



# Datainnsamling fra RIS og PACS for rekonstruksjon av stråledoser i CT

**HelsIT 2013**

Tore Sivert Istad, S. Flatabø, E.G. Friberg, K. Kjærheim, A. Liland,  
E. Meen, H.M. Olerud, T. Zhanussova



# Radiologi og strålingsrisiko (1)

- Eksponeringen for ioniserende stråling i medisinsk diagnostikk har vokst betydelig i Norge de siste tre tiårene
- Hovedsakelig på grunn av en hurtig vekst i bruk av CT
- Standard konsensus fra de fleste ekspertpanelene har vært å bruke en lineær modell for sammenhengen mellom pasientens mottatte stråledose og kreftrisiko
- Modellen kalles Linear-No-Threshold (LNT)
- Ifølge modellen vil selv den minste strålingseksponering medføre enn viss kreftrisiko

# Radiologi og strålingsrisiko (2)

- I 2007 estimerte Brenner et al<sup>1</sup> at opptil 2% av fremtidige krefttilfeller i USA vil være forårsaket av medisinsk bildediagnostikk
- Estimateret var basert på den statistiske LNT-modellen BEIR VII<sup>2</sup> for dose-kreftrisiko

1 Brenner DJ, Hall EJ. Computed tomography: an increasing source of radiation exposure. N Engl J Med 2007

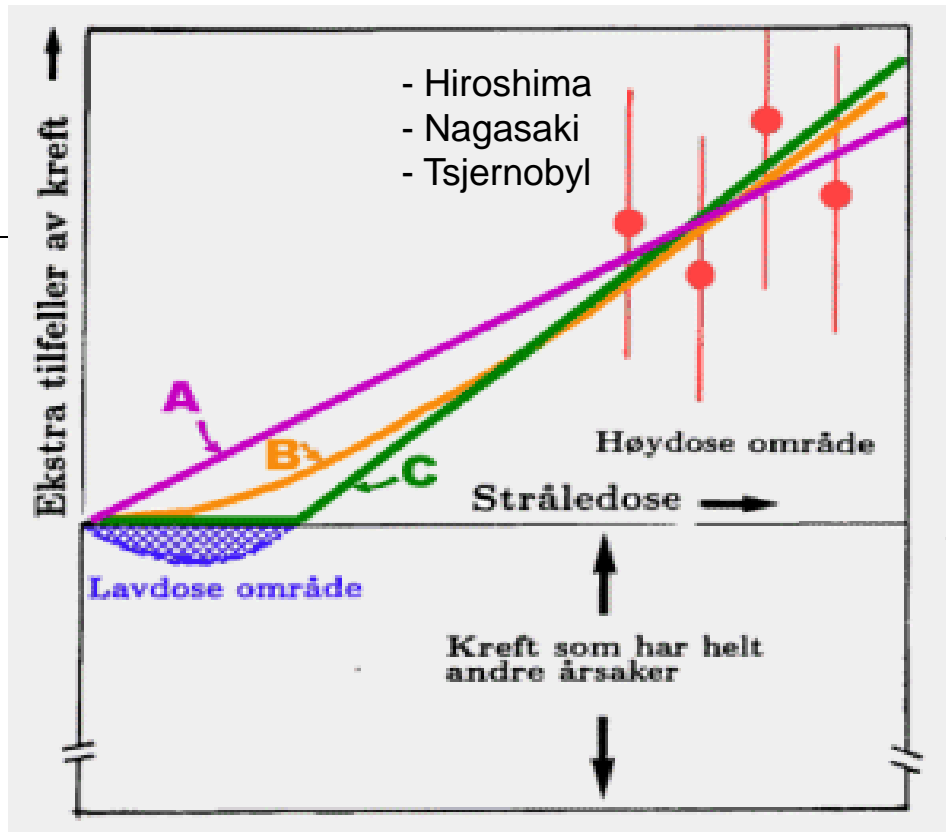
2 National Academy of Science's Committee on the Biologic Effects of Ionizing Radiation (BEIR VII)



<http://rileychildrenshospital.com/parents-and-patients/programs-and-services/radiology/ct-scan.jsp>

# Radiologi og strålingsrisiko (3)

- Problem: LNT-modellen er essensielt en ekstrapolering av data fra Hiroshima, Nagasaki og Tsjernobyl, hvor stråledosene var betydelig større og bestrålingen annerledes enn i CT
- LNT-modellen for gjentatte små stråledoser som i CT er dårlig belagt med empiriske data
- LNT-modellen er «vårt beste estimat» i mangel av empiriske data
- Derfor er det et behov for epidemiologiske studier som kan understøtte avveiningene mellom nytte og risiko ved bruk av radiologiske undersøkelser



Formen på kurven er viktig for nytte/risiko-avveining og optimalisering av CT-undersøkelser

### Modeller for dose-risiko\*:

- (A) Lineær uten terskel - LNT (brukt av Strålevernet og ICRP)
- (C) Lineær med terskel
- (B) Lineær/kvadratisk
- BLÅ+(C) Lineær/kvadratisk med hormese

# Hva er EPI-CT?

- «Epidemiologisk studie av risiko knyttet til CT av barn og optimalisering av CT»
- EPI-CT er en internasjonal helsestudie hvor målet er å kvantifisere risiko ved CT-undersøkelser av barn og unge
- Epidemiologisk metode
- Sekundært formål: Bidra til optimalisering
  - Internasjonale faglige retningslinjer for optimalisering av CT
  - Nye nasjonale retningslinjer for Norge

# EU-prosjekt

- Ledes av Det internasjonale byrået for kreftforskning (WHO)
- Deltakere: 18 institusjoner fra 11 land
- Norske partnere: Statens strålevern og Kreftregisteret, i samarbeid med norske helseforetak
- Finansiering: Norges forskningsråd / EU



<http://epi-ct.iarc.fr/index.php>



# Metode (1)

- Kohort: 1,1 millioner barn og unge
  - forventet norsk kohort: 60 000 - 80 000 pasienter
- Samle individuelle data om CT-undersøkelser fra sykehusenes databaser
  - RIS (radiologisk journalsystem)
  - PACS (radiologisk bilde-database)
- Nøyaktige individuelle organdose-beregninger basert på lagrede scan-parametre fra PACS (røntgenspektrum, eksponering, pitch, FOV, scan-lengde osv.)
- Automatisk gjenkjenning av hvilke organer som ble scannet fra CT-bildene gir mer nøyaktig estimat av organdoser i og utenfor scannet volum



# Metode (2)

- Doseberegninger med state-of-the-art Monte Carlo-simuleringer på 2-300 hybride matematiske kroppslige fantomer (hybride splines-fantomer) (National Cancer Institute USA og Univ. of Florida)



Voxel phantom series B created at the University of Florida.

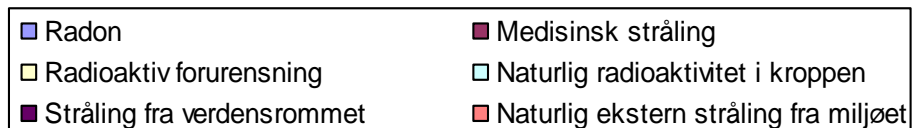
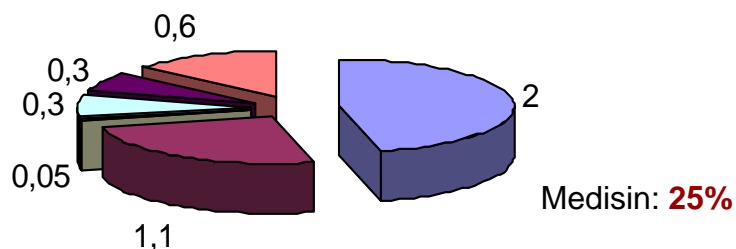
- Organ-dose beregnes for alle strålefølsomme organer
- Følge pasientene over tid, samle data om kreftforekomst, dødelighet, confounders (feks. sosioøkonomiske faktorer, indikasjon)
- Oppfølging første fase: 5 år, leukemi
- Forhåpning: etablert kohort følges opp på livstid, epidemiologi på et stort antall kreftformer (avh. av finansiering)

# Radiologi i Norge per 2008

- 4,3 millioner undersøkelser
  - ca. 1 undersøkelse per innbygger per år
- CT står for 80% av befolkningsdosen fra radiologi
- Typiske pasientdoser: 1-100 mSv
- Radiologi gir det største menneskeskapte bidraget til befolkningsdosen



Årlig dose per innbygger (mSv) - 2008



# Hvorfor barn og unge?



- Strålefølsomme individer
  - Høyere aktivitet av celledeling
- Lengre forventet levetid
  - Overlever latenstiden
- Enkeltindivider som får mange undersøkelser
  - Høye akkumulerte doser

<http://rileychildrenshospital.com/parents-and-patients/programs-and-services/radiology/ct-scan.jsp>

# Motivasjon for deltagelse

- Norge del av en stor multinasjonal pediatrik kohort
  - Økt statistisk styrke
- Norske fagmiljøer får tilgang til verktøy og databaser
- Nasjonal database: Mulighet for utarbeidelse av nasjonale DRLs og faglige anbefalinger, trendstudier
- Økt fokus på optimalisering innen pediatrik CT
- Utgangspunkt for nasjonale felles prosjekter

# Datainnsamling

Datainnsamlingen skjer i to steg:

1. **Datainnsamling fra RIS**
2. **Datainnsamling fra PACS**



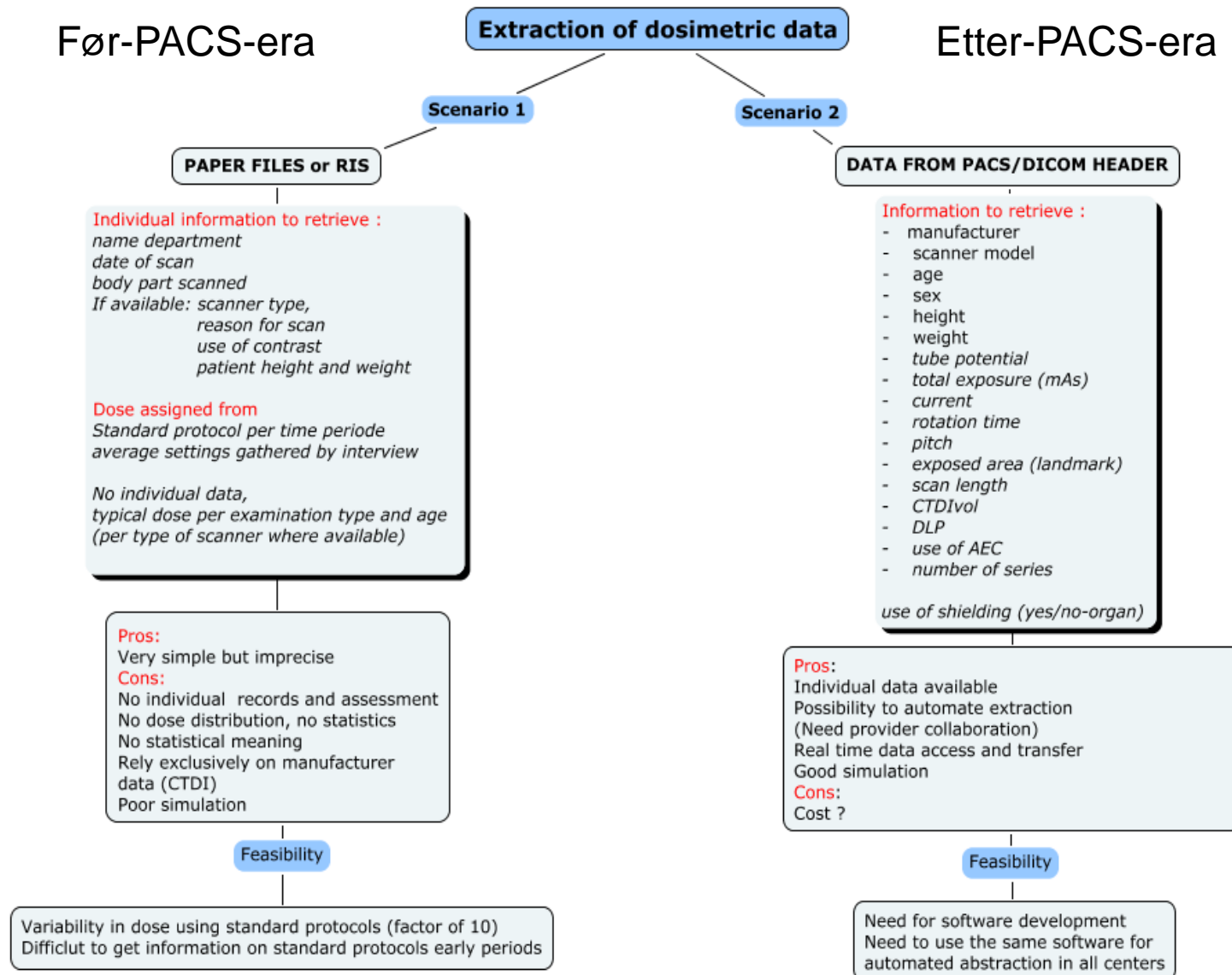
# Datainnsamling fra RIS

- Søk i RIS etter de pasientene som ønskes inkludert i studien
- Inklusjonskriterium: alle pasienter som har fått en CT-undersøkelse mens de var i alderen 0-20 år
- Henter ut en rekke variable om CT-undersøkelsen og pasienten, inkludert personidentifiserende data
- Fra RIS-søket lages også en ren tekst-fil med pasient-IDer — denne brukes senere til dataauthenting fra PACS

# Datainnsamling fra PACS

- Fra PACS henter vi ut alle relevante data som finnes i DICOM-header, bortsett fra identifiserende info
- Identifiserende info blir istedet mappet til et pasient-unikt pseudonym
- I tillegg lages konturer fra bildene for bestemmelse av hvilke organer som ble eksponert for strålefeltet
- Automatisert bildesegmentering og organ-gjenkjenning

# Datasettet vil bestå av to subsett







# Datasikkerhet og driftssikkerhet

- EPI-CT-prosjektet gir høyeste prioritet til:
  - datasikkerhet for sykehuset
  - konfidensialitet for pasientdata
  - at PACS-drift og klinisk virksomhet ikke på noen måte påvirkes

# Løsninger for sikkerhet

- PerMoS kjøres på innsiden av sikkert nettverk, ingen kommunikasjon ut av sykehuset
- IT-admin og PACS-admin har full kontroll på programvaren og på data som forlater sykehuset\*
- Velutprøvd programvare som bruker DICOM-standarden
- Radiologisk avdeling bestemmer riktige tidspunkter med lav trafikk
- Radiologisk avdeling bestemmer riktig query-frekvens
- Testkjøring først med konservative innstillinger
- La PerMoS jobbe «sakte men sikkert» over tid
- Av-identifisering ihht veletablert standard, DICOM PS 3.15-2011 Annex E
- Validering av av-identifisering før data forlater sykehuset



# Utfordring: av-identifisering av DICOM metadata

- DICOM-standarden er kompleks
- DICOM-objekter kan inneholde opptil ~ 100-200 metadatafelter
- DICOM-objekter kan inneholde opptil 20 forskjellige unike identifikatorer (UIDer) definert i DICOM
- UIDene er globalt unike identifikatorer som entydig identifiserer pasienter, CT-maskiner, CT-undersøkelser og enkeltbilder
- UIDene kan sees på som databasenøkler
- Ikke ønskelig å ha databasenøkler fra kilde-databasen i av-identifiserte forskningsdata, hverken mhp. etablerte prinsipper for av-identifisering eller lovens krav
- "Private tags" - CT-leverandørens egendefinerte datafelter
- Indentifiserende info innbrent i bildepixler
- Det enkelte PACS kan ha sine lokale særegenheter
- En CT-undersøkelse vil typisk bestå av ~ 100 - 1000 DICOM-objekter

# Løsninger for av-identifisering av DICOM metadata

- Automatisering
- Mulighet i Permos for å spesifisere skreddersydd av-identifisering etter lokale behov
- Alle potensielt problematiske UIDer slettes eller erstattes med pseudonymer
  - UIDene må erstattes med pseudonymer som oppfyller DICOM-standardens krav og bibeholder relasjonene i DICOM-datastrukturen («consistent within a set of instances»)
- DICOM-standarden PS 3.15-2011, Annex E, er en veletablert, gjennomarbeidet standard som spesifiserer hvilke tags som må håndteres og hvordan
- Scanning med tredjeparts "tag sniffer" brukes til å validere av-identifiseringen før dataene forlater sykehuset

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
	Attribute Name	Tag	Retired	In Std.	Basic Profile	Retain Safe Private Option	Retain UIDs	Retain Device	Retain Patient Chars. Option	Retain Long. Full Dates	Retain Long. Modif. Dates	Clean Desc.	Clean Struct.	Clean Gr.
2	Accession Number	(0008,0050)	N	Y	Z									
3	Acquisition Comments	(0018,4000)	Y	N	X							C		
4	Acquisition Context Sequence	(0040,0555)	N	Y	X								C	
5	Acquisition Date	(0008,0022)	N	Y	X/Z					K	C			
6	Acquisition DateTime	(0008,002A)	N	Y	X/D					K	C			
7	Acquisition Device Processing Description	(0018,1400)	N	Y	X/D							C		
8	Acquisition Protocol Description	(0018,9424)	N	Y	X							C		
9	Acquisition Time	(0008,0032)	N	Y	X/Z					K	C			
10	Actual Human Performers Sequence	(0040,4035)	N	N	X									
11	Additional Patient's History	(0010,21B0)	N	Y	X							C		
12	Admission ID	(0038,0010)	N	Y	X									
13	Admitting Date	(0038,0020)	N	N	X					K	C			
14	Admitting Diagnoses Code Sequence	(0008,1084)	N	Y	X							C		
15	Admitting Diagnoses Description	(0008,1080)	N	Y	X							C		
16	Admitting Time	(0038,0021)	N	N	X					K	C			
17	Affected SOP Instance UID	(0000,1000)	N	N	X		K							
18	Allergies	(0010,2110)	N	N	X				C			C		
19	Arbitrary	(4000,0010)	Y	N	X									
20	Author Observer Sequence	(0040,A078)	N	Y	X									
21	Branch of Service	(0010,1081)	N	N	X									
22	Cassette ID	(0018,1007)	N	Y	X			K						
23	Comments on Performed Procedure Step	(0040,0280)	N	Y	X							C		
24	Concatenation UID	(0020,9161)	N	Y	U		K							
25	Confidentiality Constraint on Patient Data De	(0040,3001)	N	N	X									
26	Content Creator's Identification Code Sequen	(0070,0086)	N	Y	X									
27	Content Creator's Name	(0070,0084)	N	Y	Z									
28	Content Date	(0008,0023)	N	Y	Z/D					K	C			
29	Content Sequence	(0040,A730)	N	Y	X								C	
30	Content Time	(0008,0033)	N	Y	Z/D					K	C			

Table E.1-1 Safe Private Elements

Ready

100%



Confidentiality Profile as a de-identifier shall protect or retain all instances of the Attributes listed in Table E.1-1, whether contained in the main dataset or embedded in an Item of a Sequence of Items. The following action codes are used in the table:

- D – replace with a non-zero length value that may be a dummy value and consistent with the VR
- Z – replace with a zero length value, or a non-zero length value that may be a dummy value and consistent with the VR
- X – remove
- K – keep (unchanged for non-sequence attributes, cleaned for sequences)
- C – clean, that is replace with values of similar meaning known not to contain identifying information and consistent with the VR
- U – replace with a non-zero length UID that is internally consistent within a set of Instances
- Z/D – Z unless D is required to maintain IOD conformance (Type 2 versus Type 1)
- X/Z – X unless Z is required to maintain IOD conformance (Type 3 versus Type 2)
- X/D – X unless D is required to maintain IOD conformance (Type 3 versus Type 1)
- X/Z/D – X unless Z or D is required to maintain IOD conformance (Type 3 versus Type 2 versus Type 1)
- X/Z/U\* - X unless Z or replacement of contained instance UIDs (U) is required to maintain IOD conformance (Type 3 versus Type 2 versus Type 1 sequences containing UID references)

These action codes are applicable to both Sequence and non-Sequence attributes; in the case of Sequences, the action is applicable to the Sequence and all of its contents. Cleaning a sequence (“C” action) may entail either changing values of attributes within that Sequence when the meaning of the Sequence within the context of its use in the IOD is understood, or recursively applying the profile rules to each Dataset in each Item of the Sequence. Keeping a Sequence (“K” action) requires recursively applying the profile rules to each Dataset in each Item of the Sequence (for example, in order to remap

# Ressurser for av-identifisering

- NCI Cancer Imaging Archive DICOM de-identification knowledgebase:  
<https://wiki.cancerimagingarchive.net/display/Public/De-identification+Knowledge+Base;jsessionid=BC436B4E3728E12F7070D617672F4ADF>
- Excel-ark som spesifiserer av-ident.-profiler ihht PS 3.15 Annex E:  
[https://wiki.cancerimagingarchive.net/download/attachments/3539047/PS-3.15-AnnexE-2011\\_ClinicalTrials\\_De-Identification.xlsx?version=1&modificationDate=1357921781800](https://wiki.cancerimagingarchive.net/download/attachments/3539047/PS-3.15-AnnexE-2011_ClinicalTrials_De-Identification.xlsx?version=1&modificationDate=1357921781800)
- DICOM-standard for av-identifisering - PS 3.15-2011 Annex E:  
[http://medical.nema.org/Dicom/2011/11\\_15pu.pdf](http://medical.nema.org/Dicom/2011/11_15pu.pdf)



# Respons fra norske helseforetak

- Vi har fått svært positiv respons fra norske helseforetak
- 12 av 20 norske HF har svart positivt til deltagelse til nå
- Alle fire RHF representert



# Resultater RIS datainnsamling (1)

- Til nå har vi mottatt RIS-data fra:
  - Helse Midt-Norge RHF (Levanger, Namsos og Volda)
  - Universitetssykehuset Nord-Norge HF
  - Stavanger Universitetssykehus HF
  - Sørlandet sykehus HF
  - OUS HF Ullevål sykehus
  - OUS HF Radiumhospitalet
  - Helse Bergen HF
- RIS-data på gang:
  - Resterende Helse Midt-Norge RHF
  - OUS HF Rikshospitalet

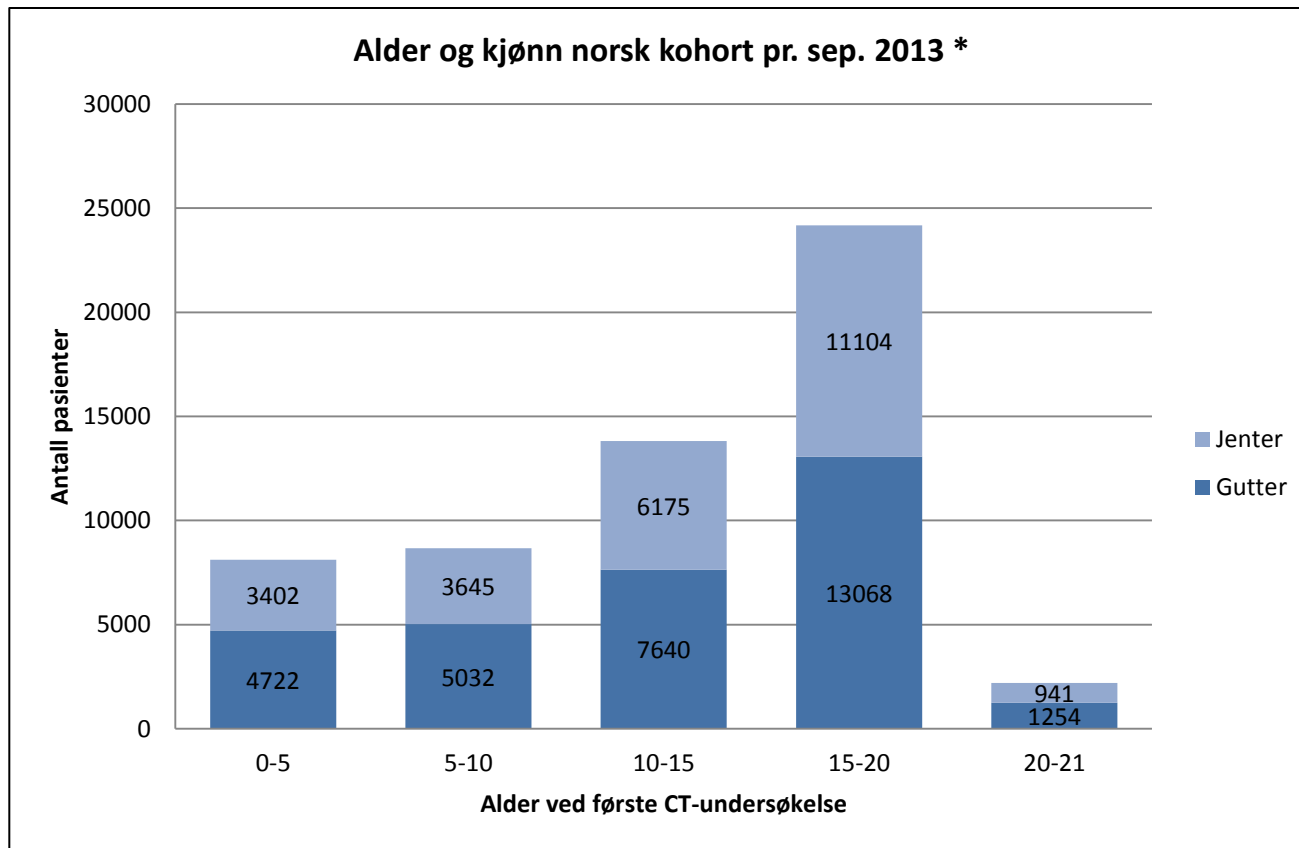
Utfordringer vi har møtt: «Ingen RIS er like», forskjellige felter og feltnavn, inkonsistente data, feil i import av gamle HIS-data, NORAKO-koding, innhenting av koblingsnøkler mellom RIS/PACS, innhenting av undersøkelser gjort etter fylte 21 år blant personer som ble undersøkt som barn/unge

# Resultater RIS datainnsamling (2)

## Norsk kohort pr. sep. 2013 \*

RHF	HFs / Hospitals	Unique patients	Total CT-scans	Age range	Time period
Helse Midt	Levanger, Namsos, Volda	5126	7845	0-20	1999-2011
Helse Nord	UNN HF / Harstad, Tromsø, Narvik	12628	32330	0-20	1980-2011
Helse Sør-Øst	Sørlandet sykehus HF / Kristiansand, Arendal, Flekkefj.	11196	22899	0-19	1992-2011
Helse Sør-Øst	OUS HF Ullevål	8749	15396	0-20	2005-2012
Helse Sør-Øst	OUS HF Radium	982	6000	0-20	1989-2011
Helse Vest	Stavanger Universitetssykeh. HF	5118	7936	0-20	1994-2012
Helse Vest	Bergen HF	13184	21149	0-20	1990-2012
<b>Total</b>		<b>56983</b>	<b>113555</b>		

# Eksempel på ekstraherte RIS-data (1)



\*) Kohort under oppbygging / estimerer

# Eksempel på ekstraherte RIS-data (2)

Antall pasienter etter alder ved første us. og årstall for første us. \*

	<2000	2000-2004	2005-2009	2010+
< 5 years at first scan	2067	1918	3055	1076
5-10 years at first scan	1883	2070	3380	1253
10-15 years at first scan	2296	3002	5672	2729
15-20 years at first scan	3054	4686	10913	5512
20-21 years at first scan	392	623	1033	369
<b>Total</b>	9692	12299	24053	10939

\*) Kohort under oppbygging / estimerer

# Erfaringer PACS datainnsamling

- Første installasjon av PerMoS har vært testkjørt i Helse Midt-Norge i 2,5 mnd.
- Ingen merkbar påvirkning på PACS-drift eller nettverksytelse
- Sakte men sikker strøm av data
- Data er ikke tatt ut av sykehuset foreløpig

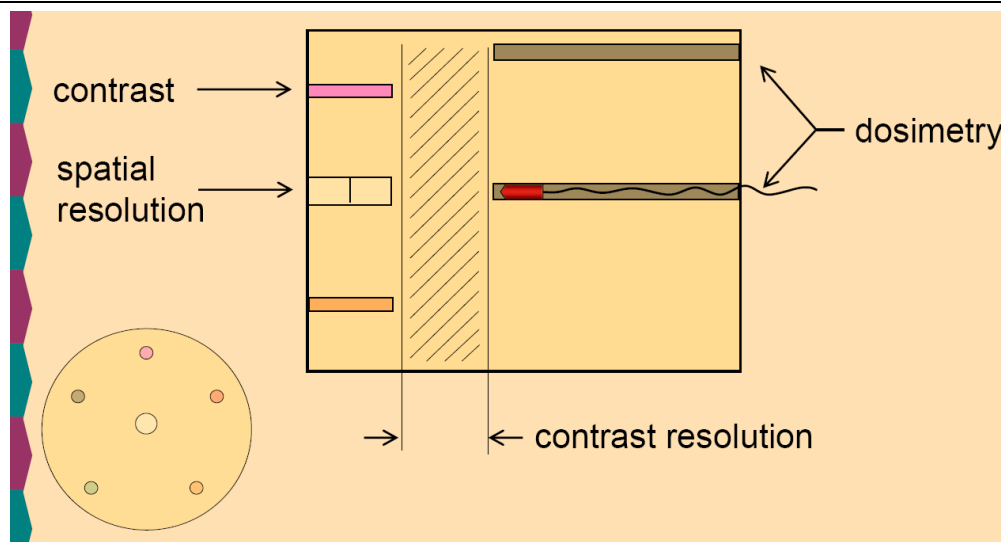
# Erfaringer PACS datainnsamling

- Å få endelig «ja» til installasjon av PerMoS i HFet og hos IKT-enhet kan ta svært lang tid
- Data-innsamling ble satt på vent etter at vår egen risiko-analyse viste behov for bedre av-identifisering
- Ny versjon av PerMoS utviklet med utvidete av-identifiserings-funksjoner
- Ny testkjøring igang i Helse Midt-Norge nå
- Neste planlagte installasjoner: Sørlandet sykehus HF, Stavanger Universitetssykehus, more to follow...



# Optimalisering – Innspill fra Norge

Samarbeidspartner: Tudor Research Centre, Luxemburg



- Nytt **ICRU fantom** for CT som evaluerer både
  - bildekvalitet (CNR, MTF)
  - dose (z-sensitivitetsprofil)
- Unngår problemene med  $CTDI_{vol}$

## Målsetning:

- Få produsert fantomet for bruk i EPI-CT
  - Videre utvikling av fantomet for pediatrisk bruk
- Skaffe/utvikle programvare for automatisk evaluering av bildekvalitet og dose

## Metode:

- Skanne fantomet, pediatriske protokoller, ulike CT-modeller
- Sammenligne resultat med studien som evaluerer kliniske bilder (sub-gruppen)
  - Kjørt med samme protokoll



# Den norske prosjektgruppen

- Eva Godske Friberg<sup>2</sup>
- Silje Flatabø<sup>2</sup>
- Tore Sivert Istad<sup>2</sup>
- Kristina Kjærheim<sup>1</sup>
- Astrid Liland<sup>2</sup>
- Eivind Meen<sup>1</sup>
- Hilde Merete Olerud<sup>2</sup>
- Tore Tynes<sup>3</sup>
- Tamara Zhanussova<sup>2</sup>

1. Kreftregisteret
2. Statens strålevern
3. Statens arbeidsmiljøinstitutt

# WP3/WP4 Data collection & Dosimetry working groups

I. Thierry-Chef<sup>1</sup>, C. Lee<sup>2</sup>, A. Jahnen<sup>3</sup>, J. Dabin<sup>4</sup>, E.G. Friberg<sup>5</sup>, J. Hermen<sup>3</sup>, T.S. Istad<sup>5</sup>, L. Krille<sup>6</sup>, C. Maccia<sup>7</sup>, A. Nordenskjöld<sup>8</sup>, H. Olerud<sup>5</sup>, J.-L. Rehel<sup>9</sup>, L. Struelens<sup>4</sup> and A. Kesminiene<sup>1</sup>

1. International Agency for Research on Cancer, France
2. Division of Cancer Epidemiology and Genetics, National Cancer Institute, National Institutes of Health, US
3. Centre de Recherche Public Henri Tudor, Luxembourg
4. Studiecentrum voor Kernenergie – Centre d'Etude de l'Energie Nucléaire, Belgium
5. Norwegian Radiation Protection Authority, Norway
6. Universitätsmedizin der Johannes Gutenberg-Universität, Germany
7. Centre d'Assurance de qualité des Applications Technologiques dans le domaine de la Santé, France
8. Karolinska Institutet Karolinska University Hospital, Stockholm, Sweden
9. Institut de Radioprotection et Sûreté Nucléaire, France

# En spesiell takk til

- **Mari Gårseth**, Helse Nord-Trøndelag HF
- **Bjørn Kothe Næss**, Helse Nord-Trøndelag HF
- **Steinar Tveiten**, Sørlandet sykehus HF
- **Heidi Andersen**, Stavanger Universitetssykehus HF
- **Åshild Sætevik**, Stavanger Universitetssykehus HF
- ... og medhjelpere på SUS, UNN, OUS, Helse Bergen

# Konklusjon

- CT er et uvurderlig diagnostisk verktøy, men kunnskapen om risikoen fra strålingseksponeringen ved CT er mangelfull
- EPI-CT-prosjektet vil gi et nytt datasett med en nøyaktighet i doseestimerer og en kohort-størrelse som langt overgår tidligere studier
- EPI-CT vil kunne gi bedre kunnskap om sammenhengen mellom CT-undersøkelser og kreftrisiko
- Viktig for at CT skal kunne brukes videre på en berettiget og optimalisert måte
- Vi er godt igang med RIS-datainnsamling i den norske delen av prosjektet
- PACS-datainnsamling fortsatt i startfase