bruk av takstkoder.

Som leder for Italias statistikkbyrå Istat mellom 1926 og 1932, måtte Corrado Gini nesten daglig forholde seg til en lunefull Benito Mussolini. Gini var på sin side mer interessert i å bli husket som opphavsmannen bak *Gini-koeffisienten*,<sup>19</sup> fremfor å innlate seg med fascister — selv om formelen hans utvilsomt var egnet til å skape noen politiske poenger (figur 9).

$$G = \frac{A}{A+B}$$

**Figur 9:** Gini-koeffisienten  $G$  tar utgangspunkt i Lorentz-kurven, som vi omtalte ovenfor. Formelen dividerer arealet under Lorentz-kurven (dvs. den blå streken i figur 7), angitt som  $A$  i formelen, med arealet under streken for lik fordeling (dvs. den stiplede linjen i figur 7), angitt som  $A+B$  i formelen.

Formelen var opprinnelig designet for å måle fordelingslikheten av inntekt eller formue i en befolkning. Dersom Gini-koeffisienten av formue er 0, har alle innbyggerne like mye; dersom Gini-koeffisienten er (nesten) 1, er alt samlet på hendene til bare én innbygger.<sup>20</sup>

Så, hvorfor ikke bruke Gini-koeffisienten på noe helt annet enn penger? For eksempel på fastlegenes bruk av enkelttakster?

Vi laget et PHP-skript som gikk gjennom hver takstkode i databasen. For hver kode hentet vi ut alle legene som hadde brukt den aktuelle taksten, og i hvilket omfang de hadde brukt den. Med disse dataene kunne vi skrive en algoritme som regnet ut Gini-koeffisienten for alle de 169 takstkodene som var i bruk blant fastlegene i 2017.

---

<sup>19</sup>[en.wikipedia.org/wiki/Gini\\_coefficient](https://en.wikipedia.org/wiki/Gini_coefficient)

<sup>20</sup>Så lenge vi ikke opererer med negative tall, vil Gini-koeffisienten aldri kunne være nøyaktig 1 eller større. Men for alle praktiske formål kan vi runde av og si at den er 1 ved fullstendig ulikhet.

```

1 $db = new MySQLi( 'mysql', 'user', 'password', 'refusjon' );
2 $db->real_query( "SELECT DISTINCT takst_kode FROM refusjon" );
3 $res = $db->store_result();
4 $takstkoder = array();
5 while( $row = $res->fetch_assoc() ) {
6     $takstkoder[] = $row[ 'takst_kode' ];
7 }
8 $res->free();
9 $db->real_query( "DROP TABLE IF EXISTS takst_gini" );
10 $db->real_query( "CREATE TABLE takst_gini ..." );
11 $gini = array();
12 foreach( $takstkoder as $takst_kode ) {
13     $db->real_query( "SELECT ..." );
14     $res = $db->store_result();
15     if( $res ) {
16         $cum = 0;
17         $prv = 0;
18         $area = 0;
19         $n = $res->num_rows;
20         while( $row = $res->fetch_assoc() ) {
21             $cur = $row[ 'antall_takster' ];
22             $area += $cum + ( 0.5 * $cur );
23             $cum += $cur;
24         }
25         $gini[ $takst_kode ] =
26             array( 'coeff' => ( ( 0.5 * $n * $cum ) - $area ) / ( 0.5 * $n * $cum ),
27                 'count' => $cum );
28     }
29     $res->free();
30 }
31 foreach( $gini as $takst_kode => $t ) {
32     $c = $t[ 'coeff' ];
33     $t = $t[ 'count' ];
34     $sql = "INSERT INTO takst_gini ...";
35     $db->real_query( $sql );
36 }
37 $db->close();

```

Gini-koeffisienten for antallet enkelttakster per lege vil bli nær 1 i de tilfellene der noen svært få leger står for det store flertallet av refusjonskravene mot denne taksten. Slik kunne vi identifisere ivrig takstbruk ut over det normale.

Skriptet laget en ny tabell i databasen, hvor koeffisienten for hver av takstene ble lagret. Vi kunne nå kjøre en hovedspørring der vi ba om å få ut takstene med høyest Gini-koeffisient, samt hvor ofte de har vært i bruk.

| Kode | Tittel                                                 | Antall | Gini    |
|------|--------------------------------------------------------|--------|---------|
| 2kd  | Skriving av opplysninger til nasjonal kjernejournal    | 23313  | 0.90780 |
| 708c | Totalkolesterol                                        | 12935  | 0.84620 |
| 708b | Analyse av ALAT og GT                                  | 23505  | 0.84090 |
| 1bk  | Enkel pasientkontakt ved brev eller telefon            | 122810 | 0.80890 |
| 704k | Trichomonas vaginalis i sekret og "clue cells"         | 3897   | 0.80100 |
| 2gd  | Tillegg ved konsultasjon for fornyet vurdering         | 2689   | 0.79400 |
| 705j | Mikroskopisk undersøkelse                              | 2379   | 0.78870 |
| 705i | Skabbmidd eller sopp                                   | 4656   | 0.78280 |
| 2dk  | Tillegg spesialistkompetanse allmennmedisin siste 5 år | 9796   | 0.78230 |
| 708d | Analyse av kreatinin                                   | 24921  | 0.78190 |

**Tabell 1:** Tabellen viser de 10 takstene som noen få leger har brukt svært mye mer enn andre leger.

Slik fikk vi et svært godt utgangspunkt for videre journalistikk. Noe av takstbruken har

en naturlig forklaring, mens bruken av andre takstkoder avslørte iøynefallende mønstre hos enkeltleger.

Listen over påfallende takstkodebruk ble et av utgangspunktene for spissere undersøkelser mot enkeltleger. Som det fremgår av tabell 1, er det flere takster som vekker vår nysgjerrighet.<sup>21</sup> Den innledende researchen rettet seg derfor primært mot å undersøke bruken av enkelttakster.

Dette er langt på vei det samme som Helfos egen kontrollavdeling gjør når de prikker ut sine kontrollobjekter, i hvert fall hvis vi skal tro seksjonssjef Ole Trapness hos Helfo:

– De som ligger høyest i bruken av en takst, er ofte de som er mest interessante for oss å se nærmere på, sier han i et av intervjuene.

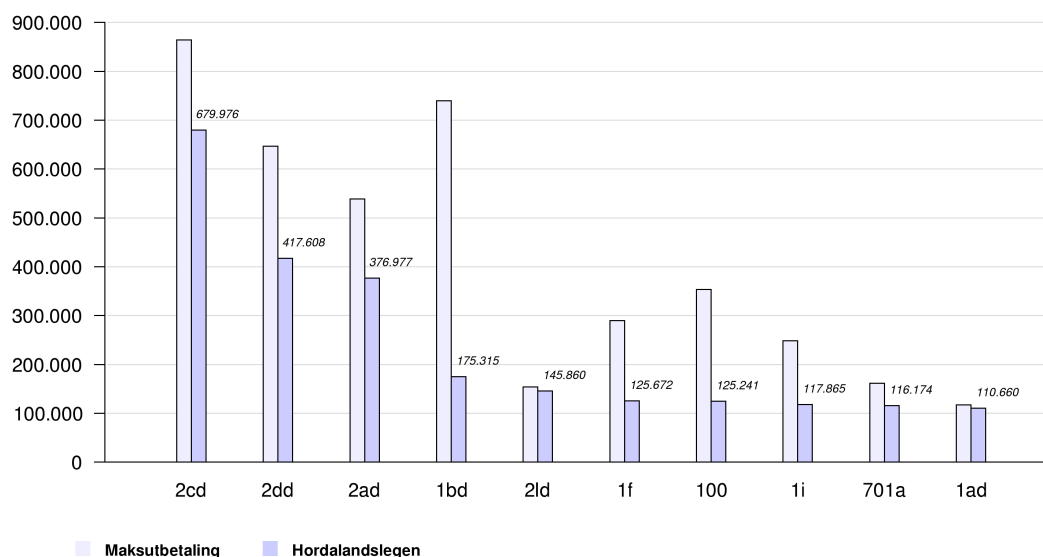
Parallelt med dataundersøkelsene, hadde også tradisjonelt kildearbeid gitt avkastning i form av konkrete tips om leger vi burde ta en nærmere titt på.

En av disse kildene, assisterende kommuneoverlege Ola Jøsendal i Bergen, kjente igjen problemstillingen med mulig takstkode-triksing fra et legekonto på Radøy for snart 20 år siden. Den gangen varslet han myndighetene om mulig misbruk av takstkoder fra en navngitt enkeltlege der, men uten at saken fikk konsekvenser for vedkommende.

Den aktuelle Hordalands-legen nådde ikke opp til topps på noen av våre kjøringer. Men vi skulle etter hvert oppdage at det var en grunn til det.

### 3.7 Kollektivt hakeslipp

Vi hadde satt sammen en samling sorterte lister over legenes bruk av enkelttakster som vi tok med oss til Bergen. Da Jøsendal fikk se rapporten, la ha straks merke til navnet på Hordalands-legen. Men han lå et stykke ned på listene våre, og fantes ikke i toppen av noen rangeringer. Andre leger var alltid “verre enn ham” (figur 10).



**Figur 10:** Hordalands-legen gjemte seg i skyggen av verstingene. Utbetalingene han fikk for enkelttakster (mørkeblå søyler) ble overskygget av andre leger med høyere utbetalinger (lyseblå søyler). Dermed forsvant han nedover på listene våre. Han lå alltid under de høyeste utbetalingene hvis vi kun så på enkelttakster.

<sup>21</sup>Tabellen vi viser her er betydelig kortere enn tabellen vi brukte videre i undersøkelsen. Flere takster ble vurdert som interessante, og vi gjorde undersøkelser knyttet til flere av dem.

Vi hadde fått tunnelsyn på vår ferd gjennom datasettet. Det var først under opptak av klippebilder til filmen at Maria oppdaget hvorfor Hordalands-legen fløy under radaren: Han “smurte utover”.



**Figur 11:** Teamet satte seg ned for å ta noen klippebilder fra databaseresearchen. Det som kun skulle være en fortellermessig rekonstruksjon for TV ble brått dataresearch i realtid. Maria fikk øye på noen kjente navn på veggen, og plutselig falt brikkene på plass.

Vi hadde gjort noen overordnede SQL-spørringer foran kamera, for å illustrere arbeidet med dataene, da Maria med ett får øye på et kjent navn fra en dom i Sarpsborg tingrett fra 2017. Her ble en fastlege fra Østfold dømt til fire års fengsel for grovt bedrageri, etter å ha krevd til sammen 7,3 millioner kroner for mye fra Helfo.<sup>22</sup>

Den dømte legen røk rett opp på topp i vår rangering over refusjonsmottakere for 2015, før anmeldelsen fra Helfo kom. Han forsvant naturlig nok fra datasettet vårt da politiet igangsatte sin etterforskning i 2016.

Men hvem overtok “førsteplassen” etter ham?

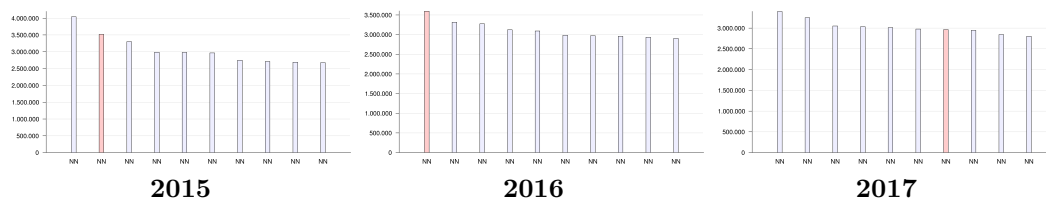
Vi fikk kollektivt hakeslipp da vi la opp en ny kjøring for 2016 på skjermen. Plutselig var Hordalands-legen på topp over alle legene i landet (figur 12). Dette var altså den samme legen som Jøsendal hadde varslet om 20 år tidligere, uten at de den gang ble noen reaksjon.

I stedet for å overdrive bruken av noen bestemte takster, slik enkelte leger gjorde, lå Hordalands-legen generelt *nokså høyt på mange takster samtidig*. Ikke høyt nok til å skille seg ut i søk på enkelttakster, men mer enn nok til å gi ham en suveren førsteplass når vi slo sammen alle takstkodene for 2016 i samme tabell.

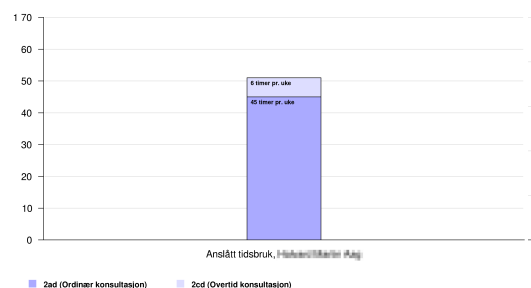
Vi startet dessuten tidlig med å se på bruken av konsultasjons- og overtidstakstene 2ad og 2cd. Var det mulig å finne leger som hadde rapportert høyere tidsbruk enn døgnet har timer, eller lignende? (Figurene 13 og 14)

---

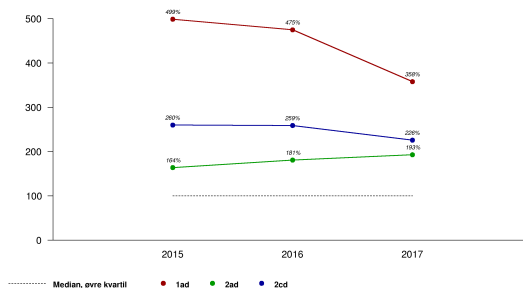
<sup>22</sup>Dommen er ikke rettskraftig. Behandlingen av ankesaken pågår i skrivende stund.



**Figur 12:** Summerte refusjonutbetalinger, de 10 høyeste for hvert år. Hordalands-legen er markert med rød søyle i diagrammene. Han var den fastlegen som mottok mest i refusjon i 2016. Legen som hadde førsteplassen i 2015, ble for øvrig anmeldt av Helfo, og senere dømt for grovt bedrageri. Året etter overtok altså Hordalands-legen førsteplassen etter ham.



**Figur 13:** Vi forsøkte å anslå tidsbruken innenfor normalarbeidsuken ved å se på hvor mange ganger legene brukte en av konsultasjonstakstene. Disse anslagene blir imidlertid veldig omtrentlige.



**Figur 14:** Utviklingen i bruk av ulike konsultasjonstakster for Hordalands-legen, beregnet som prosent av median for de 25% ivrigste brukerne. Den stiplede linjen er medianen.

### 3.8 Tall til folket

Vi hadde tall, kjøringer, lister og en journalistisk analyse, men hadde vi en fortelling? Kravene til å fortelle på TV er høyere enn i noe annet medium. Mest fordi det går fryktelig fort. Men også fordi du må fange seerens oppmerksomhet med levende bilder som treffer i magen, ikke bare hodet. Med så mange data, hadde vi en betydelig pedagogisk utfordring. Vi skulle på den ene siden fortelle om den jobben vi hadde gjort med å analysere tallene og hva dette betydde. På den andre siden skulle dette materialet angå “folk flest”, for å sitere et politisk parti. Vårt beste kort var åpenbart.

Omtrent alle har et forhold til legen sin. Så det var der vi måtte begynne. Og dersom det viste seg at leger misbrukte den tilliten som de hadde fått, både fra storsamfunnet og fra enkeltpersoner, ville denne konflikten være selve nøkkelen for å gripe seerne. Tallene måtte brytes ned til folks hverdag. Altså, det handler om at DU går til legen. Hva skjer der, når det skal betales for moroa? Hvem tjener på deg som pasient? Og foregår det alltid riktig?

Så vi måtte ut på et legekontor, og vi måtte filme der. Vi brukte mye tid på å finne legen som ville stille opp. Og vi måtte ha en pasient som skulle undersøkes for akkurat det vi ønsket, som utløste akkurat den koden og den taksten som passet vår fortelling. Her lå selve clouet: Takst og koder er ikke uforklarlige størrelser, de betyr noe for deg. Samtidig ble det viktig å samarbeid med grafisk avdeling. Hvordan synliggjøre den grådige fjerdedelen? Resultatet ble en enkel, men effektiv grafikk som kunne brukes både på nett og på TV.

Hovedpoenget var å vise at det dreide seg om enorme summer som var gitt i tillitt og at noen få skummet fløten, både på bekostning av fellesskapet og den enkelte pasient. Samme kveld som filmen ble publisert, kunne vi publisere artikkelen “Fastlegekaka blir ulikt fordelt”<sup>23</sup> som et supplement. Her kunne vi gå i dybden på tallene på en annen måte, og

<sup>23</sup>[www.nrk.no/dokumentar/fastlegekaka-blir-ulikt-fordelt-1.14032216](http://www.nrk.no/dokumentar/fastlegekaka-blir-ulikt-fordelt-1.14032216)

samtidig bruke den samme grafikken, noe som gjorde at filmen og nettartikkelen “spilte sammen”.

## 4 Stein til byrden?

Datasettet var i aller høyeste grad offentlig informasjon, siden det beskrev fordelingen av 15 mrd skattekrone. Samtidig var vi trygge på at aggregeringen gjorde det teknisk umulig å assosiere enkeltpasienter med dataene. Legene var på sin side er fullt identifisert med navn og autorisasjonsnummer, fordi de i denne sammenhengen var offisielle personer som forvaltet fellesskapets midler. Men det er stor forskjell på hva pressefolk samler inn av informasjon, og hva vi publiserer. “Vær Varsom”-plakaten har klare publiseringsregler. Presseetikken stopper oss som oftest lenge før jusen. Skulle legene navngis eller ikke? Det var absolutt ting som talte for å bruke navn. Selv om legene vi omtalte hadde ulike historier, var det noen fellestrekk. Vi så et misbruk av takster over tid, det var gitt advarsler fra Helfo gjentatte ganger uten at praksis var endret. Det som skilte dem var jussen. En lege praktiserte fremdeles som fastlege, noe som i seg selv kunne tale for identifisering, en annen hadde en rettskraftig dom. Hos den tredje var saken fortsatt i rettsapparatet. Så ja, noen av de aktuelle legene kunne selvsagt identifiseres. Men var det et stort poeng? Spørsmålet var om identifisering ville skape såpass mye støy at det kom i veien for selve kjernen i historien. Og hvorfor gi motparten anledning til å bråke om navn, slik at dette kunne få saken til å spore av. Dessuten, belastningen på den enkelte lege ville være stor nok når vi snakket med pasienter og andre i lokalmiljøet. De nærmeste visste uansett hvem det var snakk om. Vi trengte ikke legge en riksdekkende stein til byrden. Konklusjon: Faktaene ropte mer enn høyt nok.

## 5 Andre kilder

Data fra 90 millioner regninger er ikke journalistikk, det er rådata. Først når man intervjuer dataene nærmer man seg noe som gir mening. Så det grunnleggende spørsmålet var egentlig dette: Hvordan kunne vi “intervjue” dataene riktig?

SQL spørringene forutsatte en stor mengde bakgrunnskunnskap om fastlegeordningen, diagnoser, politiske føringer, oppgjørssystemer, rettspraksis osv... Det skulle vise seg at den trassige fireåringen” (side 3.2) var det vanskeligste intervjuobjektet vi hadde opplevd.

### 5.1 Koder for det meste

Siden dataene kun opererte med takstkoder og ikke priser eller diagnoser, måtte vi pugge hver av de 169 kodene for å lese dataene effektivt. 1f-telefonsamtale, 2ad-konsultasjon, 105-større kirurgiske inngrep med uttømmende liste som ”strikkbehandling av hemoroider”. Taksthefte viste prisen på hver av disse behandlingene som fastlegen brukte til å fremsette refusjonskrav overfor Helfo og veiledningen definerte hva de ulike takstkodene dekket.<sup>24</sup>

### 5.2 Mangelfulle kontroller

Helfo utførte enkelte manuelle etterkontroller på enkelttakster der de mente det var størst risiko for feilutbetalinger. I vår tidsperiode var det kun utført tre kontroller på to takster med et utvalg på 19 leger som hadde fakturert mest på de aktuelle takstene. Alle var tatt, ingen anmeldt.

---

<sup>24</sup>[normaltariffen.legeforeningen.no/pdf/Fastlegetariff.2016.pdf](https://normaltariffen.legeforeningen.no/pdf/Fastlegetariff.2016.pdf)



Rapportene var publisert på Helfos hjemmesider, men metodekapitlene var avspist med noen setninger. Helfos årsrapporter ga heller ikke flere svar om dette. For å utvikle gode SQL spørringer var det viktig med innsikt i Helfos egne metoder. Kildesamtaler med kontrollører og eieren av databasen i Helsedirektoratet ble derfor viktig.

Kildene viste oss hvor mangelfull kontrollen var, hvordan leger som måtte betale tilbake flere hundre tusen kroner, ikke ble anmeldt. Samtidig ble pasienter fengslet for trygdemisbruk over 100.000 kroner.

Viktigst for å intervju dataene riktig var å identifisere hvilke takster som hadde høy risiko for misbruk. I tillegg forstod vi hvilken profil vi skulle legge inn i dataene som sannsynliggjorde misbruk av takster. For eksempel kunne vi sette sammen antall pasienter på lista, antall konsultasjoner, utbetaling fra Helfo og regne ut hvor mye arbeidstid dette betød. Gikk det opp med antall arbeidstimer i døgnet?

### 5.3 Innsyn i et skjebnefellesskap

Gjennom intervjuer med fastleger og besøk på legekontorer fikk vi innsikt i hverdagen og kunne selv identifisere hvilke takster det var lettest å misbruke. Varslere ga tips om leger de hadde rapportert inn til Helfo og ga et detaljert bilde av hva de hadde sett som fikk dem til å reagere.

Søk i Lovdata viste at det på det aktuelle tidspunktet bare fantes to aktuelle saker i rettsapparatet. Den ene dommen var rettskraftig etter behandlingen i tingretten, den andre ikke. Legen som var blitt dømt i Asker og Bærum Tingrett, hadde ikke anket dommen. Vi brukte mye tid på å snakke både med advokaten hans, ansatte ved legekontoret og andre kollegaer. Men vi hadde få pasienter. Men det fantes en god mulighet; rettsboka. Ved å lese rettsboka, ville vi også få innsyn i vitnelista. Og noen av vitnene ville åpenbart være pasienter. Her støtte vi på en tydelig inkonsekvens i praktisering av innsynsretten i det norske rettssystemet. Da vi ba om innsyn i rettsboka i Asker og Bærum, var det bråstopp. Vi ba om merinnsyn, uten å få gehør. Litt annerledes da vi henvendte oss til tingretten i Sarpsborg hvor den andre saken hadde vært ført. Her fikk vi innsyn, og satt plutselig på nærmere 100 navn på pasienter, kollegaer, etterforskere, aktorat, forsvarere osv. Vi tok oss tid til å kartlegge bosted, telefon og arbeid på de enkelte før vi tok kontakt med dem gjennom en betydelig ringerunde. Til slutt satt vi med det vi det som var målet vårt: Pasienter som var blitt misbrukt av legen slik at han hadde økonomisk gevinst av det. Vi ba de aktuelle pasienten om å kreve innsyn i journalen sin slik at vi kunne se på enkelttakster sammen med dem og sjekke om de stemte med virkeligheten. Underveis oppdaget vi at hypotesen vår stemte: Det var et skjebnefellesskap mellom lege og pasienter. Pasientgruppa til Sarpsborg-legen delte seg tydelig i to. Den ene grupperingen var begeistret for en lege som viste stor omsorg og brukte mye tid på mange undersøkelser av dem. De var mindre opptatt av om han jukset med takstene. Han var jo så grei. Mange av disse var eldre mennesker som hadde opparbeidet et nært forhold til denne legen over lang tid. Den andre gruppa bestod av folk som var forbanna. Kanskje ikke først og fremst på pengejukset, selv om det var alvorlig nok. Det var heller misbruket av tillit som førte til raseriet. Ikke minst gjaldt det misbruk av takster hvor barna deres var involvert. Andre igjen hadde en interessant, overordnet tilnærming. De ville være med å belyse saken fordi de hadde et samfunnsansvar, og var tydelig på at "slikt gjør man ikke med fellesskapets midler".

### 5.4 Ikke grovt bedrageri?

Etter hvert som enkeltleger ble mistenkt på bakgrunn av datakjøringene, måtte vi sjekke og dokumentere avvik hos hver enkelt lege. Avisartikler, sosiale medier, spesielt Facebook der både legene og deres pasienter stort sett befant seg ble brukt til å se hvilke hverdag de levde. Kunne f.eks en takstbruk som tilsa 12 timers arbeidsdag, knapt uten ferie forenes med facebookbilder fra lange Sydenturer? Vi konfronterte hver aktuelle lege og ba dem legge fram sin dokumentasjon på at de ikke hadde misbrukt systemet. Samtalene ble foretatt både på telefon og mail, og ble dokumentert. Noen svar holdt mål, andre ikke og de som ikke svarte mistenkeliggjorde seg selv. Legen som hadde tatt ut mest fra Helfo de siste årene,

måtte vi til slutt oppsøke en morgen på jobb for å få tilsvaret til materialet vårt. Det viste seg at han hadde betalt tilbake tre ganger beløpet som er grensen for ”grovt bedrageri”, men ikke anmeldt og fortsatte som før.

Gjennom kommentarfeltet både på Facebook, men også på [www.legelisten.no](http://www.legelisten.no) og gjennom NRK-kolleger på Marienlyst og distriktskontorer fikk vi tak i pasienter av de mistenkte legene.

De var ikke klar over retten til innsyn i egen pasientregning. Vi utformet et standardbrev som de kunne sende inn til Helfo om de ønsket å sjekke om deres pasientbesøk var korrekt ført. Selv om vi informerte Helfo om brevet, var det et tidkrevende arbeide og vi var bevisste på at uansett svarene det ga, kunne ikke en legeregning som ikke stemte i seg selv dokumentere at legen misbrukte samfunnets tillit.

Noen pasienter ønsket ikke å be om innsyn i legens arbeid på grunn av det tette tillitsforholdet. De som gjorde det fikk svaret etter mange uker. Da var det flere pasienter som ble usikre på hva som egentlig hadde foregått i legetimen og dermed ble det ord-mot ord og lite håndfast materiale å bruke som dokumentasjon. Det var likevel en metodisk inngang vi lærte mye av fordi kildene gjerne stod nær legen og fortalte oss mye om legens hverdag, arbeidsinnsats og etikk.

## 5.5 Deilige regnskaper

Hva tjener egentlige norske fastleger? Dette er et komplisert spørsmål. Skattelistene gir sjelden gode svar. Der finner man skattbar inntekt, altså hvilken sum fastlegene skatter av etter fradrag for utgifter. Og fastleger har mange utgifter, både til kontor, medisinsk materiell og ansatte. Dessuten er svært mange fastleger bare registrert med enkeltmannsforetak (ENK). Her er innsynretten begrenset og du får ikke ut regnskapet for et ENK med mindre innehaveren vil det. Og det skjer bare unntaksvis. Men det finnes leger som har organisert seg med aksjeselskaper, og regnskapene fra disse er tilgjengelige i Brønnøysundregistrene. Så det var bare å sette igang. Vi kjørte topplistene våre mot Brønnøysundregistrene for å se hvor mange av legene med høyest refusjon fra Helfo som hadde aksjeselskaper. Resultatet var for så vidt overraskende. Relativt mange hadde valgt denne organisasjonsformen. Kanskje det var for å minimere risiko? I alle fall er det slik at dersom du går konkurs med et AS, så følger ansvaret selskapet. Med et ENK står du personlig ansvarlig. Regnskaper er deilige. Enda deiligere er notene hvor man både spesifiserer lønnsutgifter, eierforhold og ikke minst utbytte. Det skulle vise seg at flere av de som toppet refusjonsstatistikken, hadde en særdeles offensiv utbyttepolitikk. Over år dreide det seg om uttak av titalls millioner kroner. Noe som igjen var med på å underbygge hypotesen vår. Nemlig at det var en sammenheng mellom svært høy taksbruk og personlig inntekt. Dermed kunne vi også se at motivene til fastlegene ikke bare lå i den Hippokratiske ed, men også i det økonomiske.

## 5.6 Helsetilsynet leverer — mer i vente

Vi begynte å nærme oss valide data. Leger med navn og legeid, adresse og telefonnummer pekte seg ut og vi stilte oss to spørsmål: Dersom vi har rett at disse ikke har rent mel i posen, hvilken etikk har de ellers i legevirket? Hvilke kontroller er gjort på dem tidligere? Gjennom innsynsbegjæringer i Helsetilsynet, samt HELFO/helsedirektoratets sakshistorikk på hver enkelt lege, mottok vi dokumentasjon fra deres legepraksis som kunne være med å bygge opp under eller kontrollere dataene vi hadde fått ut av datasettet. Svarene her gikk rimelig raskt.

Fastleger har varierende antall pasienter på listene sine. Disse listene er heller ikke alltid stabile over tid. Derfor måtte vi komplettere datasettet vårt med informasjon fra *fastlegedatabasen* til Helsedirektoratet. Fersk informasjon om antall pasienter på legenes lister finnes tilgjengelig på nettet, men vi måtte ha de *historiske* tallene for 2015, 2016 og 2017, slik at de kunne holdes sammen med KUHR-dataene fra de samme årene.

I tillegg ba vi om og fikk innsyn i samtlige tilsynsaker som var utført på norske fastleger de tre siste årene. Det viste seg å være et betydelig antall, totalt 352 enkeltsaker. Dette vil gi NRK mange gode nyhetssaker i ukene som kommer.

## 6 Tilbake til kilden

På Marienlyst var det begrenset med kolleger som hadde den datakyndigheten som skulle til for å gi dataene mostand. Uansett, et korrigerende “blikk utenfra” er alltid best. Det var åpenbart hvor vi skulle gå — til kilden — eieren av hele fastlegedatabasen.

Turen gikk til Helsedirektoratet i Oslo sentrum. Vi bragte med oss en kopi av den foreløpige rapporten vi hadde laget. Resultat: Vi hadde ikke misforstått eller feiltolket dataene.

## 7 En passende cv

Vi trengte også et korrigerende blikk fra en lege med kunnskap om takster, gjerne en forsker med innsikt i store datamengder, som også var byråkrat og samfunnsmedisiner. Vi fant det i en og samme person. Ola Jøsendal var ikke bare varsler, men nå også assisterende overlege i Bergen og kunne relatere funnene til dagens virkelighet. Vi møttes i Bergen, presenterte dataene, fikk tips til korreksjoner og nye kjøringer som kunne gi ny kunnskap. Det ble mange samtaler på telefon og mail. Validiteten ble i tillegg testet på fastleger vi stolte på, forskere som tok kontakt da de så “programmtalen” og redaktører i NRK.

### 7.1 Sjekk din lege — eller ikke?

Med tilgang til samtlige norske fastlegers refusjonskrav, ønsket teamet bak Legekoden å lage en offentlig database og publisere denne på nett. Hver enkelt pasient skulle få sjekke sin lege, hvilke takster de benyttet mest og hvilken inntjening de hadde på hver takst. Vi mente dette var vårt mandat som avdekkende journalister. Sammen med en innføring i takstenes innhold, ville en publisering opplyse befolkningen om hvilken viktig rolle, tillit og ansvar fastlegene hadde. Det ble lange diskusjoner med ledere og redaktør. Konklusjon: Vi offentliggjør ikke databasen. Hvilke vurderinger veiet tyngst? Det var Legeforeningen som hadde inngått avtalen med Helfo om direkte oppgjør på vegne av fastlegene. Deres motstand var stor gjennom hele prosessen. De hadde laget en fortelling som de ønsket skulle feste seg. Den handlet om at landet hadde for få fastleger og at de som fantes var overarbeidet, samvittighetsfulle og fikk for dårlig betalt. Vår dokumentasjon om leger som svindlet til seg midler fra det offentlige på pasienters bekostning, ødela denne ”fortellingen”. Gravejournalister trigges som regel av motstand, men redaktøren var redd for at viktig journalistikk skulle forsvinne i en offentlig debatt om metode. Det viktigste argumentet var likevel at pasientene ikke ville ha mulighet til å få hele forklaringen på hvorfor nettopp deres fastlege hadde mye eller lite bruk av ulike takster. Den relativt heftige diskusjonen endte i et slags kompromiss. Vi bestemte oss for å se på de ti fastlegene som fakturerte Helfo for mest penger, og deretter gjøre journalistikk på funnene. Hensikten var igjen den samme: Å synliggjøre at det var noen få som fikk mye, ofte på bekostning av andre legers rykte og renommé. Men det gikk enda mer ut over pasientene. De opplevde at fastlegene sovnet foran dem under konsultasjonen. Vi stod altså overfor to ting. For det første fastleger som jukset med takstene. For det andre, fastleger som jobbet så mye og ble så utslitte, at det gikk ut over pasientsikkerheten. Begge deler var funn vi ikke ville ha sett uten tallene i databasen.

### 7.2 Veien videre

Nei, tilgang til 90 millioner legeregninger er ikke journalistikk. Men det er et fantastisk utgangspunkt for å lage undersøkende, datastøttet journalistisk arbeid. Vi har gått noen blindveier og stanget hodet i noen murvegger. Men etterhvert lærte vi oss å snakke med dataene og intervju dem på den mest formålstjenlige måten. Det viser seg lønnsomt nå, når vi sitter med en tilsvarende innholdsrik database. Med forbedrede “språkferdigheter” tror vi det er mulig å lage enda bedre undersøkende saker på en raskere og mer effektiv måte. Så her er det bare å lytte til Tordenskiolds ord: “Hva fanden nøler I efter”