universität freiburg

Semantic Querying of DICOM Structured Reports using the QLever Graph Database

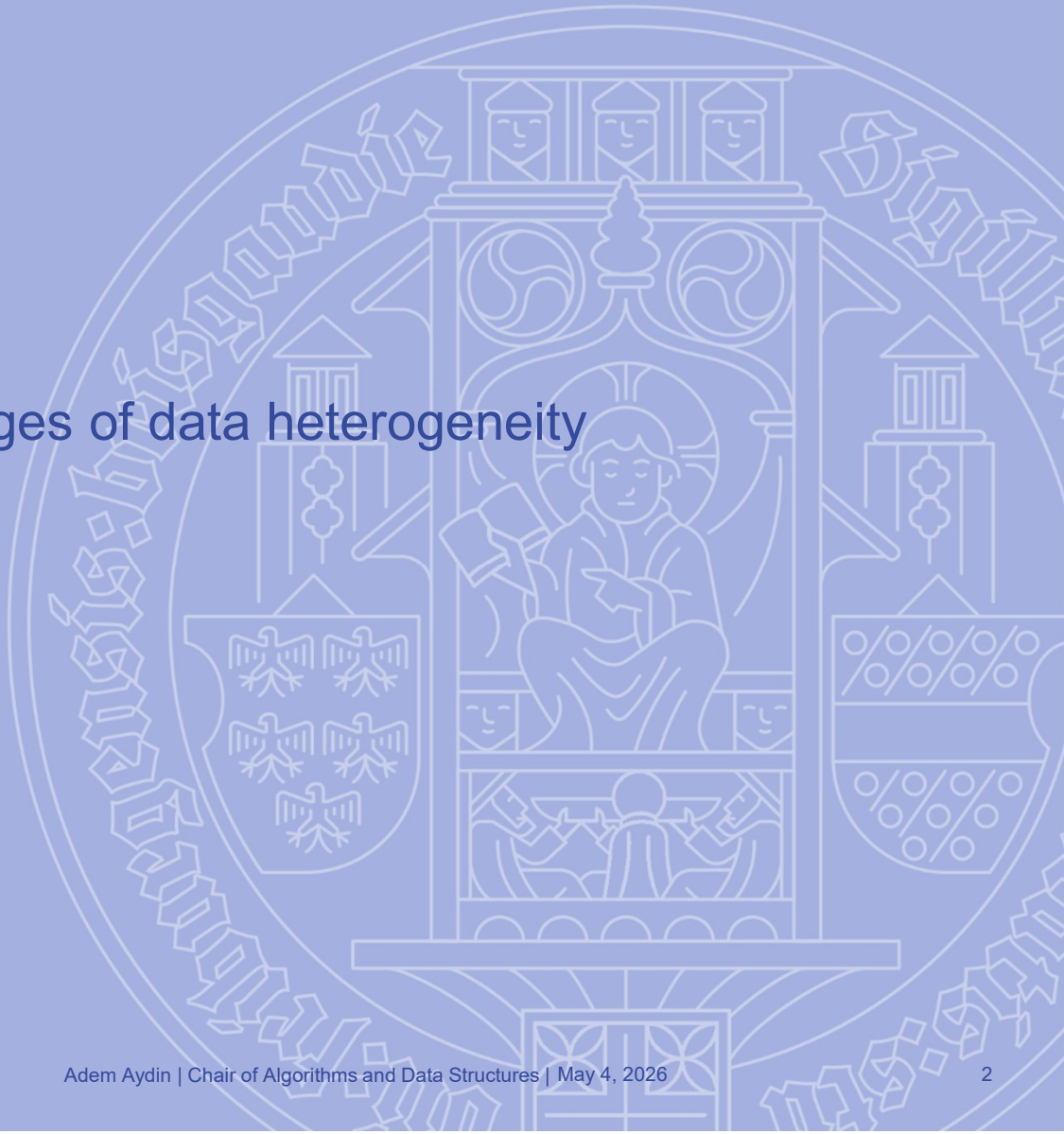
Bachelor's Thesis

Adem Aydin

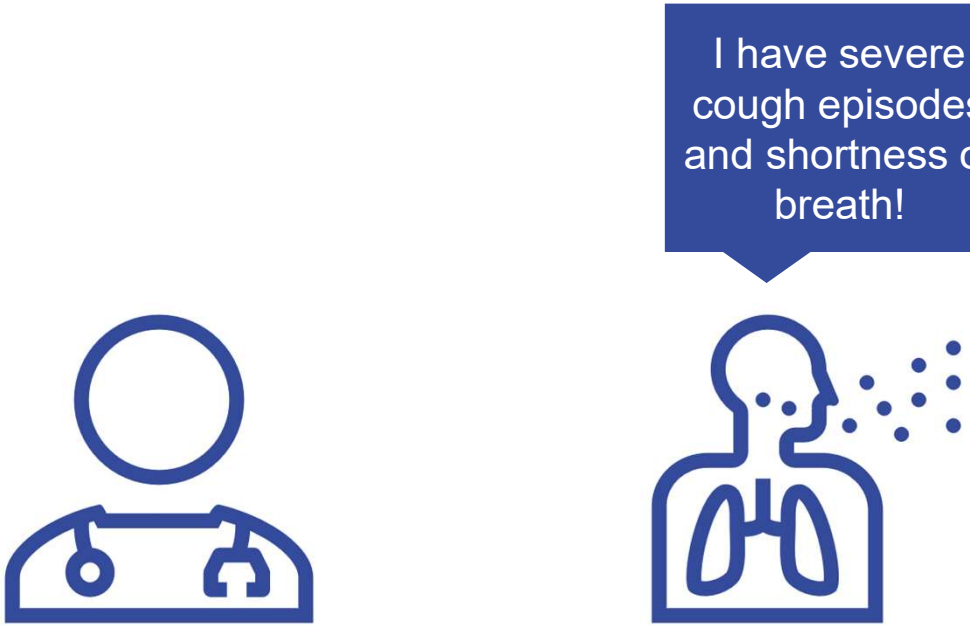
April 27, 2026

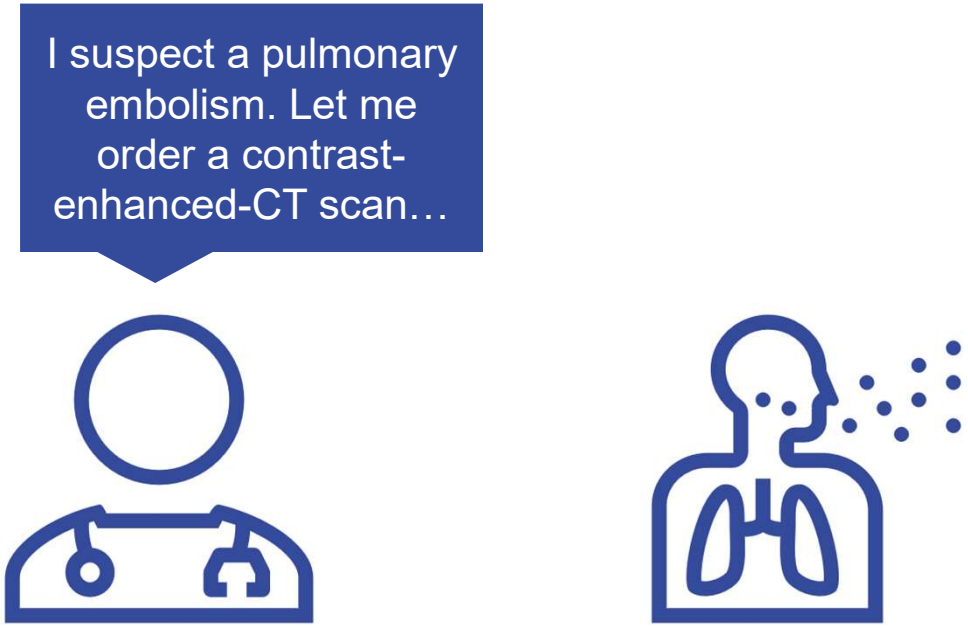
Problem

Patients, radiation doses, and the challenges of data heterogeneity



Imagine the following scenario…


I have severe cough episodes and shortness of breath!


I suspect a pulmonary embolism. Let me order a contrast-enhanced-CT scan…

CT Scanning Process

Progress

Measurement	Value
$CTDI_{vol}$	…
Scan Length	…
Dose Length Product	…
⋮	⋮

CT Scanning Process

Phase: Topogram (non-diagnostic)

Progress

Measurement	Value
$CTDI_{vol}$	$1.5\ mGy$
Scan Length	$50\ cm$
Dose Length Product	$7.5\ mGy.cm$
⋮	⋮

CT Scanning Process

Phase: PreMonitoring (non-diagnostic)

Progress

Measurement	Value
$CTDI_{vol}$	$3\ mGy$
Scan Length	$1\ cm$
Dose Length Product	$3\ mGy.cm$
⋮	⋮

CT Scanning Process

Phase: Monitoring (non-diagnostic)

Progress

Measurement	Value
$CTDI_{vol}$	$3\ mGy$
Scan Length	$10\ cm$
Dose Length Product	$30\ mGy.cm$
⋮	⋮

CT Scanning Process

Phase: Lung (Diagnostic)

Progress

Measurement	Value
$CTDI_{vol}$	$11.2\ mGy$
Scan Length	$10\ cm$
Dose Length Product	$150\ mGy.cm$
⋮	⋮

CT Scanning Process

Phase: Lung (Diagnostic)

Progress

Measurement	Value
$CTDI_{vol}$	$12.8\ mGy$
Scan Length	$20\ cm$
Dose Length Product	$240\ mGy.cm$
⋮	⋮

CT Scanning Process

Phase: Lung (Diagnostic)

Progress

Measurement	Value
$CTDI_{vol}$	$10.5\ mGy$
Scan Length	$30\ cm$
Dose Length Product	$350\ mGy.cm$
⋮	⋮

CT Scanning Process

Phase: Lung (Diagnostic)

Progress

Measurement	Value
$CTDI_{vol}$	$12\ mGy$
Scan Length	$40\ cm$
Dose Length Product	$470\ mGy.cm$
⋮	⋮

CT Scanning Process

Phase: Lung (Diagnostic)

Progress

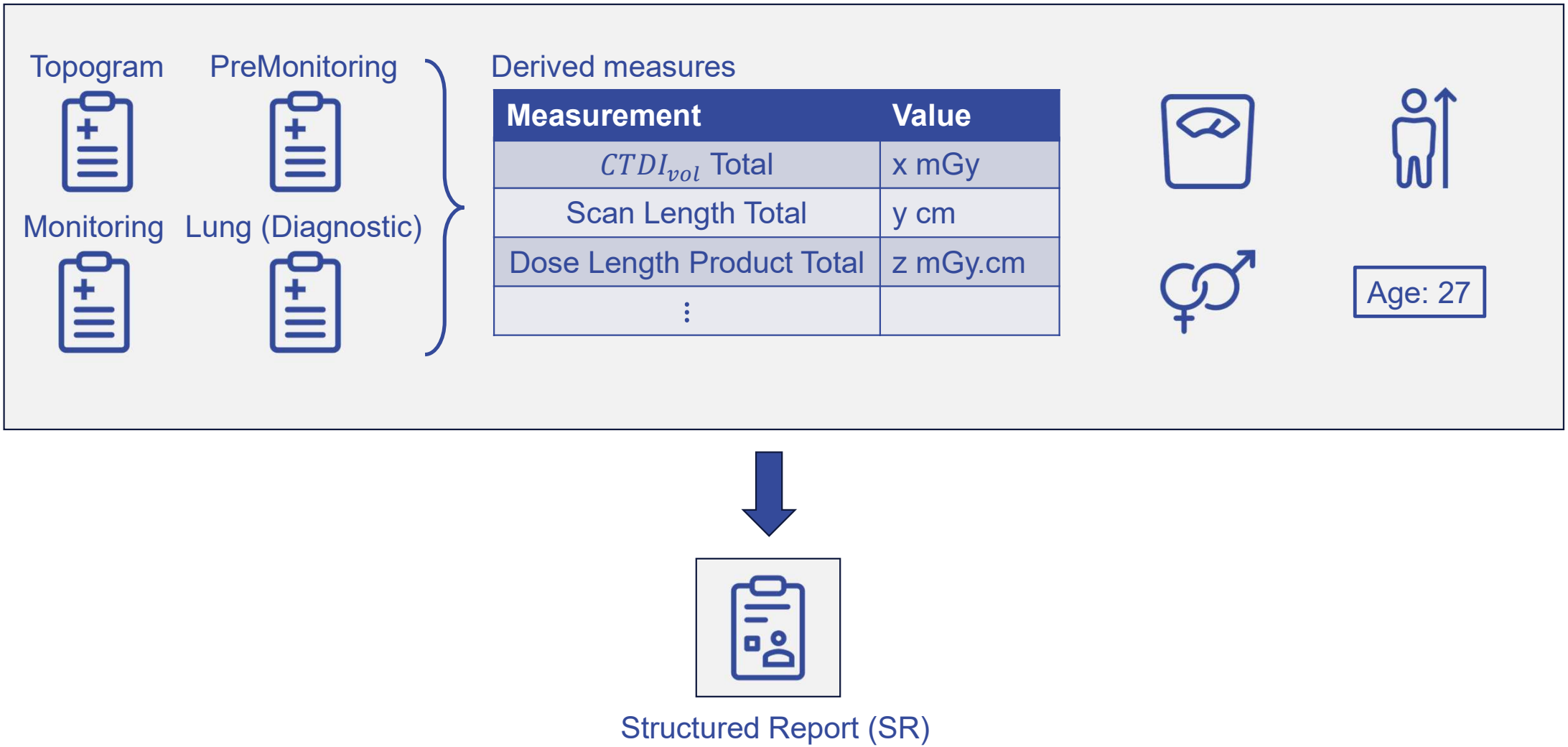
Measurement	Value
$CTDI_{vol}$	$12\ mGy$
Scan Length	$40\ cm$
Dose Length Product	$470\ mGy.cm$
⋮	⋮

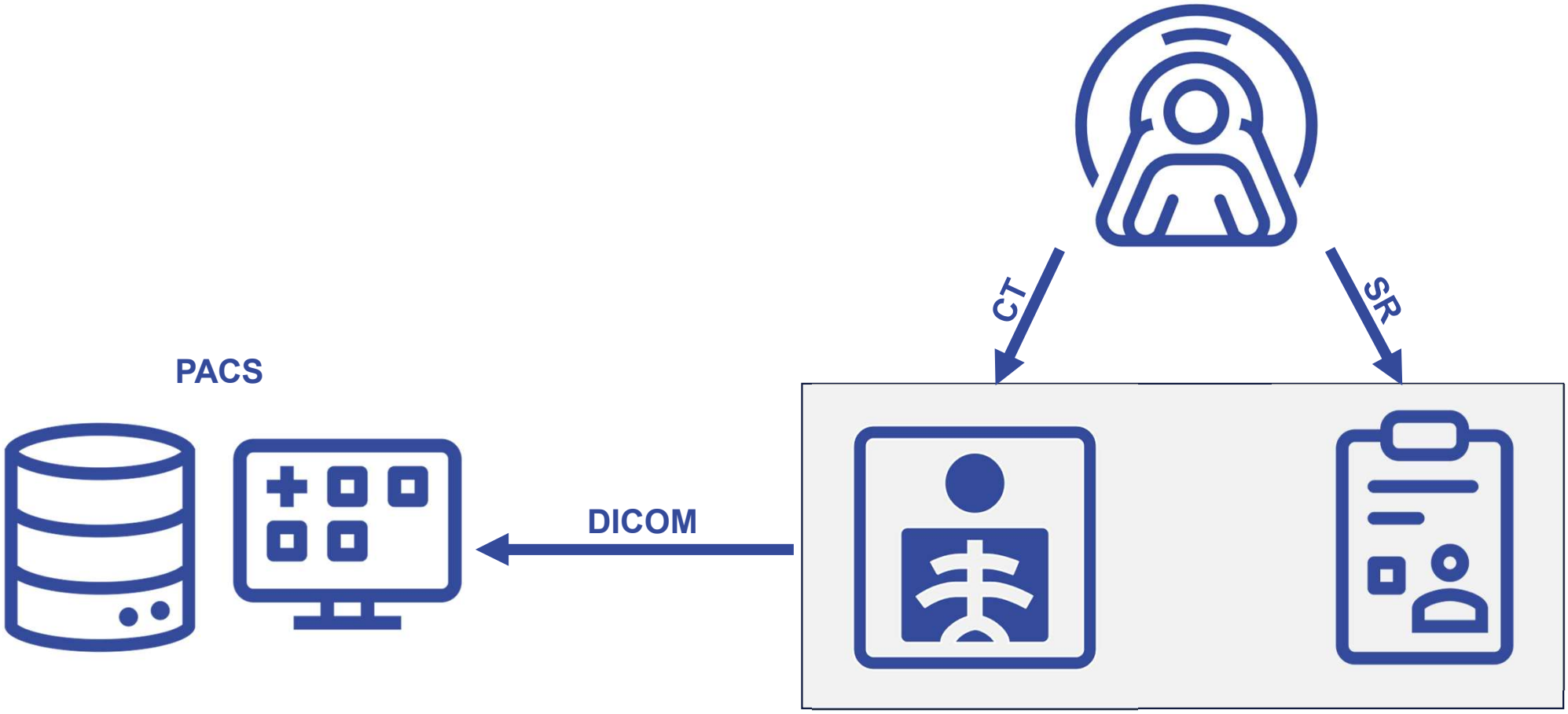
CT scan

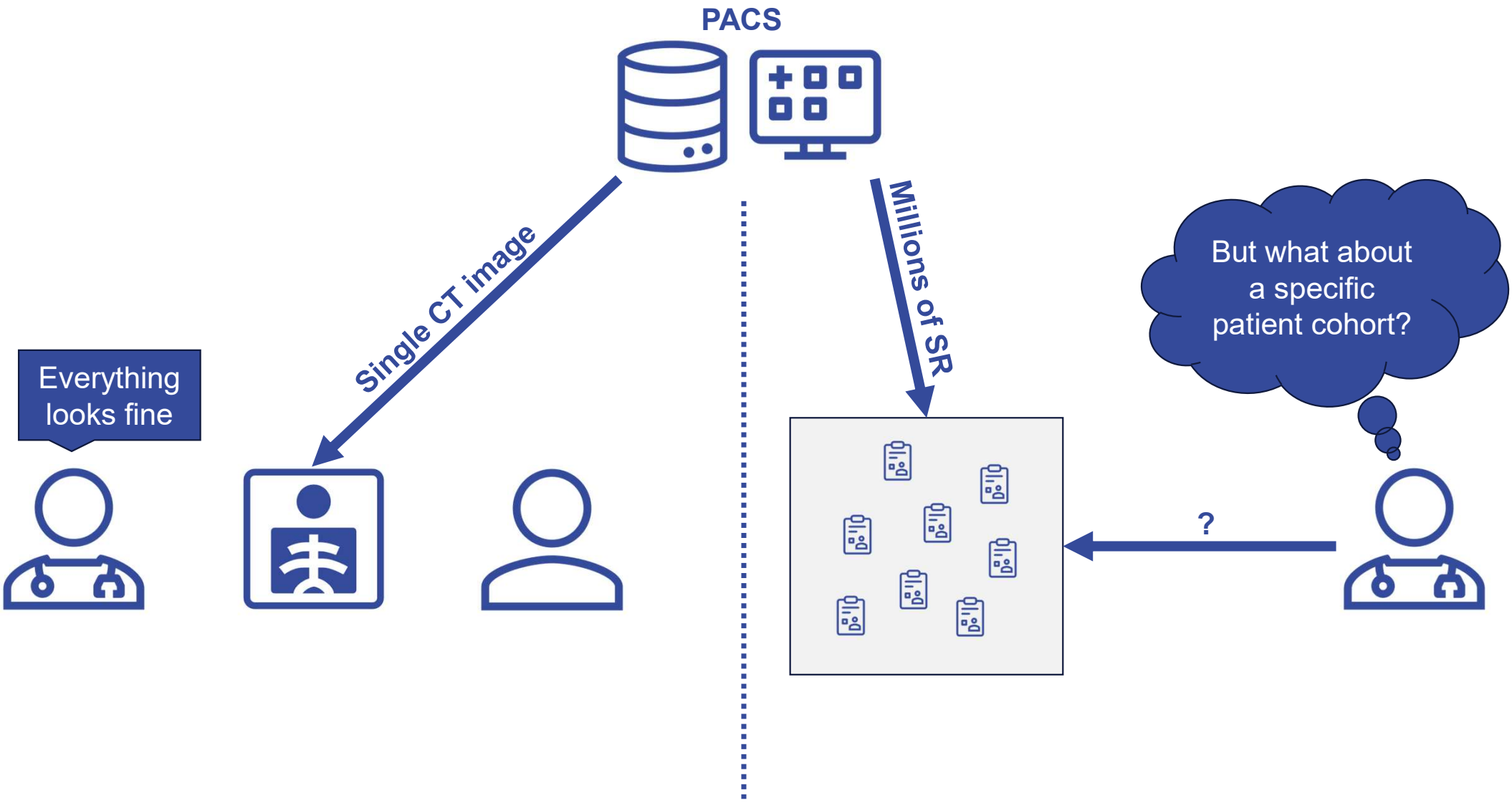
Approximately 9% of all diagnostic studies involving ionizing radiation **are CT studies**.

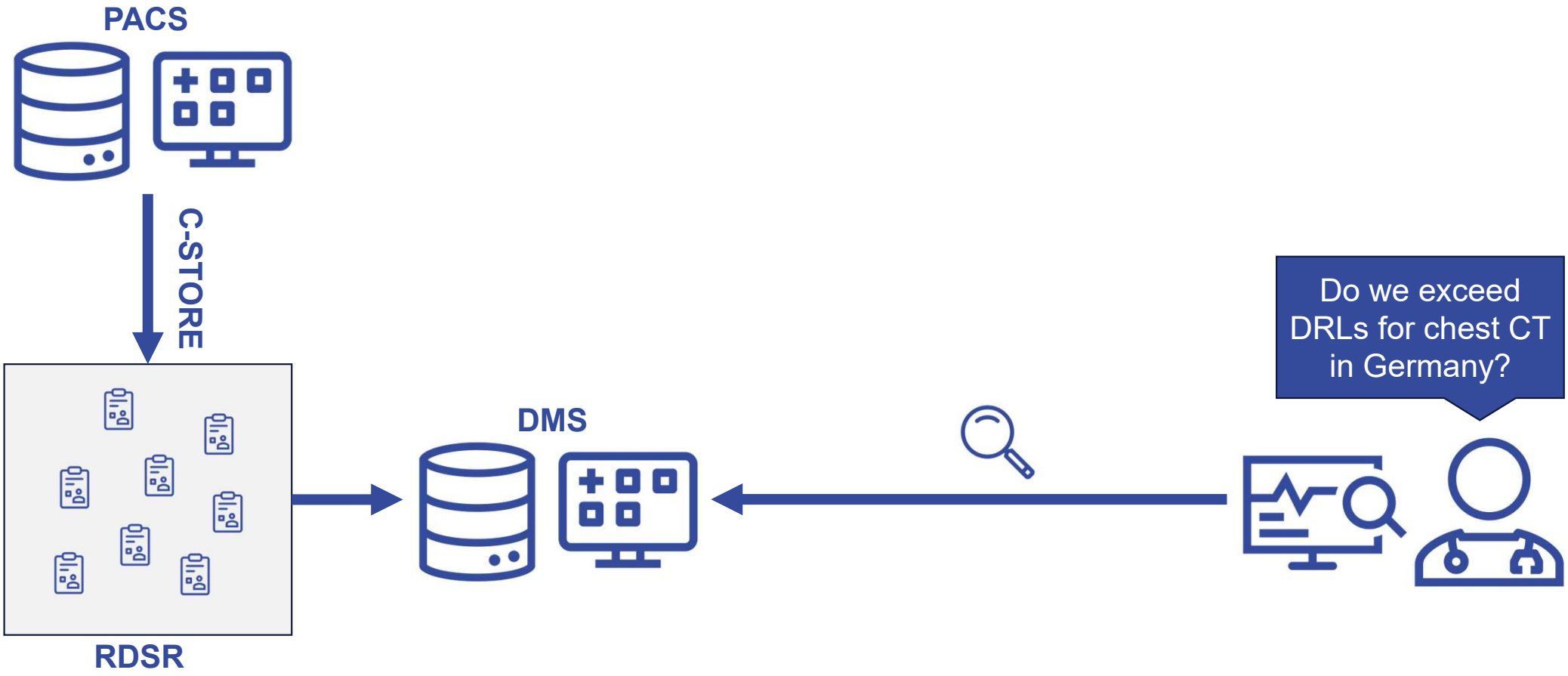
On average, **more than 60% of the collective effective dose** due to medical radiation exposure **is attributable to CT**.

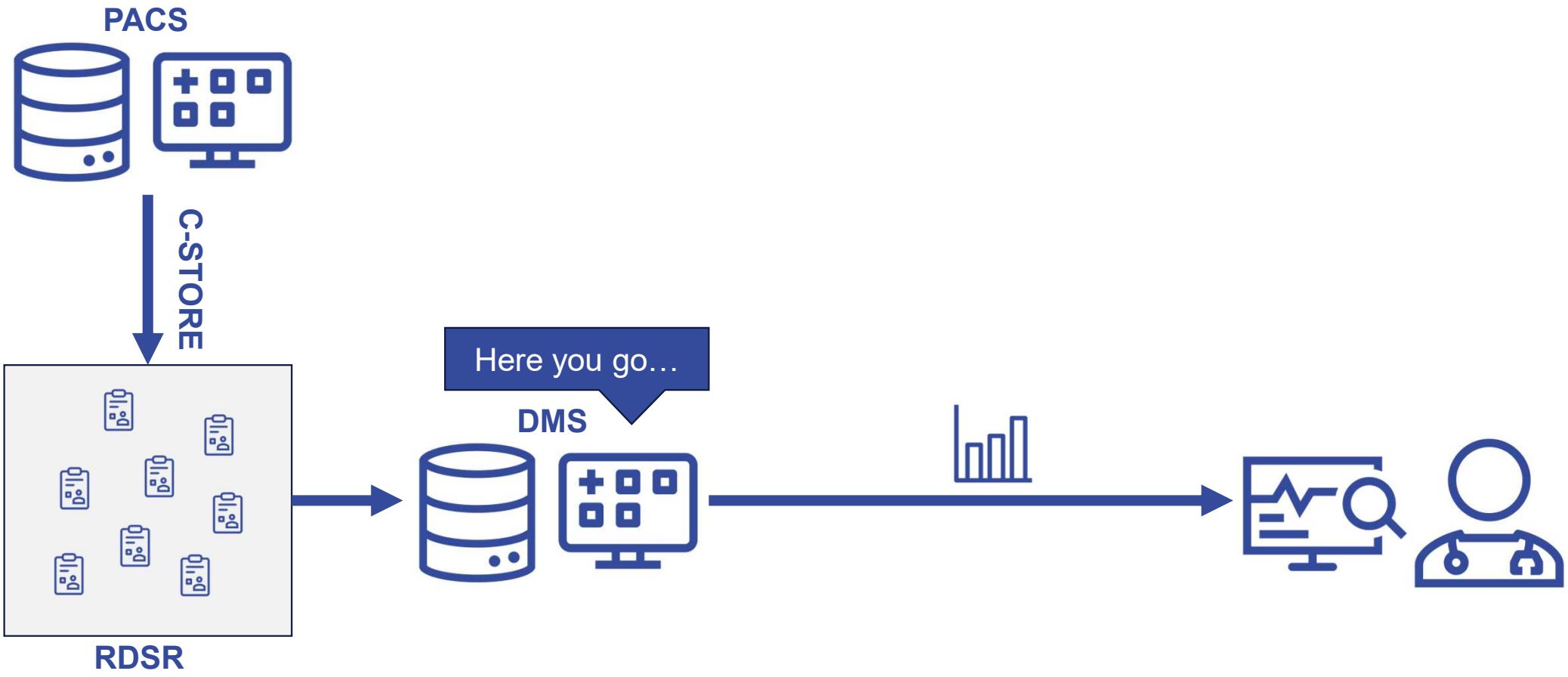
➔ Strict monitoring of radiation doses

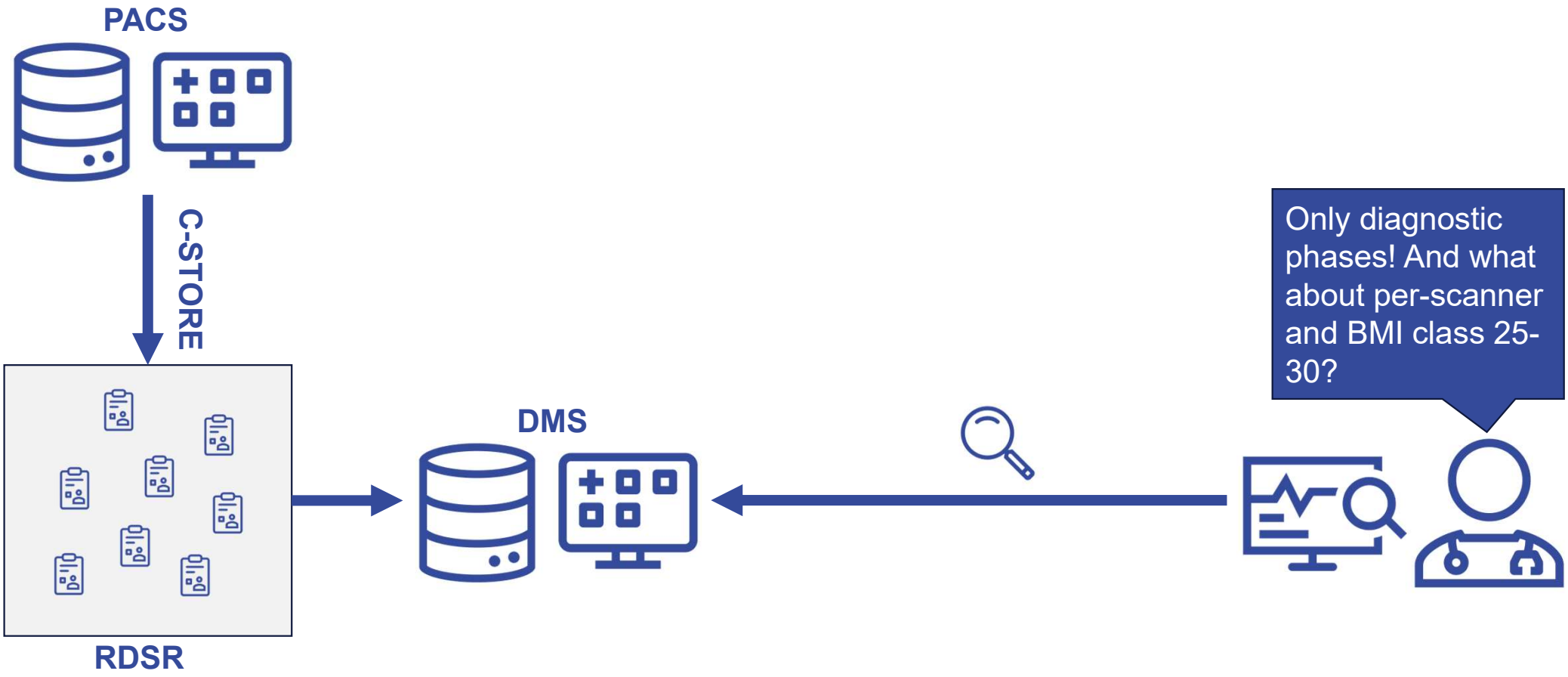

Topogram
PreMonitoring
Monitoring
Lung (Diagnostic)
Derived measures
Measurement
Value
$CTDI_{vol}$ Total
x mGy
Scan Length Total
y cm
Dose Length Product Total
z mGy.cm
⋮
Age: 27
Structured Report (SR)

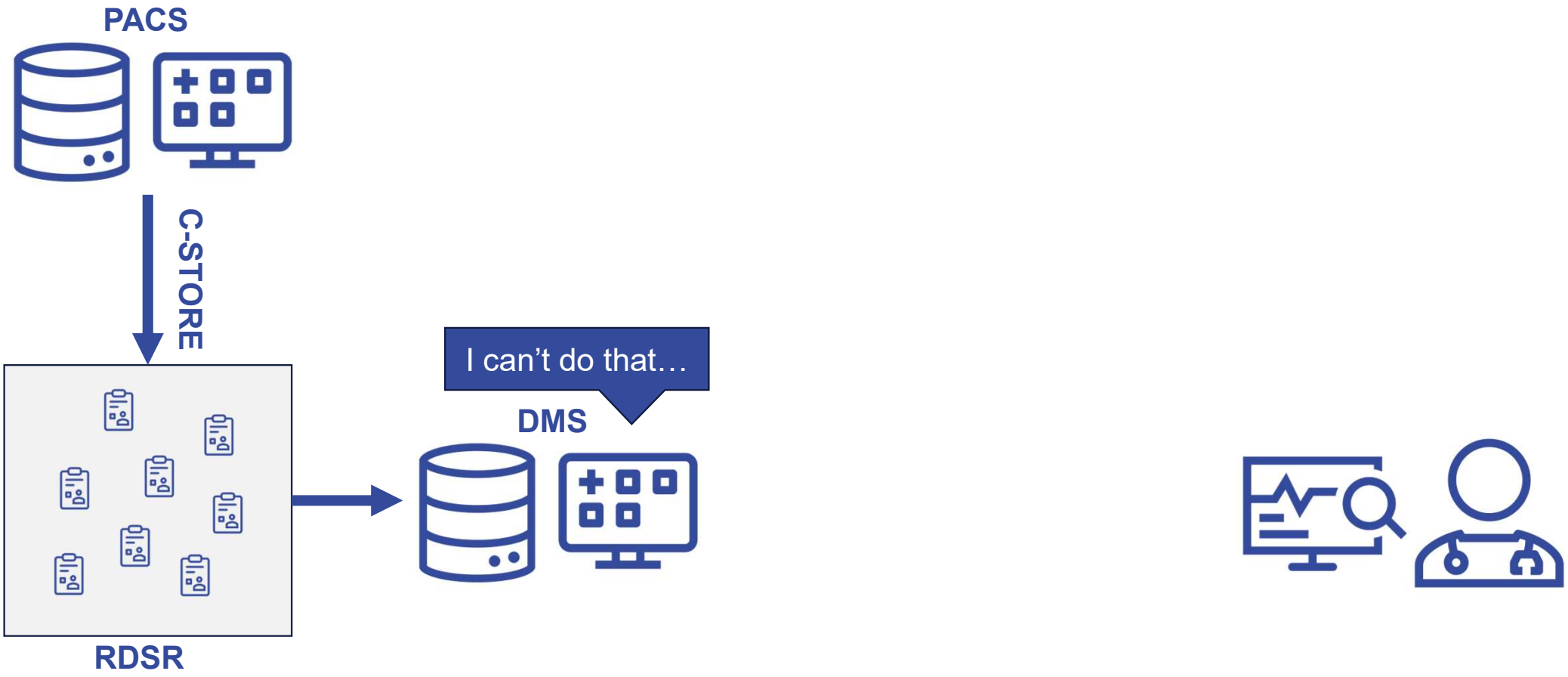

PACS
DICOM
CT
SR

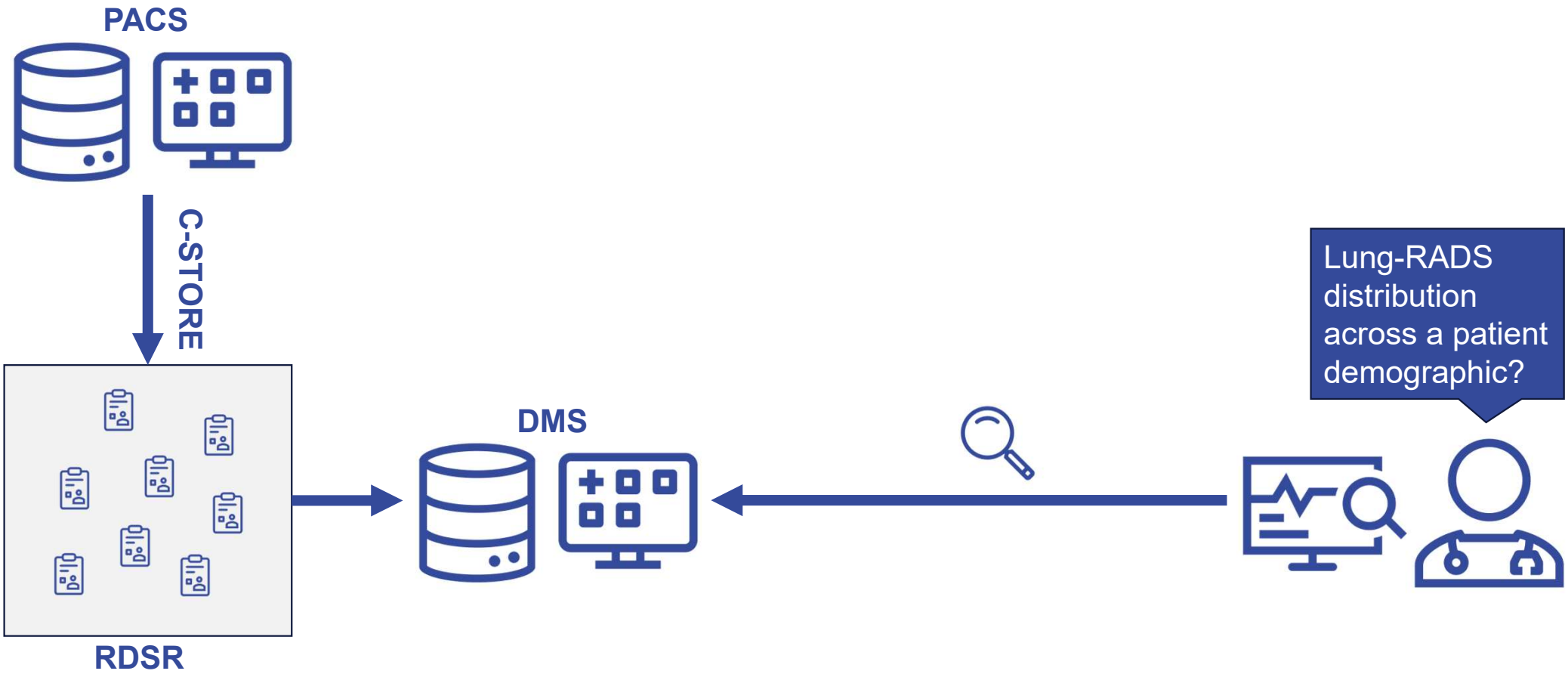

PACS
Single CT image
Millions of SR
Everything looks fine
But what about a specific patient cohort?
?

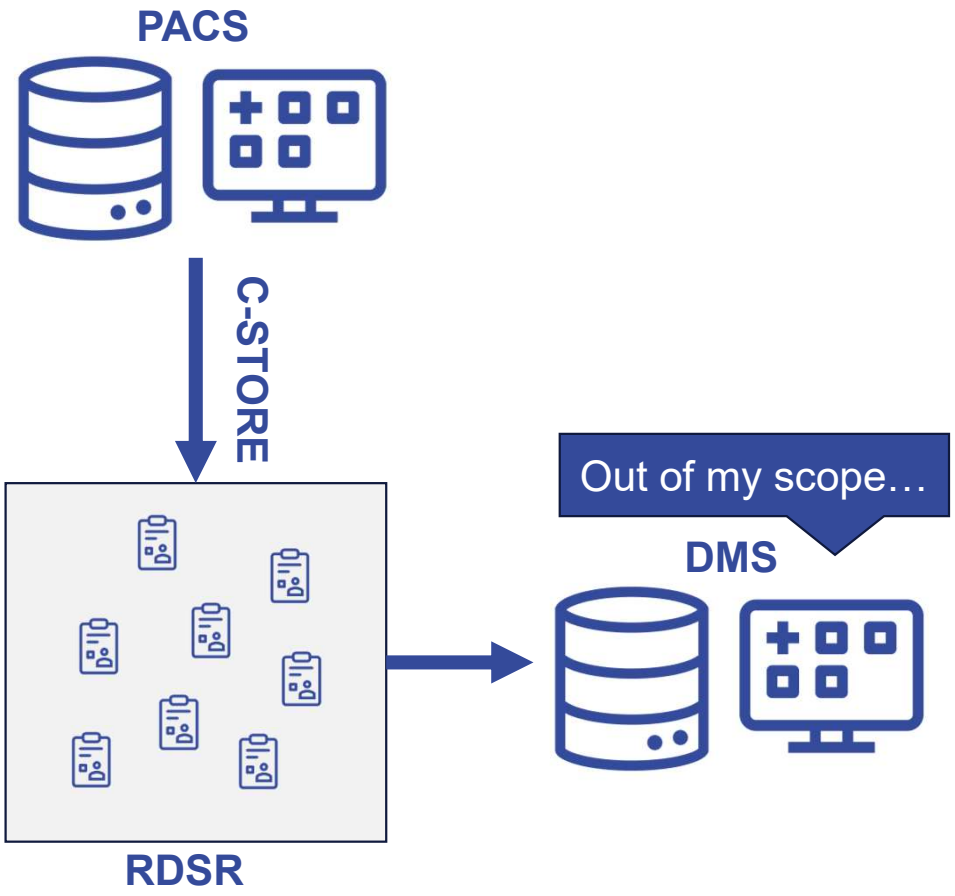

PACS
C-STORE
RDSR
DMS
Do we exceed DRLs for chest CT in Germany?


PACS
C-STORE
RDSR
Here you go…
DMS


PACS
C-STORE
RDSR
DMS
Only diagnostic phases! And what about per-scanner and BMI class 25-30?

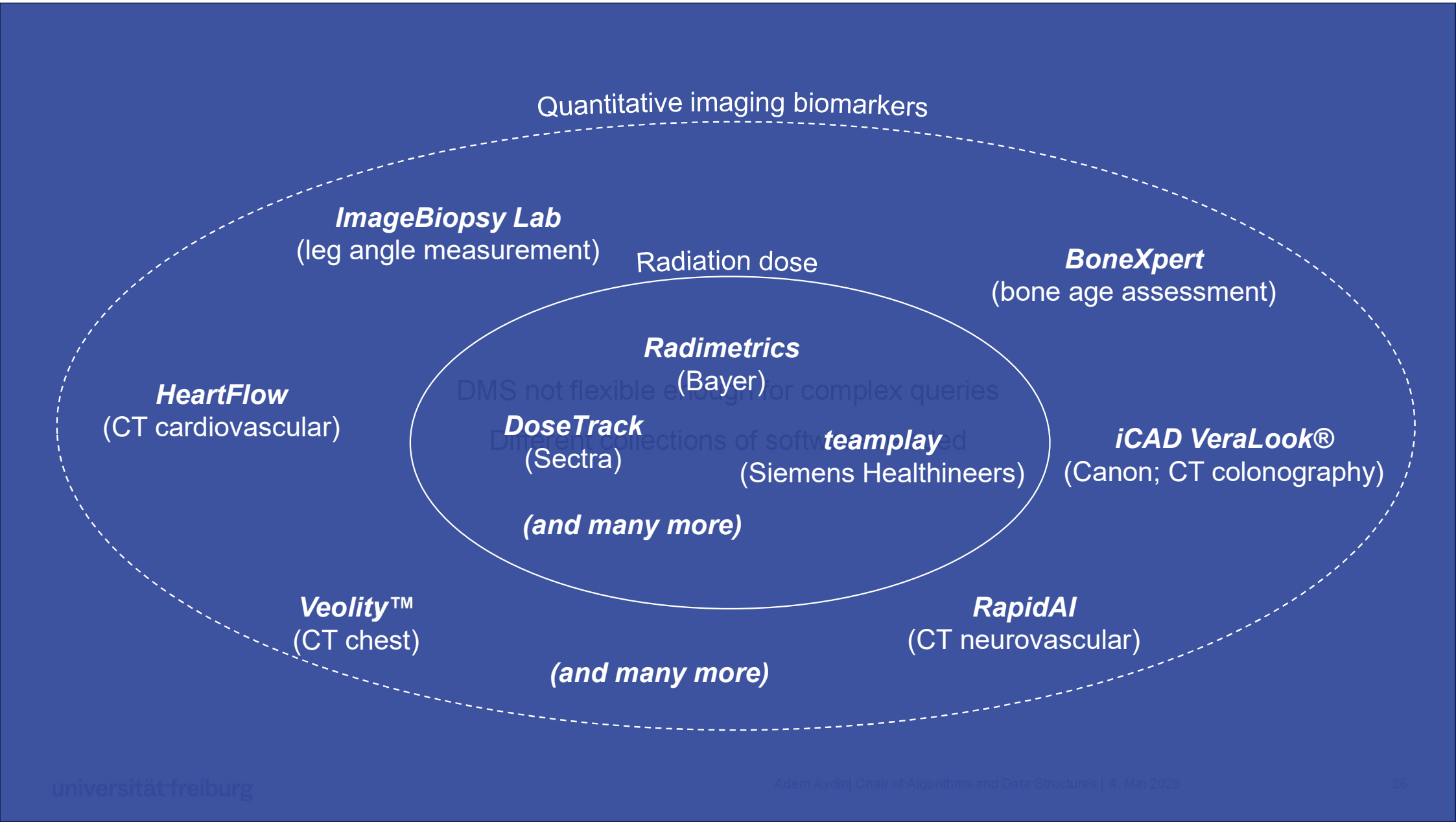

PACS
C-STORE
I can’t do that…
DMS
RDSR


PACS
C-STORE
RDSR
DMS
Lung-RADS distribution across a patient demographic?


PACS
C-STORE
RDSR
Out of my scope…
DMS

DMS not flexible enough for complex queries

Different collections of software needed


Quantitative imaging biomarkers
ImageBiopsy Lab
(leg angle measurement)
Radiation dose
BoneXpert
(bone age assessment)
Radimetrics
(Bayer)
HeartFlow
(CT cardiovascular)
DoseTrack
(Sectra)
teamplay
(Siemens Healthineers)
iCAD VeraLook®
(Canon; CT colonography)
(and many more)
Veolity™
(CT chest)
RapidAI
(CT neurovascular)
(and many more)

Problem Statement

Given millions of DICOM SR documents, how can we

a) index and query **all** measurements

b) provide an **easy-to-use** and **fast** search interface

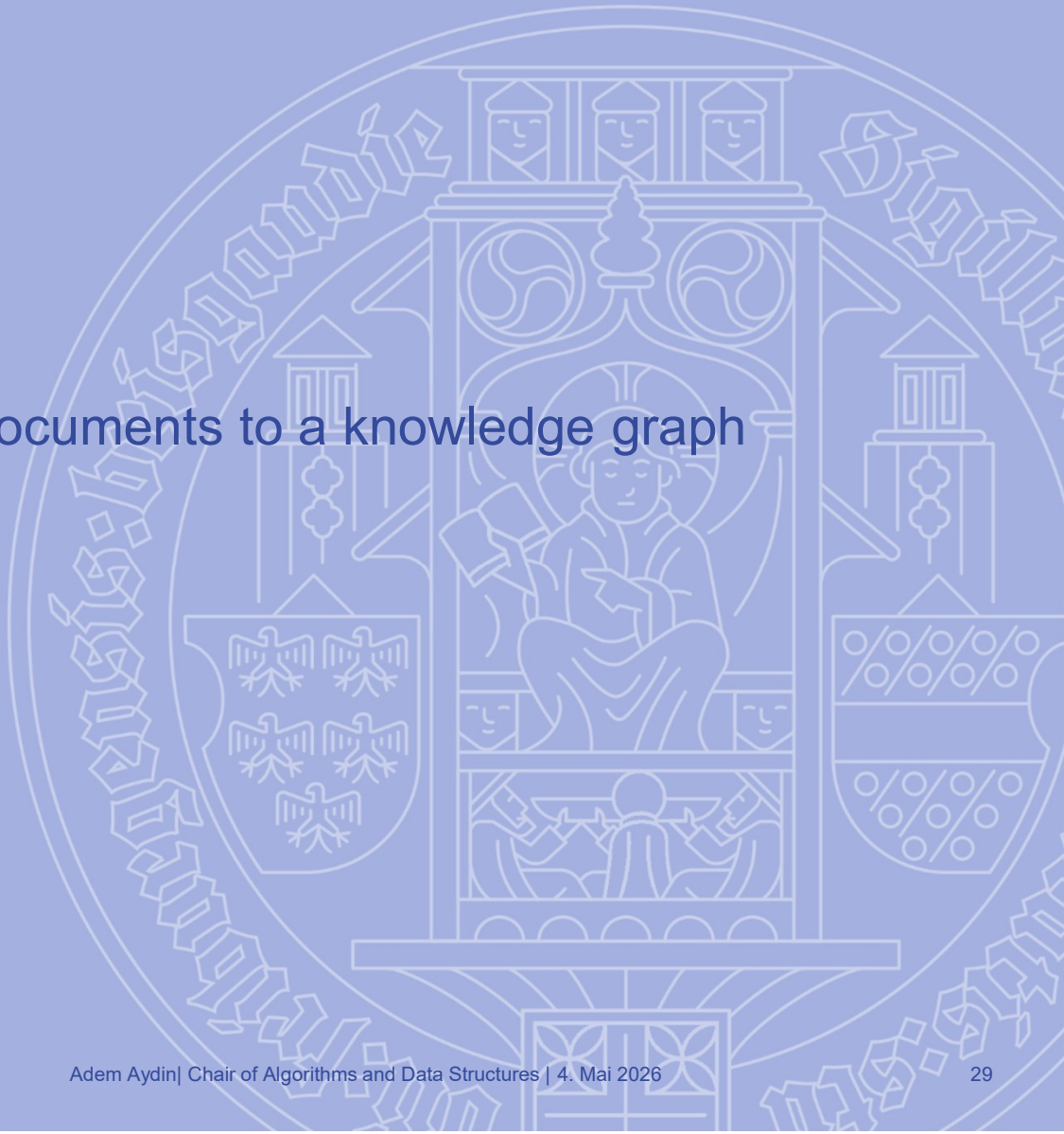
regardless of the clinical context?

Demo!

dicom2rdf – A two-stage data pipeline that turns millions of DICOM SR files into a knowledge graph

Solution

The journey from millions of DICOM SR documents to a knowledge graph



Just enough DICOM

dose-report.dcm

Tag	Meaning	VR	Value
(0010,0010)	Patient's Name	PN	"DOE^JANE"
(0010,1010)	Patient's Age	AS	"042Y"
(0010,1020)	Patient's Size	DS	1.62
(0040,A730)	Content Sequence	SQ	<List of DataSets>

dose-report.dcm

Tag	Meaning	VR	Value
(0010,0010)	Patient's Name	PN	"DOE^JANE"
(0010,1010)	Patient's Age	AS	"042Y"
(0010,1020)	Patient's Size	DS	1.62
(0040,A730)	Content Sequence	SQ	<List of DataSets>

dose-report.dcm

Tag	Meaning	VR	Value
(0010,0010)	Patient's Name	PN	"DOE^JANE"
(0010,1010)	Patient's Age	AS	"042Y"
(0010,1020)	Patient's Size	DS	1.62
(0040,A730)	Content Sequence	SQ	<List of DataSets>

dose-report.dcm

Tag	Meaning	VR	Value
(0010,0010)	Patient's Name	PN	"DOE^JANE"
(0010,1010)	Patient's Age	AS	"042Y"
(0010,1020)	Patient's Size	DS	1.62
(0040,A730)	Content Sequence	SQ	<List of DataSets>

dose-report.dcm

Data Element

Tag	Meaning	VR	Value
(0010,0010)	Patient's Name	PN	"DOE^JANE"
(0010,1010)	Patient's Age	AS	"042Y"
(0010,1020)	Patient's Size	DS	1.62
(0040,A730)	Content Sequence	SQ	<List of DataSets>

dose-report.dcm

Data Set

Tag	Meaning	VR	Value
(0010,0010)	Patient's Name	PN	"DOE^JANE"
(0010,1010)	Patient's Age	AS	"042Y"
(0010,1020)	Patient's Size	DS	1.62
(0040,A730)	Content Sequence	SQ	<List of DataSets>

dose-report.dcm

Tag	Meaning	VR	Value
(0010,0010)	Patient's Name	PN	"DOE^JANE"
(0010,1010)	Patient's Age	AS	042Y
(0018,1150)	Exposure Time	IS	1000
(0040,A730)	Content Sequence	SQ	<List of DataSets>

"What are the $CTDI_{vol}$ measurements of a given Structured Report?"

dose-report.dcm

Tag	Meaning	VR	Value
(0010,0010)	Patient's Name	PN	"DOE^JANE"
(0010,1010)	Patient's Age	AS	"042Y"
(0010,1020)	Patient's Size	DS	1.62
(0040,A730)	Content Sequence	SQ	<List of DataSets>

dose-report.dcm

Tag	Meaning	VR	Value
(0010,0010)	Patient's Name	PN	"DOE^JANE"
(0010,1010)	Patient's Age	AS	"042Y"
(0010,1020)	Patient's Size	DS	1.62
(0040,A730)	Content Sequence	SQ	<List of DataSets>

Tag	Meaning	VR	Value
…	…	…	…
…	…	…	…
…	…	…	…

Tag	Meaning	VR	Value
…	…	…	…
…	…	…	…
…	…	…	…

Tag	Meaning	VR	Value
…	…	…	…
…	…	…	…
…	…	…	…

dose-report.dcm

Tag	Meaning	VR	Value
(0010,0010)	Patient's Name	PN	"DOE^JANE"
(0010,1010)	Patient's Age	AS	"042Y"
(0010,1020)	Patient's Size	DS	1.62
(0040,A730)	Content Sequence	SQ	<List of DataSets>

Tag	Meaning	VR	Value
…	…	…	…
…	…	…	…
…	…	…	…

Tag	Meaning	VR	Value
(0040,A040)	Value Type	CS	"CONTAINER"
(0040,A043)	Concept Name Code Sequence	SQ	<DataSet>
(0040,A730)	Content Sequence	SQ	<List of DataSets>

Tag	Meaning	VR	Value
…	…	…	…
…	…	…	…
…	…	…	…

Some Content Item?

Tag	Meaning	VR	Value
(0040,A040)	Value Type	CS	“CONTAINER”
(0040,A043)	Concept Name Code Sequence	SQ	<DataSet>
(0040,A730)	Content Sequence	SQ	<List of DataSets>

CT Acquisition Content Item!

Tag	Meaning	VR	Value
(0040,A040)	Value Type	CS	“CONTAINER”
(0040,A043)	Concept Name Code Sequence	SQ	<DataSet>
(0040,A730)	Content Sequence	SQ	<List of DataSets>

Tag	Meaning	VR	Value
(0008,0100)	Code Value	SH	“113819”
(0008,0102)	Coding Scheme Designator	SH	“DCM”
(0008,0104)	Code Meaning	LO	“CT Acquisition”

CT Acquisition Content Item!

Tag	Meaning	VR	Value
(0040,A040)	Value Type	CS	“CONTAINER”
(0040,A043)	Concept Name Code Sequence	SQ	<DataSet>
(0040,A730)	Content Sequence	SQ	<List of DataSets>

Tag	Meaning	VR	Value
(0008,0100)	Code Value	SH	“113819”
(0008,0102)	Coding Scheme Designator	SH	“DCM”
(0008,0104)	Code Meaning	LO	“CT Acquisition”

Tag	Meaning	VR	Value
(0040,A040)	Value Type	CS	“CONTAINER”
(0040,A043)	Concept Name Code Sequence	SQ	<DataSet>
(0040,A730)	Content Sequence	SQ	<List of DataSets>

Some Content Item!

Tag	Meaning	VR	Value
(0040,A040)	Value Type	CS	“CONTAINER”
(0040,A043)	Concept Name Code Sequence	SQ	<DataSet>
(0040,A730)	Content Sequence	SQ	<List of DataSets>

Tag	Meaning	VR	Value
(0008,0100)	Code Value	SH	“113819”
(0008,0102)	Coding Scheme Designator	SH	“DCM”
(0008,0104)	Code Meaning	LO	“CT Acquisition”

Tag	Meaning	VR	Value
(0040,A040)	Value Type	CS	“CONTAINER”
(0040,A043)	Concept Name Code Sequence	SQ	<DataSet>
(0040,A730)	Content Sequence	SQ	<List of DataSets>

CT Dose Content Item!

Tag	Meaning	VR	Value
(0040,A040)	Value Type	CS	“CONTAINER”
(0040,A043)	Concept Name Code Sequence	SQ	<DataSet>
(0040,A730)	Content Sequence	SQ	<List of DataSets>

Tag	Meaning	VR	Value
(0008,0100)	Code Value	SH	“113829”
(0008,0102)	Coding Scheme Designator	SH	“DCM”
(0008,0104)	Code Meaning	LO	“CT Dose”

CT Dose Content Item!

Tag	Meaning	VR	Value
(0040,A040)	Value Type	CS	“CONTAINER”
(0040,A043)	Concept Name Code Sequence	SQ	<DataSet>
(0040,A730)	Content Sequence	SQ	<List of DataSets>

Tag	Meaning	VR	Value
(0008,0100)	Code Value	SH	“113829”
(0008,0102)	Coding Scheme Designator	SH	“DCM”
(0008,0104)	Code Meaning	LO	“CT Dose”

Tag	Meaning	VR	Value
(0040,A040)	Value Type	CS	“NUM”
(0040,A043)	Concept Name Code Sequence	SQ	<DataSet>
(0040,A300)	Measured Value Sequence	SQ	<List of DataSets>

CT Dose Container!

Tag	Meaning	VR	Value
(0040,A040)	Value Type	CS	“CONTAINER”
(0040,A043)	Concept Name Code Sequence	SQ	<DataSet>
(0040,A730)	Content Sequence	SQ	<List of DataSets>

Tag	Meaning	VR	Value
(0008,0100)	Code Value	SH	“113829”
(0008,0102)	Coding Scheme Designator	SH	“DCM”
(0008,0104)	Code Meaning	LO	“CT Dose”

Tag	Meaning	VR	Value
(0040,A040)	Value Type	CS	“NUM”
(0040,A043)	Concept Name Code Sequence	SQ	<DataSet>
(0040,A300)	Measured Value Sequence	SQ	<List of DataSets>

Mean CTDIvol Content Item!

Tag	Meaning	VR	Value
(0040,A040)	Value Type	CS	“NUM”
(0040,A043)	Concept Name Code Sequence	SQ	<DataSet>
(0040,A300)	Measured Value Sequence	SQ	<List of DataSets>

Tag	Meaning	VR	Value
(0008,0100)	Code Value	SH	“113830”
(0008,0102)	Coding Scheme Designator	SH	“DCM”
(0008,0104)	Code Meaning	LO	“Mean CTDIvol”

Mean CTDIvol Content Item!

Tag	Meaning	VR	Value
(0040,A040)	Value Type	CS	“NUM”
(0040,A043)	Concept Name Code Sequence	SQ	<DataSet>
(0040,A300)	Measured Value Sequence	SQ	<List of DataSets>

Tag	Meaning	VR	Value
(0008,0100)	Code Value	SH	“113830”
(0008,0102)	Coding Scheme Designator	SH	“DCM”
(0008,0104)	Code Meaning	LO	“Mean CTDIvol”

Tag	Meaning	VR	Value
(0040,08EA)	Measurement Units Code Sequence	SQ	<DataSet>
(0040,A30A)	Numeric Value	DS	0.01

Mean CTDIvol Content Item!

Tag	Meaning	VR	Value
(0040,A040)	Value Type	CS	“NUM”
(0040,A043)	Concept Name Code Sequence	SQ	<DataSet>
(0040,A300)	Measured Value Sequence	SQ	<List of DataSets>

Tag	Meaning	VR	Value
(0008,0100)	Code Value	SH	“113830”
(0008,0102)	Coding Scheme Designator	SH	“DCM”
(0008,0104)	Code Meaning	LO	“Mean CTDIvol”

Tag	Meaning	VR	Value
(0040,08EA)	Measurement Units Code Sequence	SQ	<DataSet>
(0040,A30A)	Numeric Value	DS	0.01

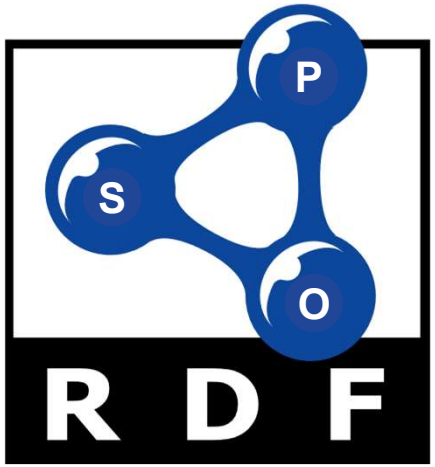
Sequence?

Tag	Meaning	VR	Value
(0040,08EA)	Measurement Units Code Sequence	SQ	<DataSet>
(0040,A30A)	Numeric Value	DS	0.01

Tag	Meaning	VR	Value
(0008,0100)	Code Value	SH	“mGy”
(0008,0102)	Coding Scheme Designator	SH	“UCUM”
(0008,0104)	Code Meaning	LO	“mGy”

DICOM files form a directed acyclic graph (DAG)

→ Use **R**esource **D**escription **F**ramework (RDF) to interpret DICOM files as a knowledge graph


P
S
O
R D F

Subject **Predicate** **Object**

Subject
Predicate
Object

RDF data model

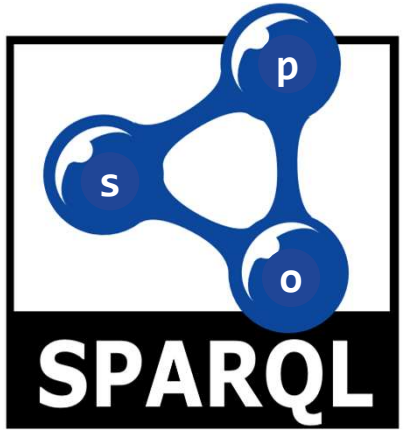
RDF data model

Statement about the world

Statement in a domain

DICOM files are modeled as RDF triples

→ Use **S**PARQL **P**rotocol **A**nd **R**DF **Q**uery **L**anguage (SPARQL) to perform queries on DICOM files

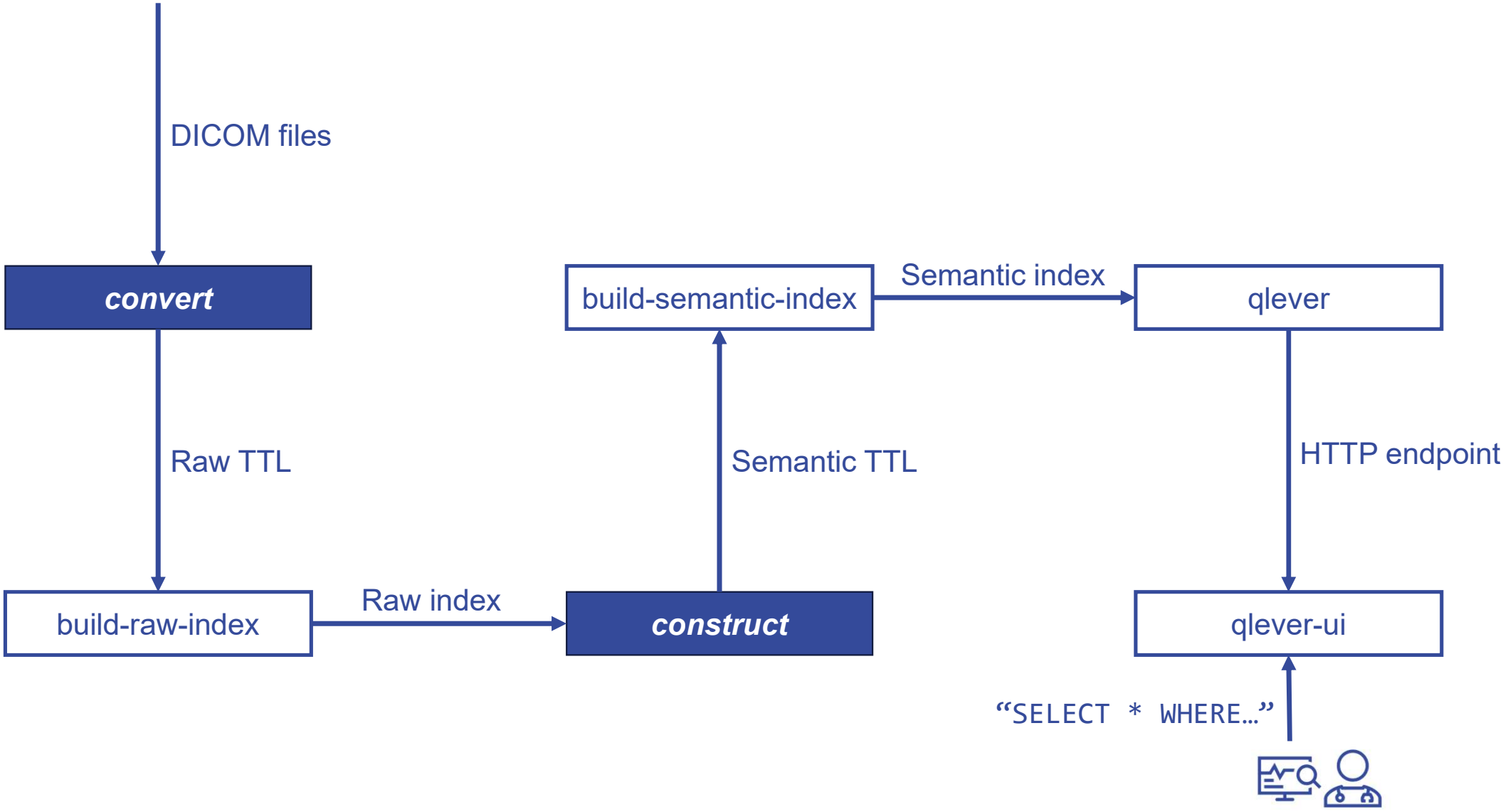

p
s
o
SPARQL

```
SELECT * WHERE {
    ?s ?p ?o .
}
```

```
SELECT * WHERE {
    ?s ?p ?o .
}
```

```
SELECT * WHERE {
    ?ct_dose <has CTDIvol> ?ctdivol .
    FILTER(?ctdivol > 0.01)
}
```

dicom2rdf Pipeline Architecture


DICOM files
convert
Raw TTL
build-raw-index
Raw index
construct
Semantic TTL
build-semantic-index
Semantic index
qlever
HTTP endpoint
qlever-ui
“SELECT * WHERE…”

Stage 1 – *convert*

dose-report.dcm

Tag	Meaning	VR	Value
(0010,0010)	Patient's Name	PN	"DOE^JANE"
(0010,1010)	Patient's Age	AS	"042Y"
(0010,1020)	Patient's Size	DS	1.62
(0040,A730)	Content Sequence	SQ	<List of DataSets>

convert(

dose-report.dcm

Tag	Meaning	VR	Value
(0010,0010)	Patient's Name	PN	"DOE^JANE"
(0010,1010)	Patient's Age	AS	"042Y"
(0010,1020)	Patient's Size	DS	1.62
(0040,A730)	Content Sequence	SQ	<List of DataSets>

)

dose-report.dcm

Tag	Meaning	VR	Value
(0010,0010)	Patient's Name	PN	"DOE^JANE"
(0010,1010)	Patient's Age	AS	"042Y"
(0010,1020)	Patient's Size	DS	1.62
(0040,A730)	Content Sequence	SQ	<List of DataSets>

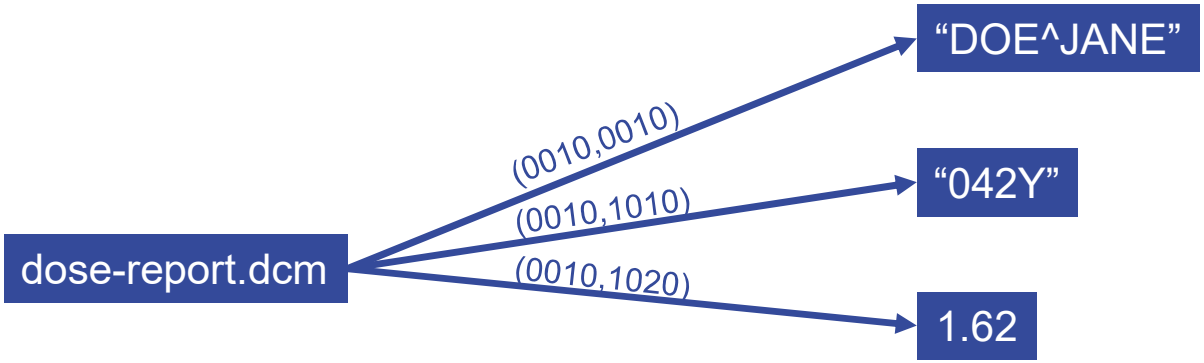
dose-report.dcm

Tag	Meaning	VR	Value
(0010,0010)	Patient's Name	PN	"DOE^JANE"
(0010,1010)	Patient's Age	AS	"042Y"
(0010,1020)	Patient's Size	DS	1.62
(0040,A730)	Content Sequence	SQ	<List of DataSets>

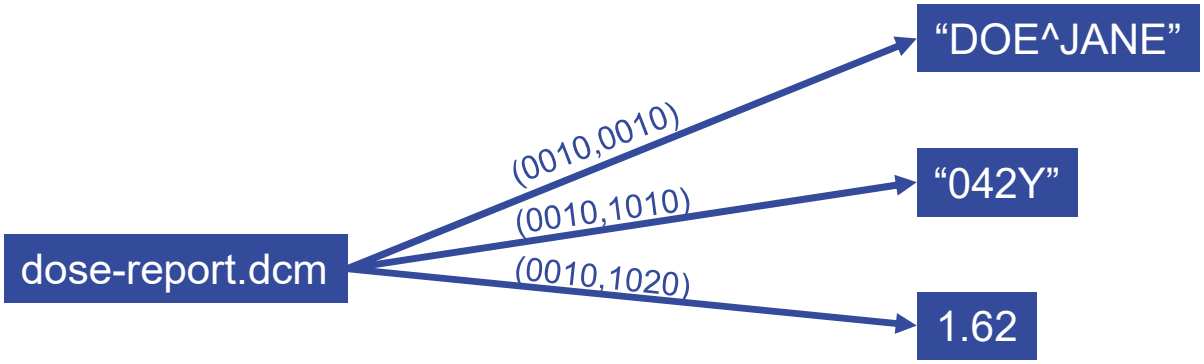
Tag	Meaning	VR	Value
(0010,0010)	Patient's Name	PN	"DOE^JANE"
(0010,1010)	Patient's Age	AS	"042Y"
(0010,1020)	Patient's Size	DS	1.62
(0040,A730)	Content Sequence	SQ	<List of DataSets>

dose-report.dcm

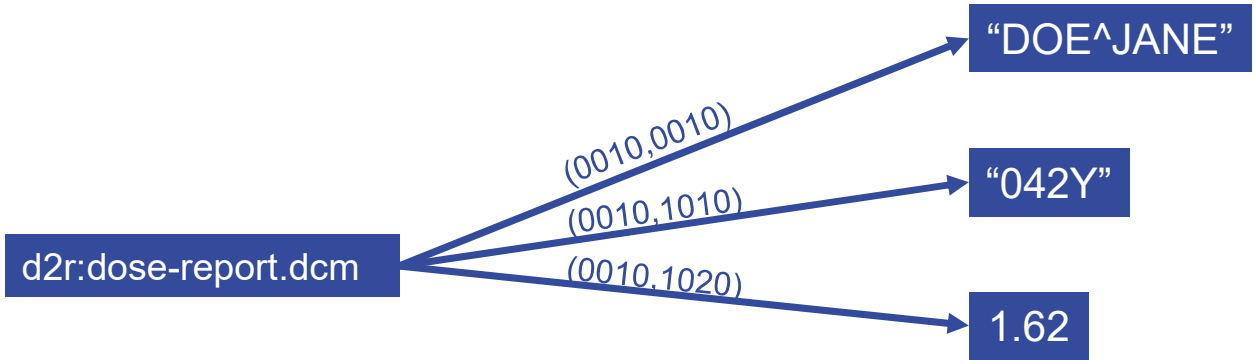
Tag	Meaning	VR	Value
	Patient's Name	PN	
	Patient's Age	AS	
	Patient's Size	DS	
(0040,A730)	Content Sequence	SQ	<List of DataSets>



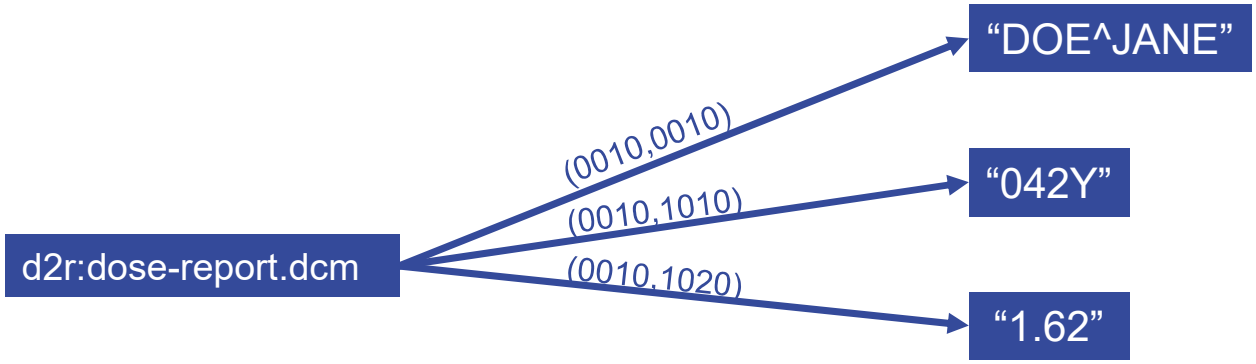
Tag	Meaning	VR	Value
		PN	
		AS	
		DS	
(0040,A730)	Content Sequence	SQ	<List of DataSets>



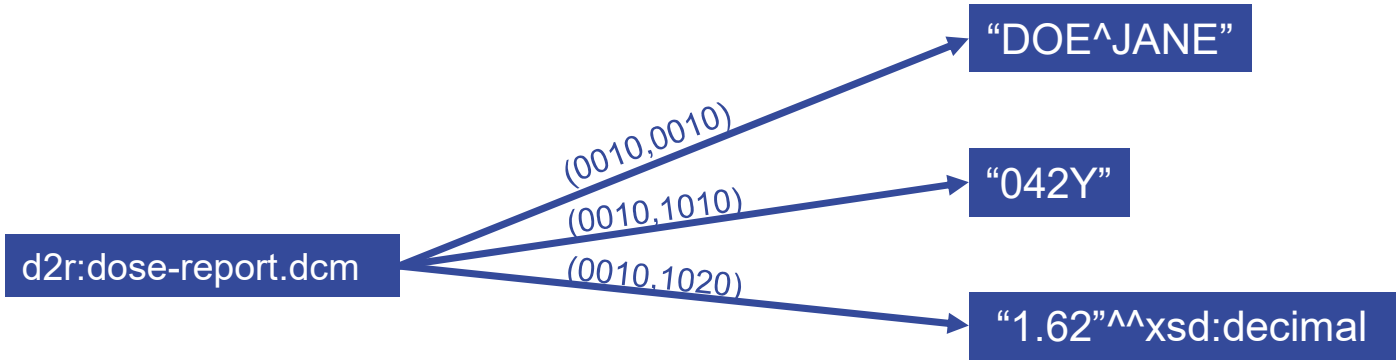
Tag	Meaning	VR	Value
		PN	
		AS	
		DS	
(0040,A730)	Content Sequence	SQ	<List of DataSets>



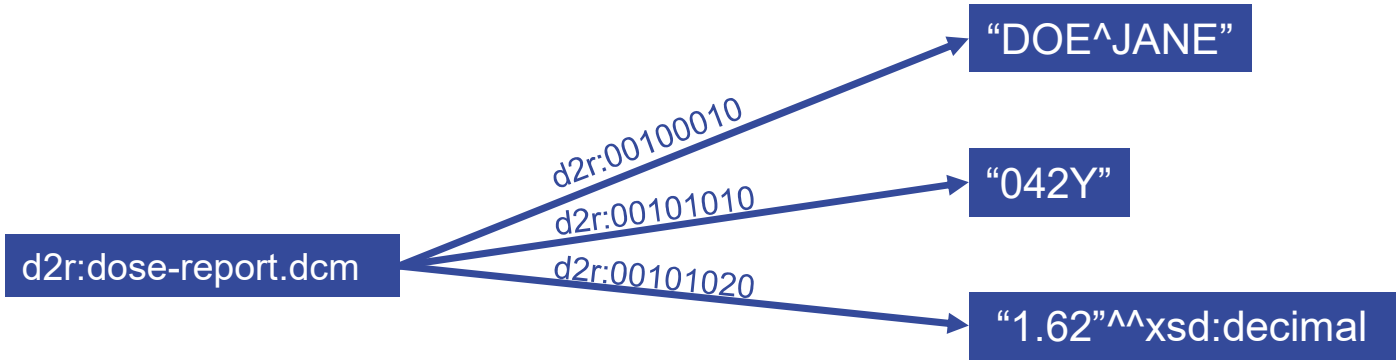
Tag	Meaning	VR	Value
		PN	
		AS	
		DS	
(0040,A730)	Content Sequence	SQ	<List of DataSets>



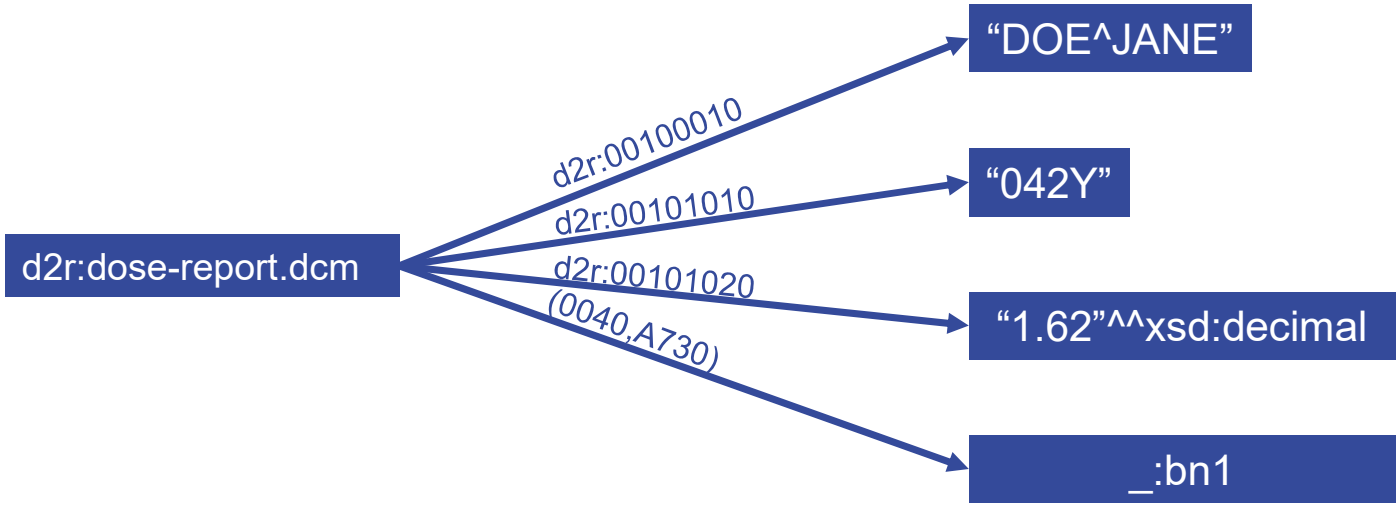
Tag	Meaning	VR	Value
		PN	
		AS	
		DS	
(0040,A730)	Content Sequence	SQ	<List of DataSets>



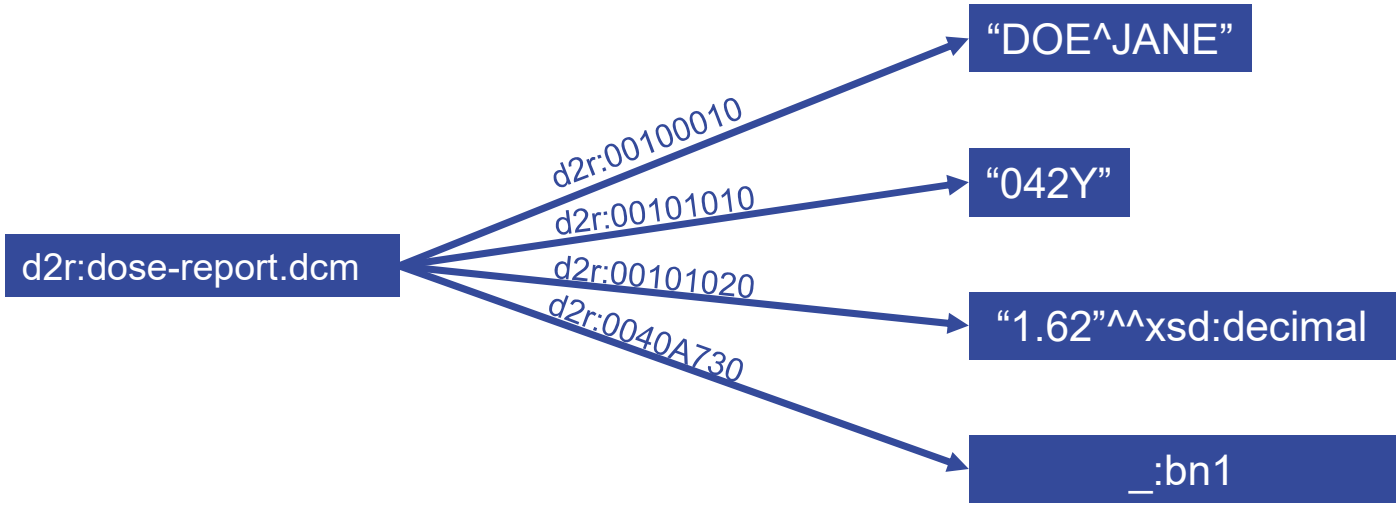
Tag	Meaning	VR	Value
		PN	
		AS	
		DS	
(0040,A730)	Content Sequence	SQ	<List of DataSets>



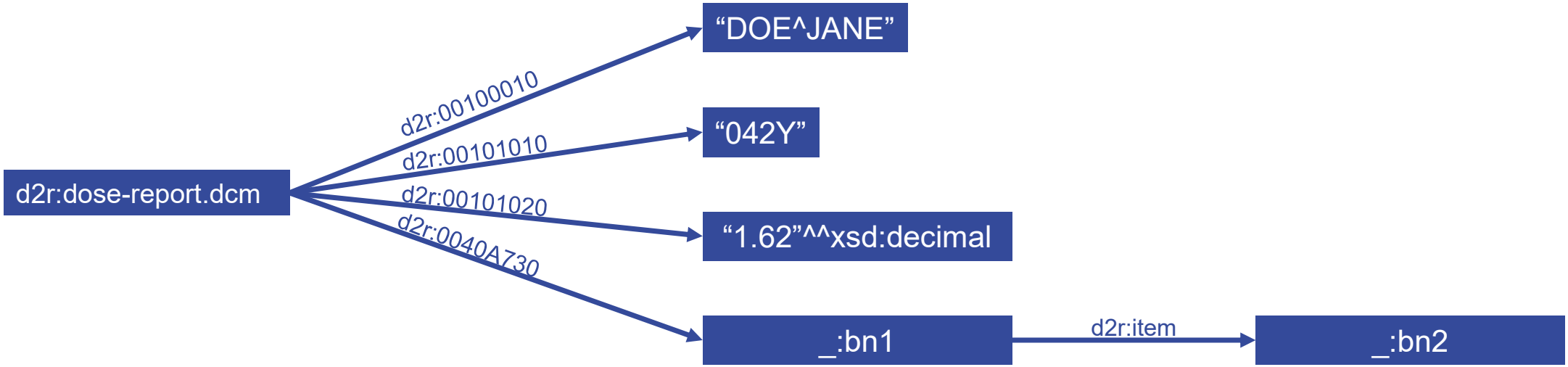
Tag	Meaning	VR	Value
		PN	
		AS	
		DS	
	Content Sequence	SQ	<List of DataSets>



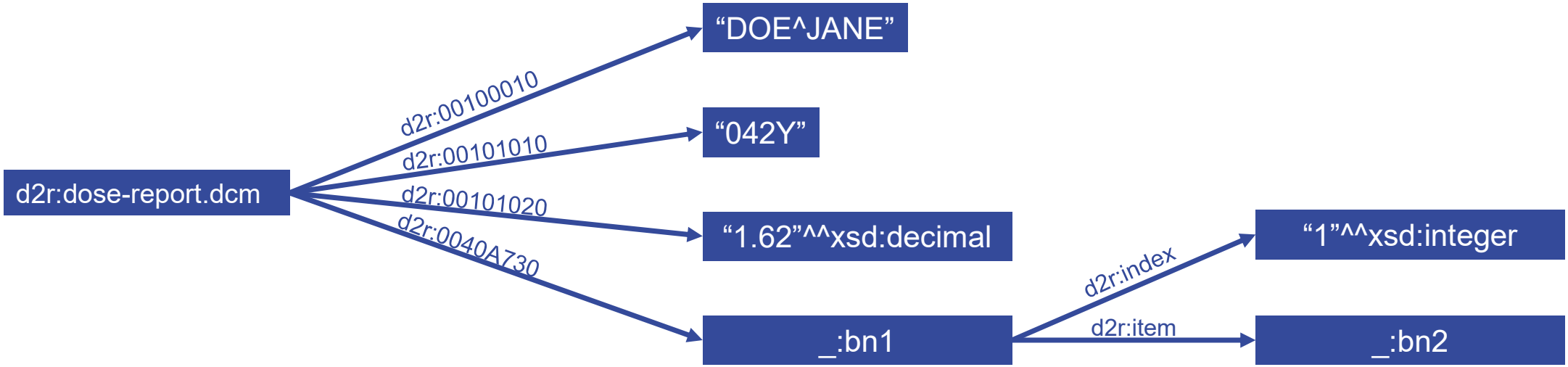
Tag	Meaning	VR	Value
		PN	
		AS	
		DS	
	Content Sequence	SQ	<List of DataSets>

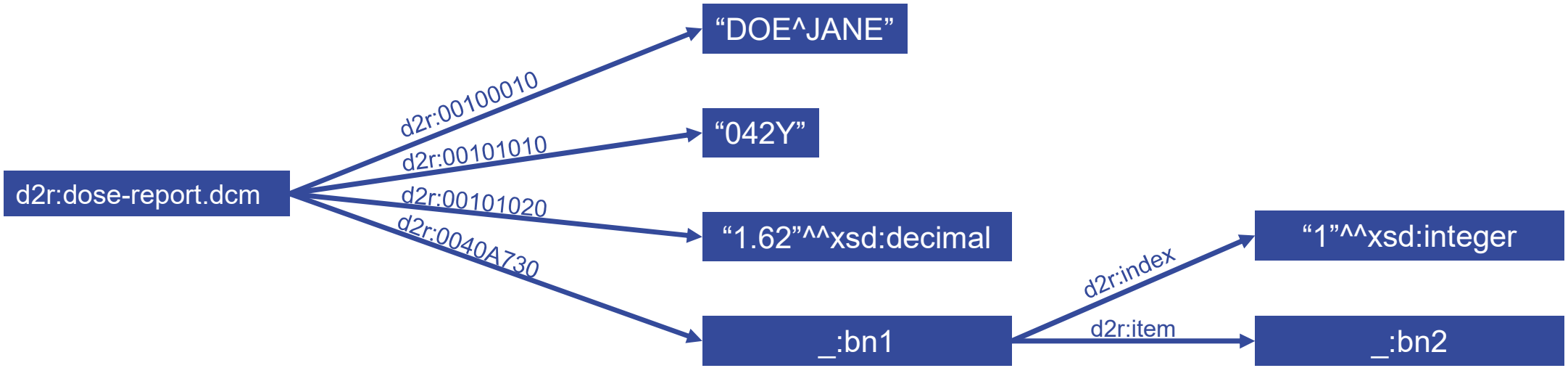


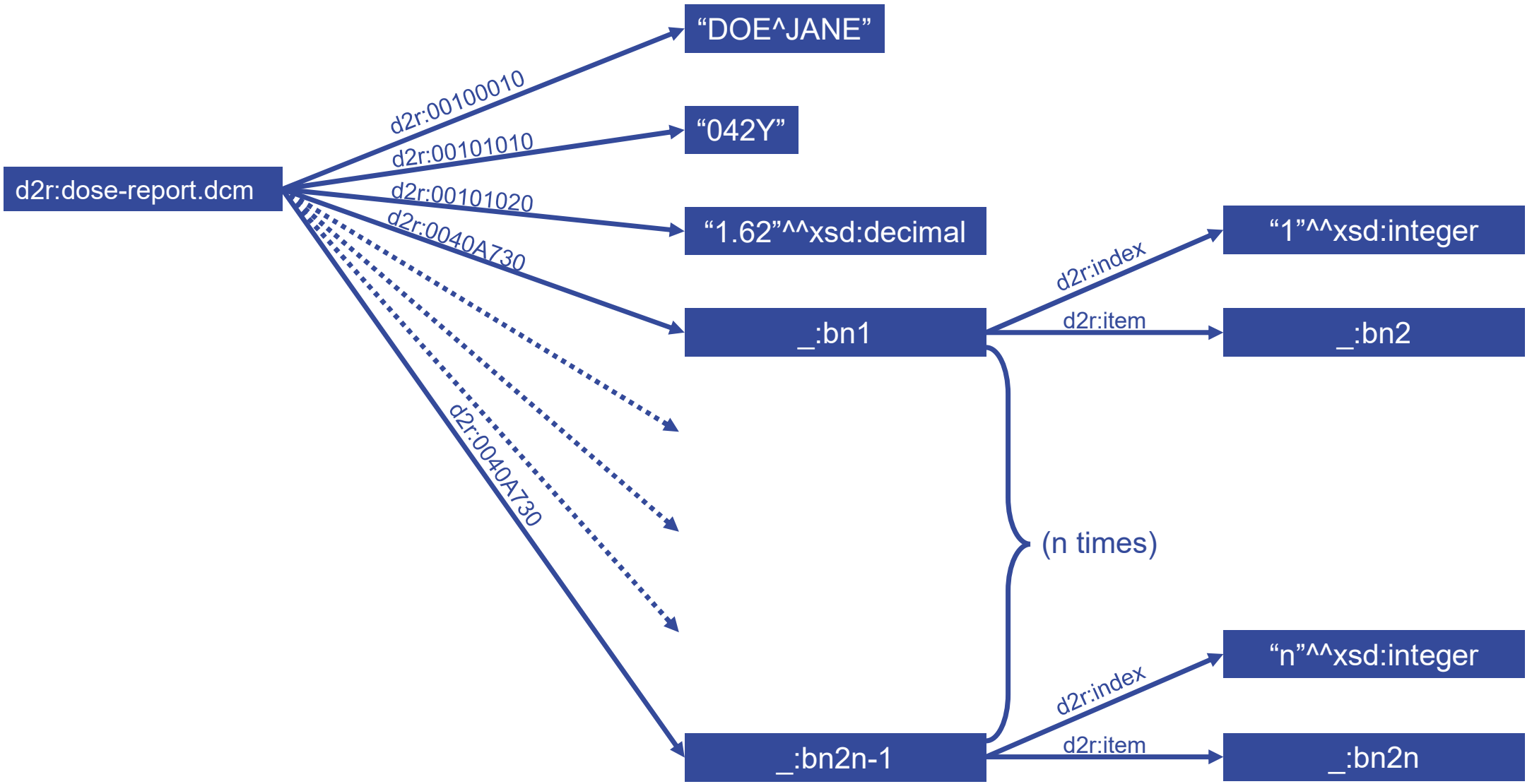
Tag	Meaning	VR	Value
		PN	
		AS	
		DS	
	Content Sequence	SQ	<List of DataSets>

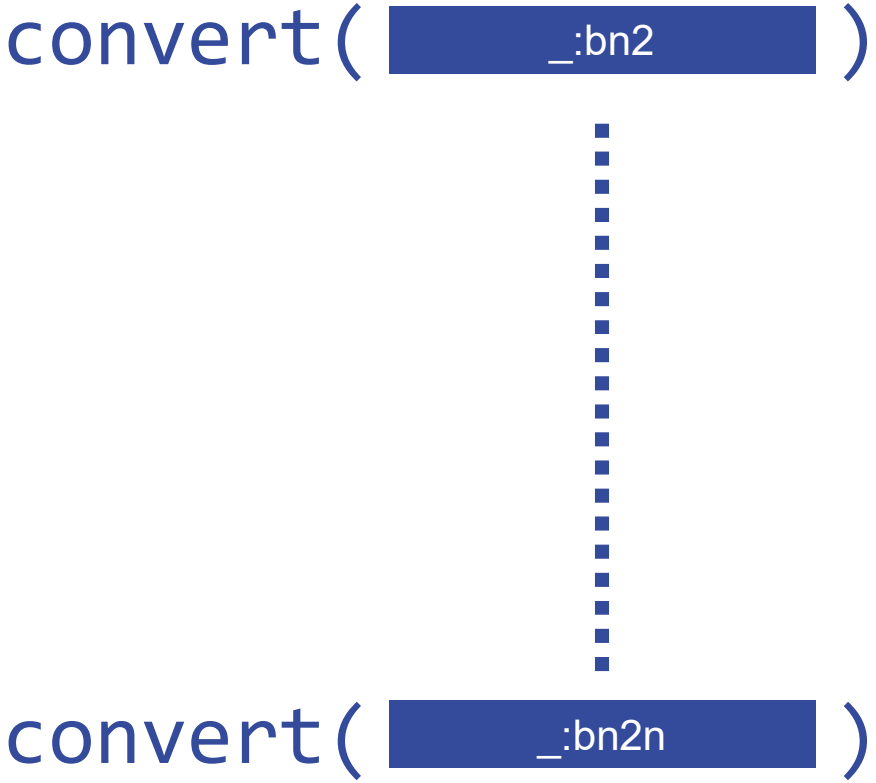


Tag	Meaning	VR	Value
		PN	
		AS	
		DS	
	Content Sequence	SQ	<List of DataSets>




d2r:dose-report.dcm
d2r:00100010
d2r:00101010
d2r:00101020
d2r:0040A730
“DOE^JANE”
“042Y”
“1.62”^^xsd:decimal
_:bn1
d2r:index
d2r:item
“1”^^xsd:integer
_:bn2


d2r:dose-report.dcm
d2r:00100010
“DOE^JANE”
d2r:00101010
“042Y”
d2r:00101020
“1.62”^^xsd:decimal
d2r:0040A730
_:bn1
d2r:index
“1”^^xsd:integer
d2r:item
_:bn2
(n times)
d2r:0040A730
_:bn2n-1
d2r:index
“n”^^xsd:integer
d2r:item
_:bn2n


convert(_:bn2)
convert(_:bn2n)

Are we done yet?

Are we done yet?
Not quite…

Are we done yet?
Not quite…

"What are the $CTDI_{vol}$ measurements of a given Structured Report?"

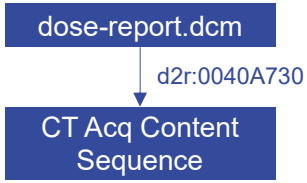
"What are the $CTDI_{vol}$ measurements of a given Structured Report?"

"What are the $CTDI_{vol}$ measurements of a given Structured Report?"

```
SELECT * WHERE {




}
```

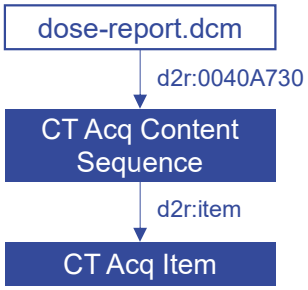
"What are the $CTDI_{vol}$ measurements of a given Structured Report?"

```
SELECT * WHERE {
    ?dose_report d2r:0040A730 ?ct_acq_cont_seq .



}
```



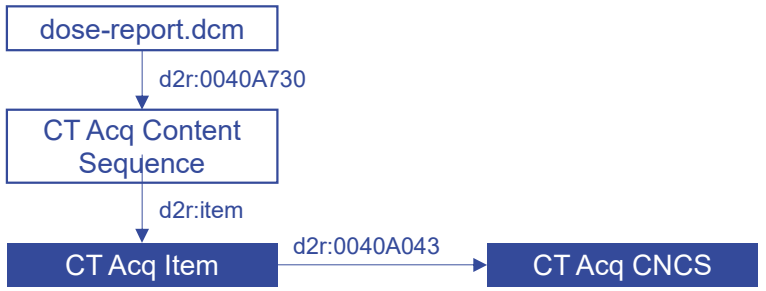
"What are the $CTDI_{vol}$ measurements of a given Structured Report?"

```
1  SELECT * WHERE {
2      ?dose_report d2r:0040A730 ?ct_acq_cont_seq .
3      ?ct_acq_cont_seq d2r:item ?ct_acq_item .
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22 }
```



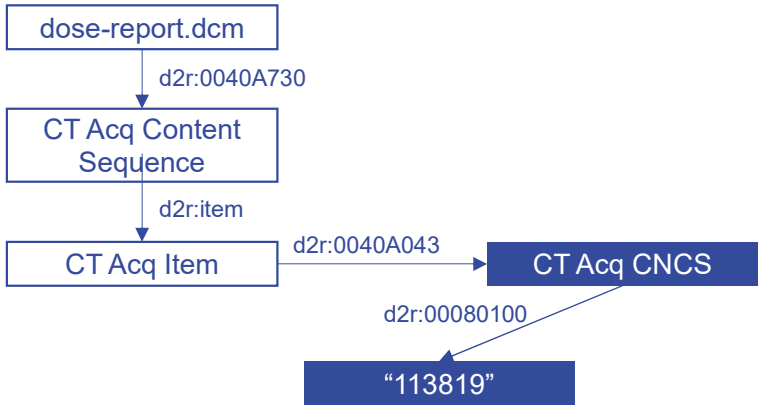
"What are the $CTDI_{vol}$ measurements of a given Structured Report?"

```
SELECT * WHERE {
    ?dose_report d2r:0040A730 ?ct_acq_cont_seq .
    ?ct_acq_cont_seq d2r:item ?ct_acq_item .
    ?ct_acq_item d2r:0040A043 ?ct_acq_cncs .
}
```



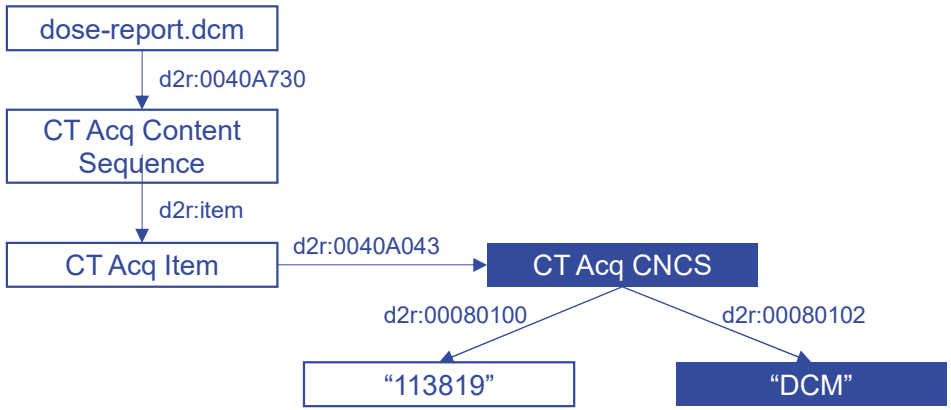
"What are the $CTDI_{vol}$ measurements of a given Structured Report?"

```
SELECT * WHERE {
    ?dose_report d2r:0040A730 ?ct_acq_cont_seq .
    ?ct_acq_cont_seq d2r:item ?ct_acq_item .
    ?ct_acq_item d2r:0040A043 ?ct_acq_cncs .
    ?ct_acq_cncs d2r:00080100 "113819" .
}
```



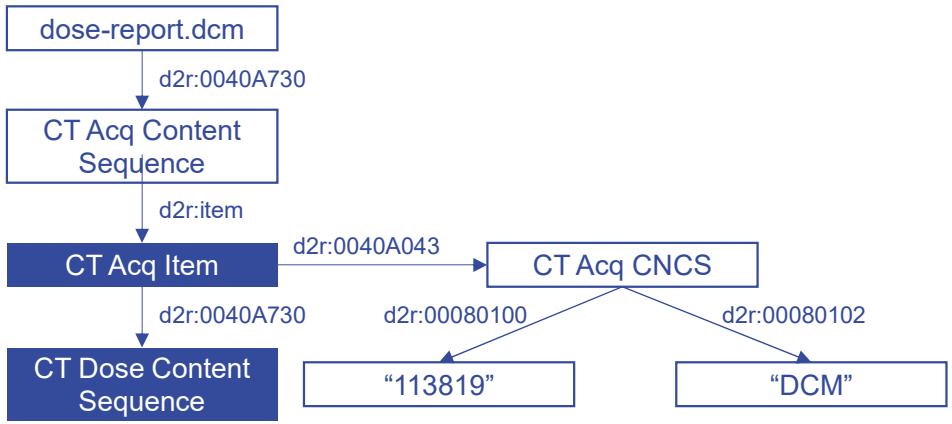
"What are the $CTDI_{vol}$ measurements of a given Structured Report?"

```
SELECT * WHERE {
    ?dose_report d2r:0040A730 ?ct_acq_cont_seq .
    ?ct_acq_cont_seq d2r:item ?ct_acq_item .
    ?ct_acq_item d2r:0040A043 ?ct_acq_cncs .
    ?ct_acq_cncs d2r:00080100 "113819" .
    ?ct_acq_cncs d2r:00080102 "DCM" .
}
```



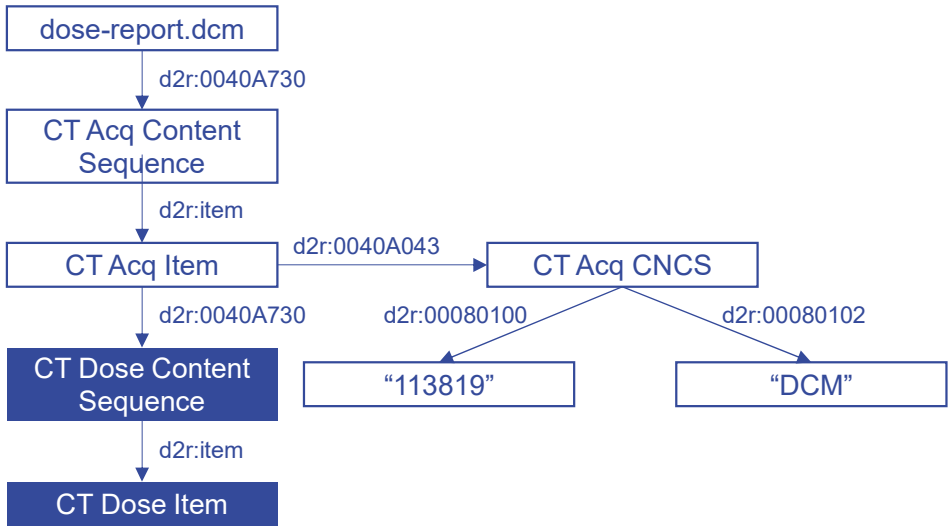
"What are the $CTDI_{vol}$ measurements of a given Structured Report?"

```
SELECT * WHERE {
    ?dose_report d2r:0040A730 ?ct_acq_cont_seq .
    ?ct_acq_cont_seq d2r:item ?ct_acq_item .
    ?ct_acq_item d2r:0040A043 ?ct_acq_cncs .
    ?ct_acq_cncs d2r:00080100 "113819" .
    ?ct_acq_cncs d2r:00080102 "DCM" .
    ?ct_acq_item d2r:0040A730 ?ct_dose_cont_seq .
}
```



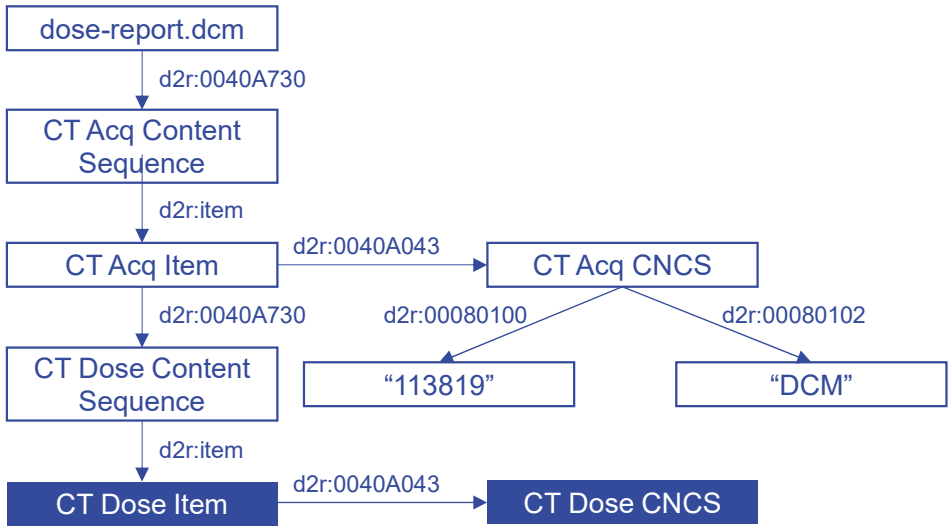
"What are the $CTDI_{vol}$ measurements of a given Structured Report?"

```
SELECT * WHERE {
    ?dose_report d2r:0040A730 ?ct_acq_cont_seq .
    ?ct_acq_cont_seq d2r:item ?ct_acq_item .
    ?ct_acq_item d2r:0040A043 ?ct_acq_cncs .
    ?ct_acq_cncs d2r:00080100 "113819" .
    ?ct_acq_cncs d2r:00080102 "DCM" .
    ?ct_acq_item d2r:0040A730 ?ct_dose_cont_seq .
    ?ct_dose_cont_seq d2r:item ?ct_dose_item .



}
```



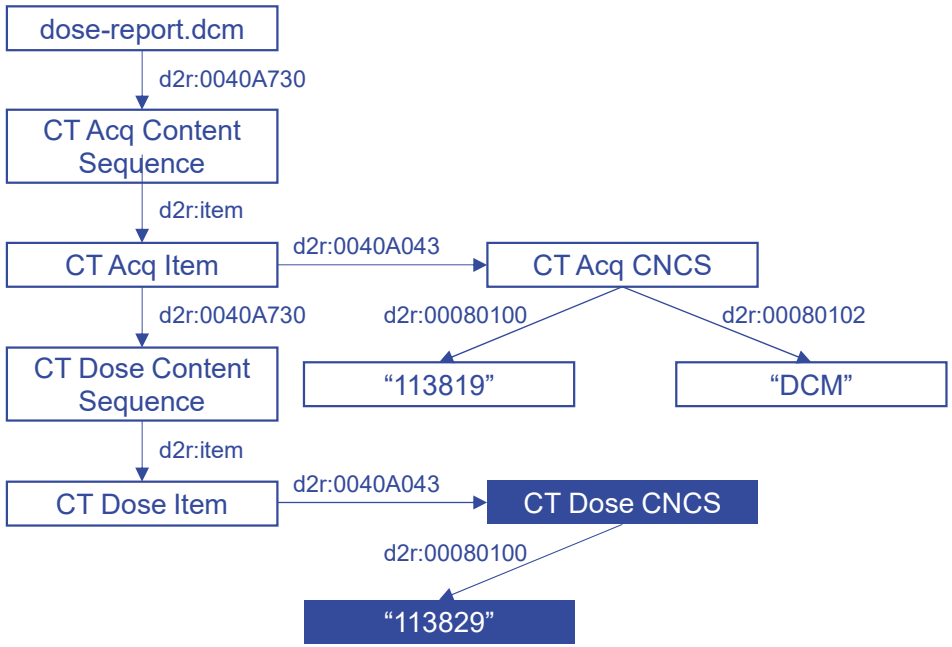
"What are the $CTDI_{vol}$ measurements of a given Structured Report?"

```
SELECT * WHERE {
    ?dose_report d2r:0040A730 ?ct_acq_cont_seq .
    ?ct_acq_cont_seq d2r:item ?ct_acq_item .
    ?ct_acq_item d2r:0040A043 ?ct_acq_cncs .
    ?ct_acq_cncs d2r:00080100 "113819" .
    ?ct_acq_cncs d2r:00080102 "DCM" .
    ?ct_acq_item d2r:0040A730 ?ct_dose_cont_seq .
    ?ct_dose_cont_seq d2r:item ?ct_dose_item .
    ?ct_dose_item d2r:0040A043 ?ct_dose_cncs .



}
```



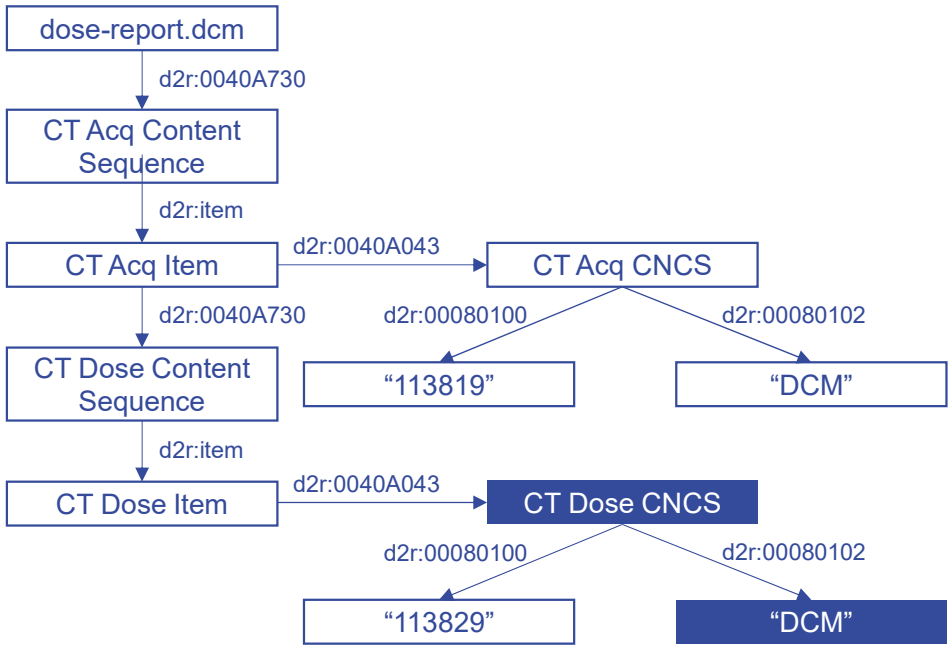
"What are the $CTDI_{vol}$ measurements of a given Structured Report?"

```
SELECT * WHERE {
    ?dose_report d2r:0040A730 ?ct_acq_cont_seq .
    ?ct_acq_cont_seq d2r:item ?ct_acq_item .
    ?ct_acq_item d2r:0040A043 ?ct_acq_cncs .
    ?ct_acq_cncs d2r:00080100 "113819" .
    ?ct_acq_cncs d2r:00080102 "DCM" .
    ?ct_acq_item d2r:0040A730 ?ct_dose_cont_seq .
    ?ct_dose_cont_seq d2r:item ?ct_dose_item .
    ?ct_dose_item d2r:0040A043 ?ct_dose_cncs .
    ?ct_dose_cncs d2r:00080100 "113829" .


}
```



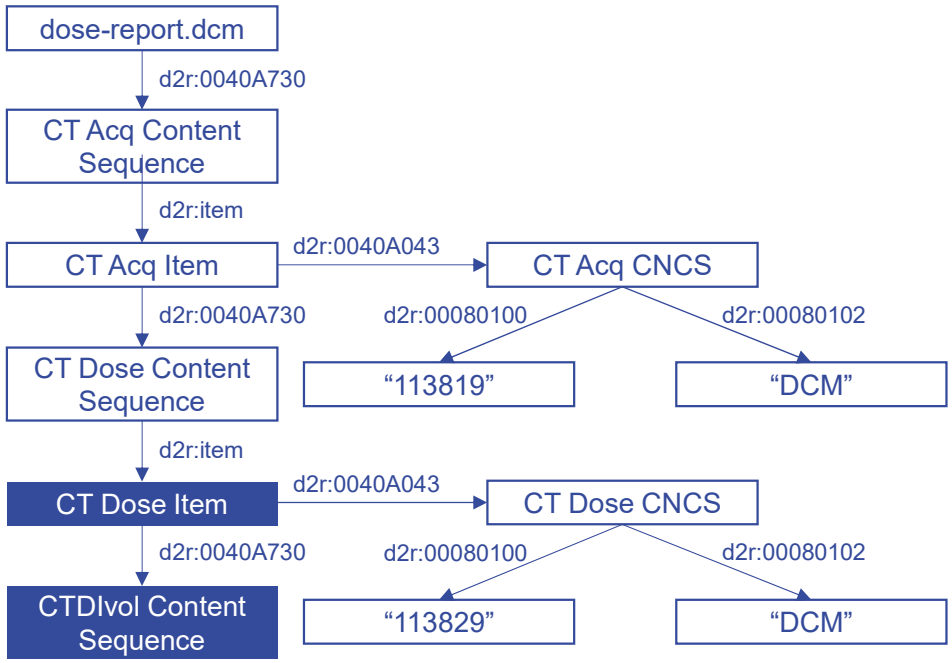
"What are the $CTDI_{vol}$ measurements of a given Structured Report?"

```
SELECT * WHERE {
    ?dose_report d2r:0040A730 ?ct_acq_cont_seq .
    ?ct_acq_cont_seq d2r:item ?ct_acq_item .
    ?ct_acq_item d2r:0040A043 ?ct_acq_cncs .
    ?ct_acq_cncs d2r:00080100 "113819" .
    ?ct_acq_cncs d2r:00080102 "DCM" .
    ?ct_acq_item d2r:0040A730 ?ct_dose_cont_seq .
    ?ct_dose_cont_seq d2r:item ?ct_dose_item .
    ?ct_dose_item d2r:0040A043 ?ct_dose_cncs .
    ?ct_dose_cncs d2r:00080100 "113829" .
    ?ct_dose_cncs d2r:00080102 "DCM" .


}
```



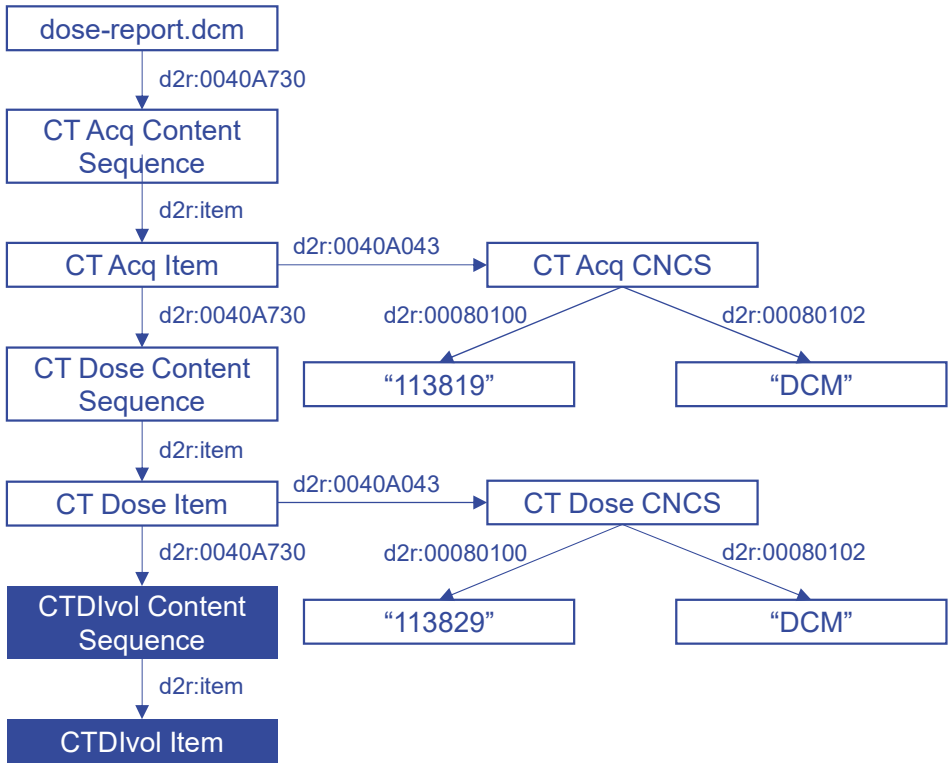
"What are the $CTDI_{vol}$ measurements of a given Structured Report?"

```
1  SELECT * WHERE {
2      ?dose_report d2r:0040A730 ?ct_acq_cont_seq .
3      ?ct_acq_cont_seq d2r:item ?ct_acq_item .
4      ?ct_acq_item d2r:0040A043 ?ct_acq_cncs .
5      ?ct_acq_cncs d2r:00080100 "113819" .
6      ?ct_acq_cncs d2r:00080102 "DCM" .
7      ?ct_acq_item d2r:0040A730 ?ct_dose_cont_seq .
8      ?ct_dose_cont_seq d2r:item ?ct_dose_item .
9      ?ct_dose_item d2r:0040A043 ?ct_dose_cncs .
10     ?ct_dose_cncs d2r:00080100 "113829" .
11     ?ct_dose_cncs d2r:00080102 "DCM" .
12     ?ct_dose_item d2r:0040A730 ?ctdivol_cont_seq .
13
14
15
16
17
18
19
20
21
22 }
```



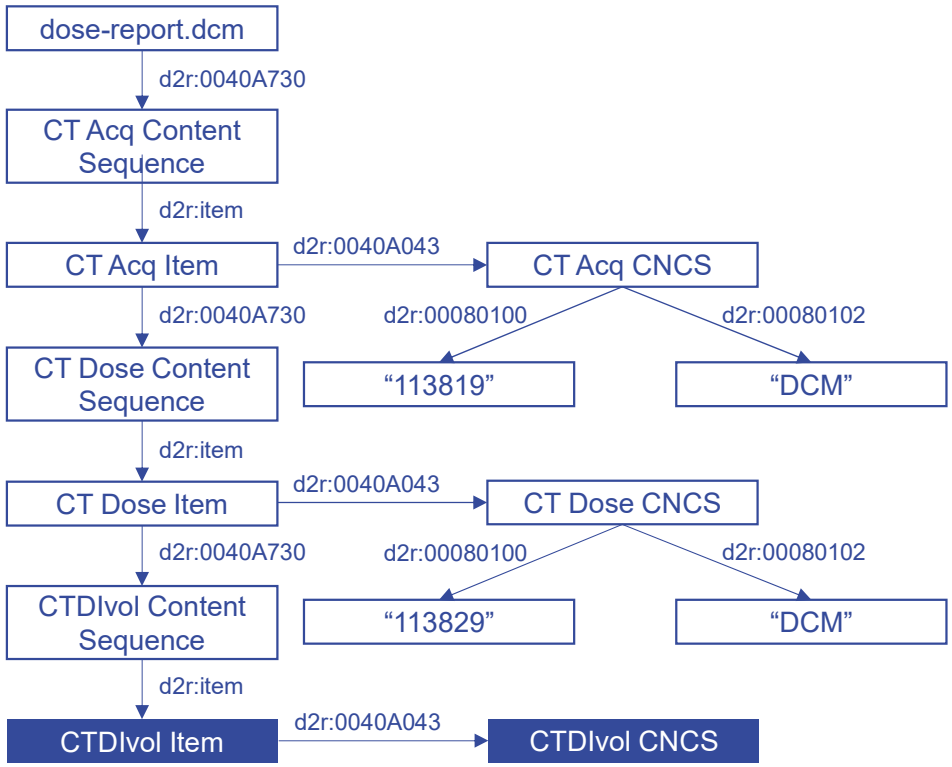
"What are the $CTDI_{vol}$ measurements of a given Structured Report?"

```
1  SELECT * WHERE {
2      ?dose_report d2r:0040A730 ?ct_acq_cont_seq .
3      ?ct_acq_cont_seq d2r:item ?ct_acq_item .
4      ?ct_acq_item d2r:0040A043 ?ct_acq_cncs .
5      ?ct_acq_cncs d2r:00080100 "113819" .
6      ?ct_acq_cncs d2r:00080102 "DCM" .
7      ?ct_acq_item d2r:0040A730 ?ct_dose_cont_seq .
8      ?ct_dose_cont_seq d2r:item ?ct_dose_item .
9      ?ct_dose_item d2r:0040A043 ?ct_dose_cncs .
10     ?ct_dose_cncs d2r:00080100 "113829" .
11     ?ct_dose_cncs d2r:00080102 "DCM" .
12     ?ct_dose_item d2r:0040A730 ?ctdivol_cont_seq .
13     ?ctdivol_cont_seq d2r:item ?ctdivol_item .
14
15
16
17
18
19
20
21
22 }
```



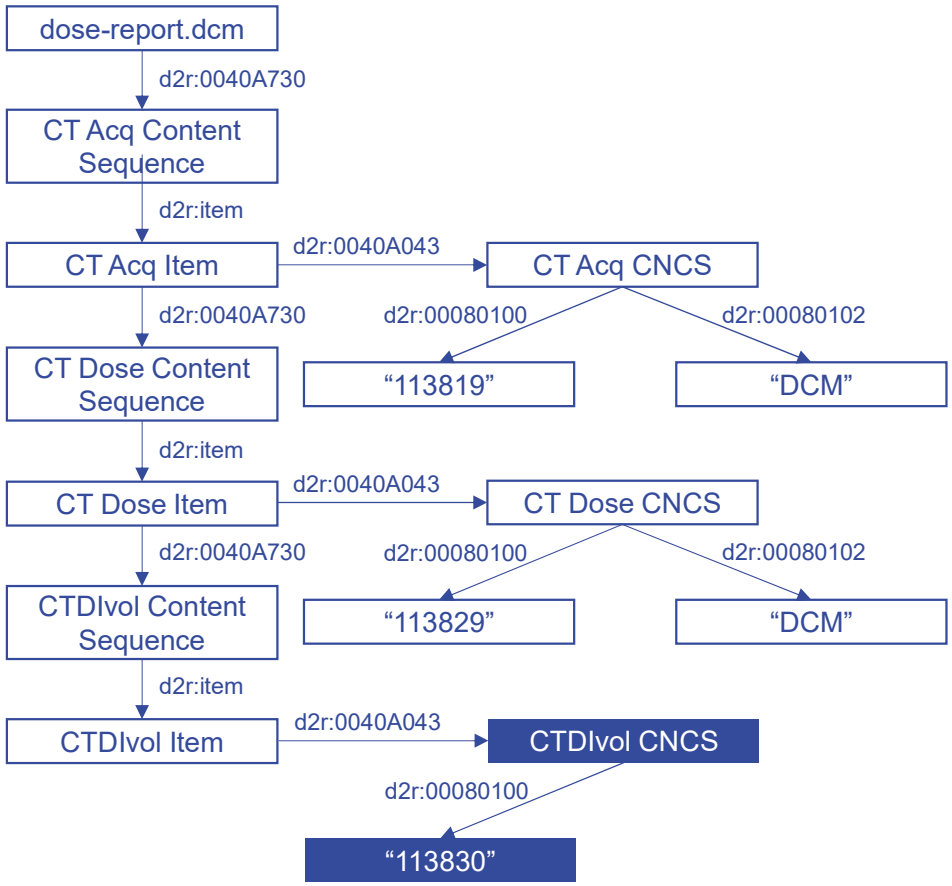
"What are the $CTDI_{vol}$ measurements of a given Structured Report?"

```
1  SELECT * WHERE {
2      ?dose_report d2r:0040A730 ?ct_acq_cont_seq .
3      ?ct_acq_cont_seq d2r:item ?ct_acq_item .
4      ?ct_acq_item d2r:0040A043 ?ct_acq_cncs .
5      ?ct_acq_cncs d2r:00080100 "113819" .
6      ?ct_acq_cncs d2r:00080102 "DCM" .
7      ?ct_acq_item d2r:0040A730 ?ct_dose_cont_seq .
8      ?ct_dose_cont_seq d2r:item ?ct_dose_item .
9      ?ct_dose_item d2r:0040A043 ?ct_dose_cncs .
10     ?ct_dose_cncs d2r:00080100 "113829" .
11     ?ct_dose_cncs d2r:00080102 "DCM" .
12     ?ct_dose_item d2r:0040A730 ?ctdivol_cont_seq .
13     ?ctdivol_cont_seq d2r:item ?ctdivol_item .
14     ?ctdivol_item d2r:0040A043 ?ctdivol_cncs .
15
16
17
18
19
20
21
22 }
```



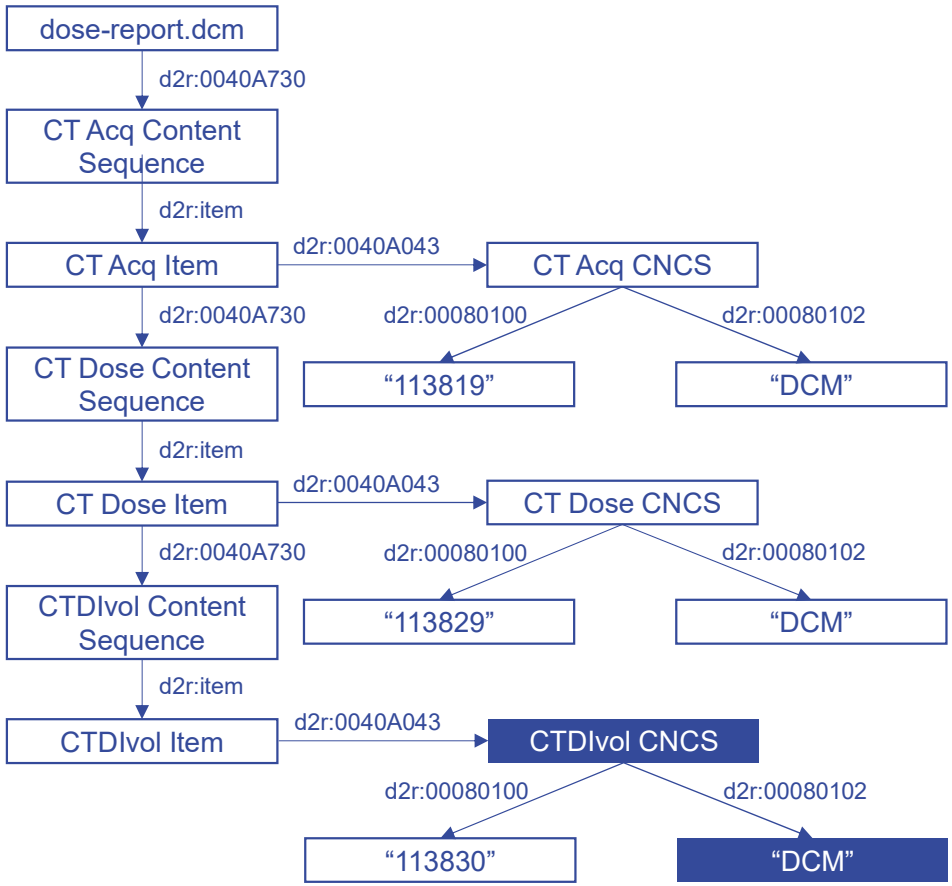
"What are the $CTDI_{vol}$ measurements of a given Structured Report?"

```
SELECT * WHERE {
    ?dose_report d2r:0040A730 ?ct_acq_cont_seq .
    ?ct_acq_cont_seq d2r:item ?ct_acq_item .
    ?ct_acq_item d2r:0040A043 ?ct_acq_cncs .
    ?ct_acq_cncs d2r:00080100 "113819" .
    ?ct_acq_cncs d2r:00080102 "DCM" .
    ?ct_acq_item d2r:0040A730 ?ct_dose_cont_seq .
    ?ct_dose_cont_seq d2r:item ?ct_dose_item .
    ?ct_dose_item d2r:0040A043 ?ct_dose_cncs .
    ?ct_dose_cncs d2r:00080100 "113829" .
    ?ct_dose_cncs d2r:00080102 "DCM" .
    ?ct_dose_item d2r:0040A730 ?ctdivol_cont_seq .
    ?ctdivol_cont_seq d2r:item ?ctdivol_item .
    ?ctdivol_item d2r:0040A043 ?ctdivol_cncs .
    ?ctdivol_cncs d2r:00080100 "113830" .



}
```



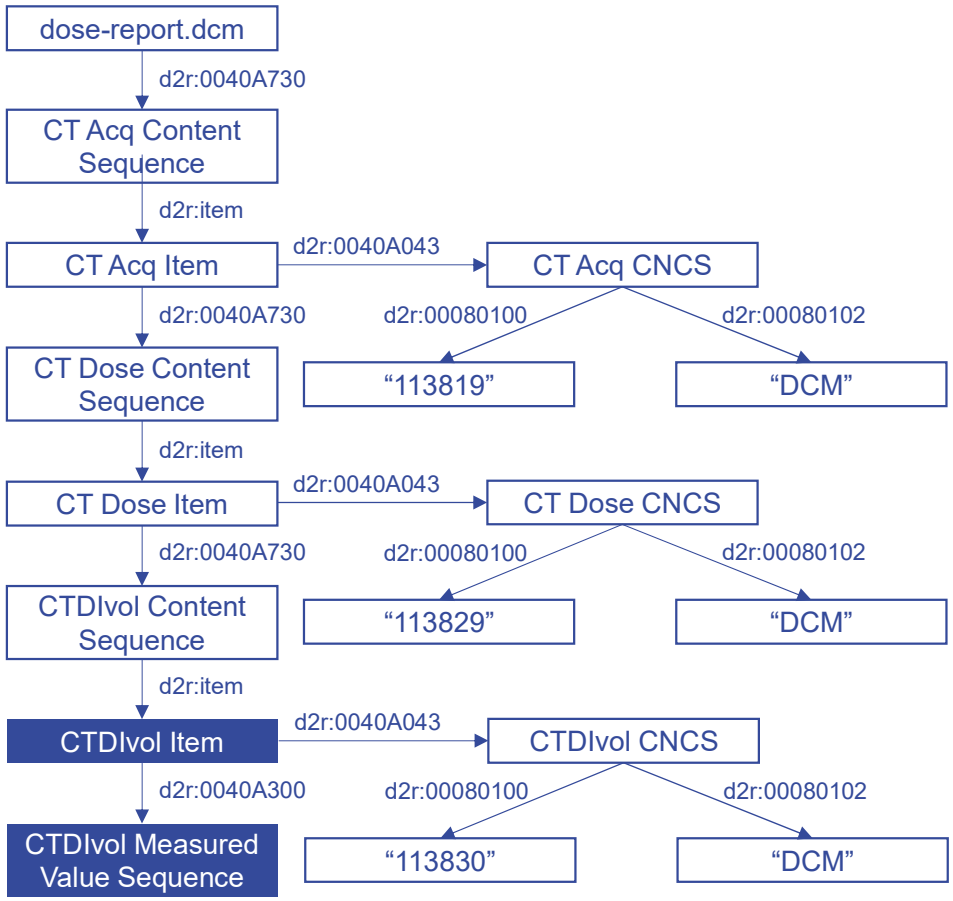
"What are the $CTDI_{vol}$ measurements of a given Structured Report?"

```
SELECT * WHERE {
    ?dose_report d2r:0040A730 ?ct_acq_cont_seq .
    ?ct_acq_cont_seq d2r:item ?ct_acq_item .
    ?ct_acq_item d2r:0040A043 ?ct_acq_cncs .
    ?ct_acq_cncs d2r:00080100 "113819" .
    ?ct_acq_cncs d2r:00080102 "DCM" .
    ?ct_acq_item d2r:0040A730 ?ct_dose_cont_seq .
    ?ct_dose_cont_seq d2r:item ?ct_dose_item .
    ?ct_dose_item d2r:0040A043 ?ct_dose_cncs .
    ?ct_dose_cncs d2r:00080100 "113829" .
    ?ct_dose_cncs d2r:00080102 "DCM" .
    ?ct_dose_item d2r:0040A730 ?ctdivol_cont_seq .
    ?ctdivol_cont_seq d2r:item ?ctdivol_item .
    ?ctdivol_item d2r:0040A043 ?ctdivol_cncs .
    ?ctdivol_cncs d2r:00080100 "113830" .
    ?ctdivol_cncs d2r:00080102 "DCM" .


}
```



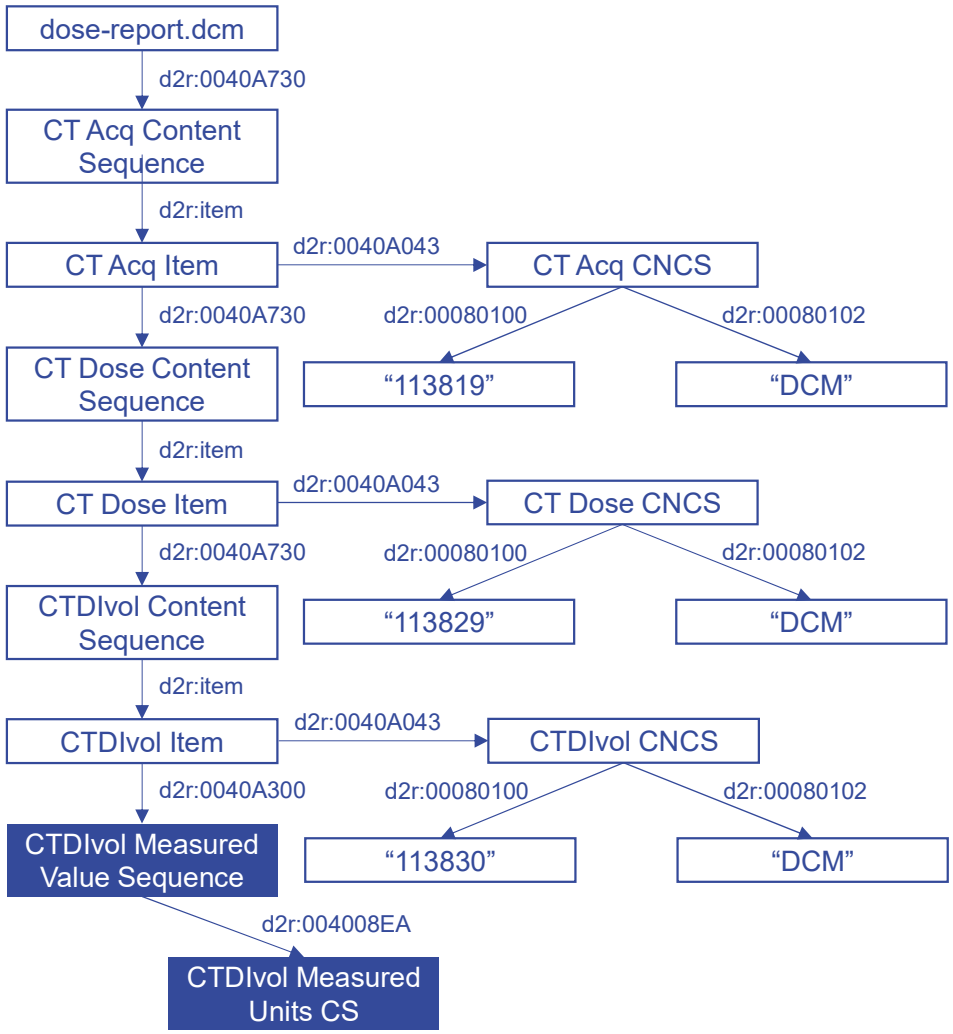
"What are the $CTDI_{vol}$ measurements of a given Structured Report?"

```
SELECT * WHERE {
    ?dose_report d2r:0040A730 ?ct_acq_cont_seq .
    ?ct_acq_cont_seq d2r:item ?ct_acq_item .
    ?ct_acq_item d2r:0040A043 ?ct_acq_cncs .
    ?ct_acq_cncs d2r:00080100 "113819" .
    ?ct_acq_cncs d2r:00080102 "DCM" .
    ?ct_acq_item d2r:0040A730 ?ct_dose_cont_seq .
    ?ct_dose_cont_seq d2r:item ?ct_dose_item .
    ?ct_dose_item d2r:0040A043 ?ct_dose_cncs .
    ?ct_dose_cncs d2r:00080100 "113829" .
    ?ct_dose_cncs d2r:00080102 "DCM" .
    ?ct_dose_item d2r:0040A730 ?ctdivol_cont_seq .
    ?ctdivol_cont_seq d2r:item ?ctdivol_item .
    ?ctdivol_item d2r:0040A043 ?ctdivol_cncs .
    ?ctdivol_cncs d2r:00080100 "113830" .
    ?ctdivol_cncs d2r:00080102 "DCM" .
    ?ctdivol_item d2r:0040A300 ?ctdivol_mv_seq .



}
```



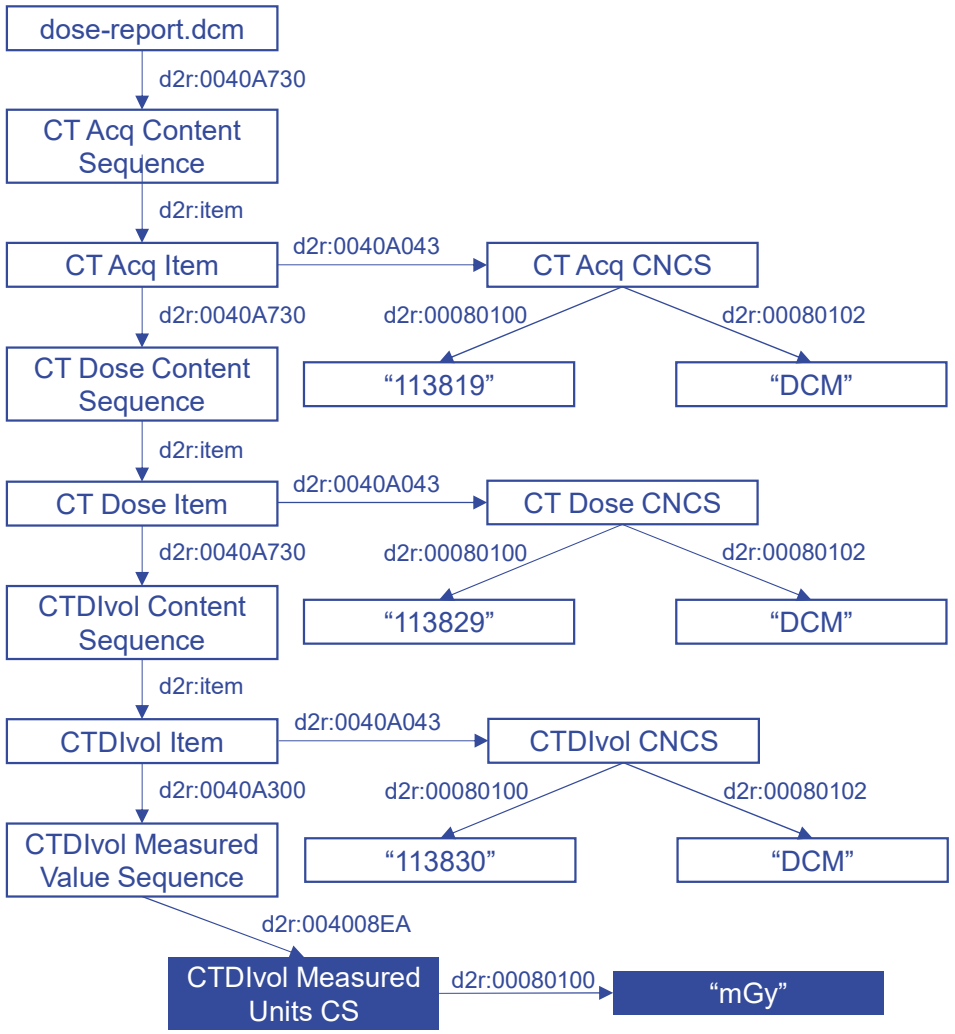
"What are the $CTDI_{vol}$ measurements of a given Structured Report?"

```
SELECT * WHERE {
    ?dose_report d2r:0040A730 ?ct_acq_cont_seq .
    ?ct_acq_cont_seq d2r:item ?ct_acq_item .
    ?ct_acq_item d2r:0040A043 ?ct_acq_cncs .
    ?ct_acq_cncs d2r:00080100 "113819" .
    ?ct_acq_cncs d2r:00080102 "DCM" .
    ?ct_acq_item d2r:0040A730 ?ct_dose_cont_seq .
    ?ct_dose_cont_seq d2r:item ?ct_dose_item .
    ?ct_dose_item d2r:0040A043 ?ct_dose_cncs .
    ?ct_dose_cncs d2r:00080100 "113829" .
    ?ct_dose_cncs d2r:00080102 "DCM" .
    ?ct_dose_item d2r:0040A730 ?ctdivol_cont_seq .
    ?ctdivol_cont_seq d2r:item ?ctdivol_item .
    ?ctdivol_item d2r:0040A043 ?ctdivol_cncs .
    ?ctdivol_cncs d2r:00080100 "113830" .
    ?ctdivol_cncs d2r:00080102 "DCM" .
    ?ctdivol_item d2r:0040A300 ?ctdivol_mv_seq .
    ?ctdivol_mv_seq d2r:004008EA ?ctdivol_mu_cs .


}
```



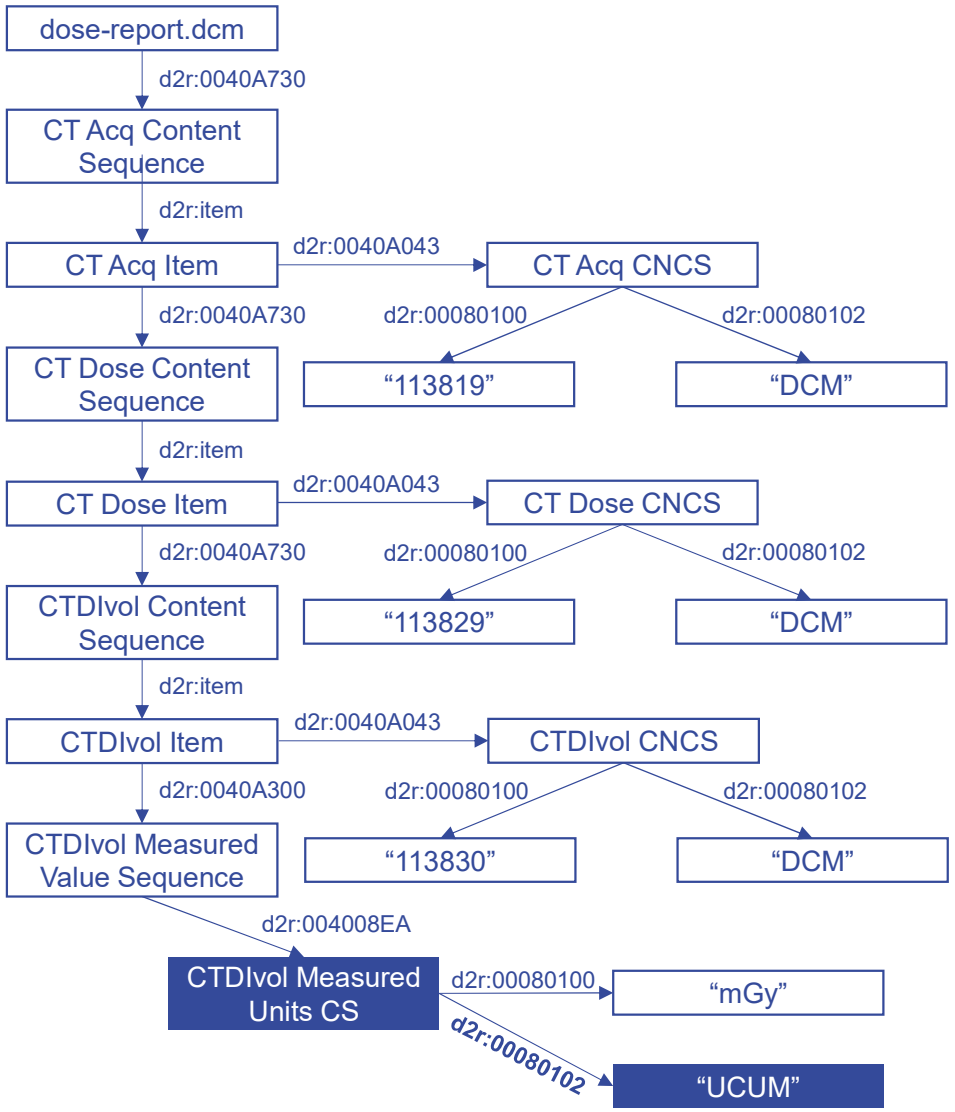
"What are the $CTDI_{vol}$ measurements of a given Structured Report?"

```
SELECT * WHERE {
    ?dose_report d2r:0040A730 ?ct_acq_cont_seq .
    ?ct_acq_cont_seq d2r:item ?ct_acq_item .
    ?ct_acq_item d2r:0040A043 ?ct_acq_cncs .
    ?ct_acq_cncs d2r:00080100 "113819" .
    ?ct_acq_cncs d2r:00080102 "DCM" .
    ?ct_acq_item d2r:0040A730 ?ct_dose_cont_seq .
    ?ct_dose_cont_seq d2r:item ?ct_dose_item .
    ?ct_dose_item d2r:0040A043 ?ct_dose_cncs .
    ?ct_dose_cncs d2r:00080100 "113829" .
    ?ct_dose_cncs d2r:00080102 "DCM" .
    ?ct_dose_item d2r:0040A730 ?ctdivol_cont_seq .
    ?ctdivol_cont_seq d2r:item ?ctdivol_item .
    ?ctdivol_item d2r:0040A043 ?ctdivol_cncs .
    ?ctdivol_cncs d2r:00080100 "113830" .
    ?ctdivol_cncs d2r:00080102 "DCM" .
    ?ctdivol_item d2r:0040A300 ?ctdivol_mv_seq .
    ?ctdivol_mv_seq d2r:004008EA ?ctdivol_mu_cs .
    ?ctdivol_mu_cs d2r:00080100 "mGy" .


}
```



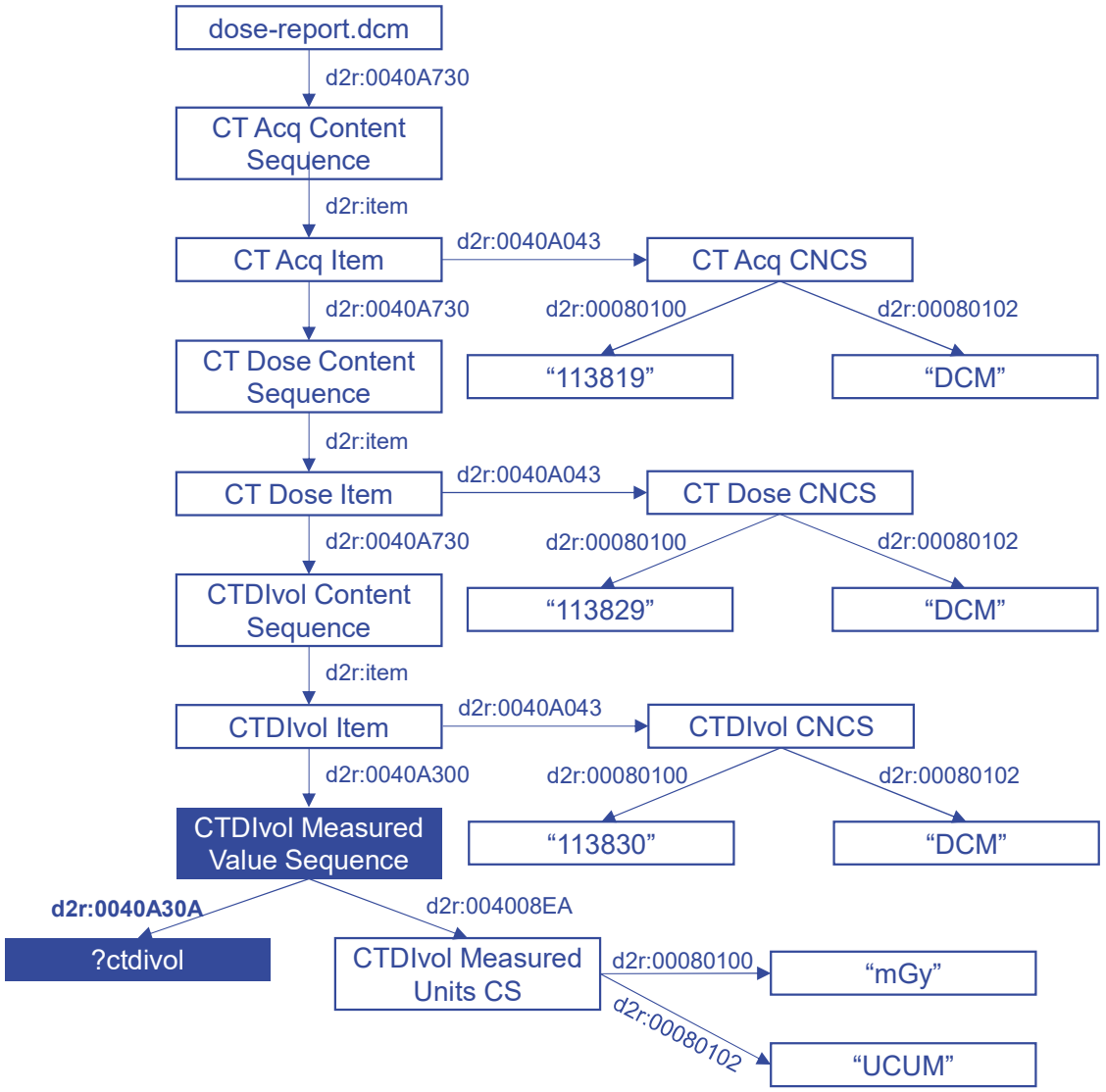
"What are the $CTDI_{vol}$ measurements of a given Structured Report?"

```
SELECT * WHERE {
    ?dose_report d2r:0040A730 ?ct_acq_cont_seq .
    ?ct_acq_cont_seq d2r:item ?ct_acq_item .
    ?ct_acq_item d2r:0040A043 ?ct_acq_cncs .
    ?ct_acq_cncs d2r:00080100 "113819" .
    ?ct_acq_cncs d2r:00080102 "DCM" .
    ?ct_acq_item d2r:0040A730 ?ct_dose_cont_seq .
    ?ct_dose_cont_seq d2r:item ?ct_dose_item .
    ?ct_dose_item d2r:0040A043 ?ct_dose_cncs .
    ?ct_dose_cncs d2r:00080100 "113829" .
    ?ct_dose_cncs d2r:00080102 "DCM" .
    ?ct_dose_item d2r:0040A730 ?ctdivol_cont_seq .
    ?ctdivol_cont_seq d2r:item ?ctdivol_item .
    ?ctdivol_item d2r:0040A043 ?ctdivol_cncs .
    ?ctdivol_cncs d2r:00080100 "113830" .
    ?ctdivol_cncs d2r:00080102 "DCM" .
    ?ctdivol_item d2r:0040A300 ?ctdivol_mv_seq .
    ?ctdivol_mv_seq d2r:004008EA ?ctdivol_mu_cs .
    ?ctdivol_mu_cs d2r:00080100 "mGy" .
    ?ctdivol_mu_cs d2r:00080102 "UCUM" .

}
```



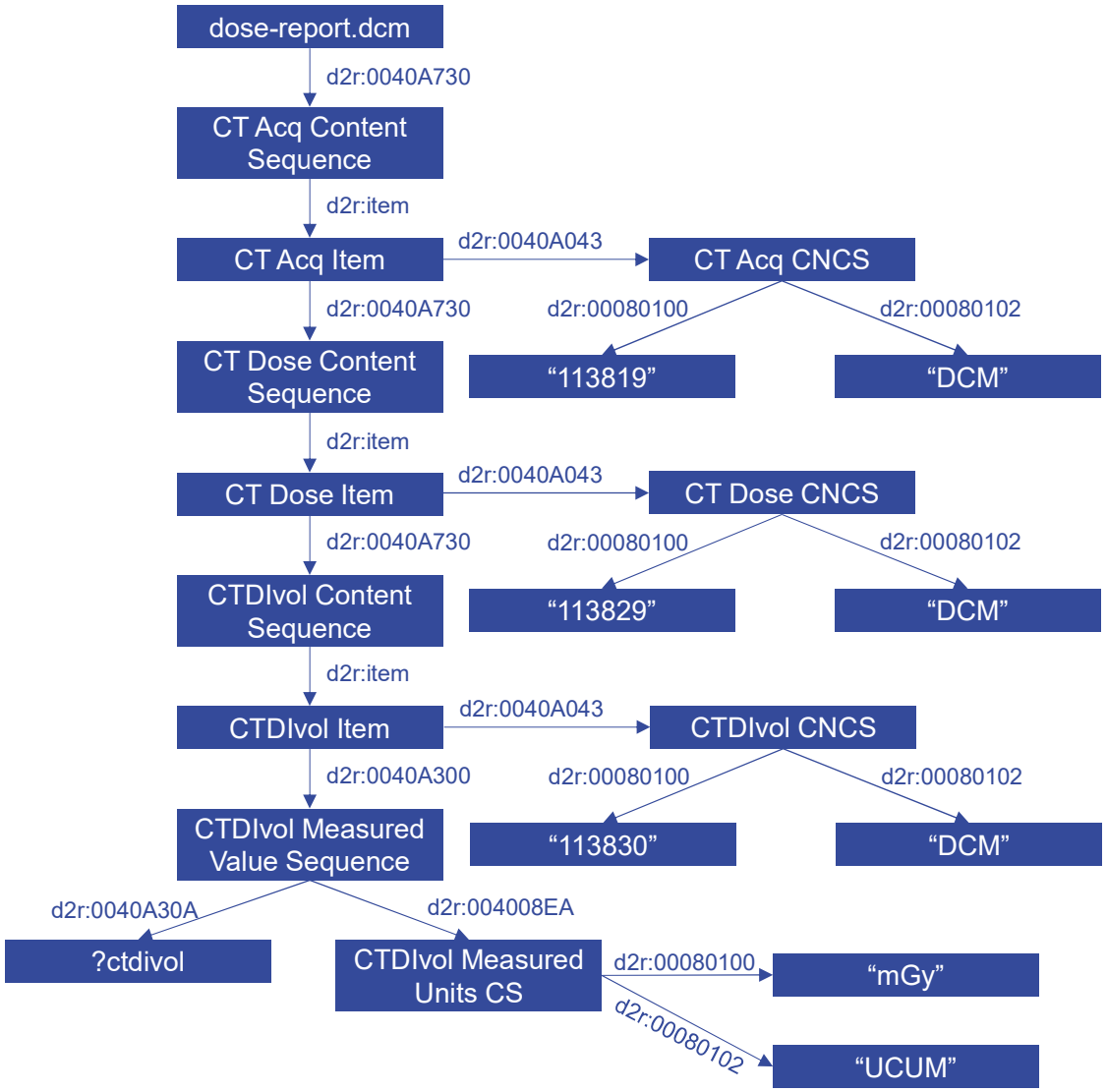
"What are the $CTDI_{vol}$ measurements of a given Structured Report?"

```
SELECT * WHERE {
    ?dose_report d2r:0040A730 ?ct_acq_cont_seq .
    ?ct_acq_cont_seq d2r:item ?ct_acq_item .
    ?ct_acq_item d2r:0040A043 ?ct_acq_cncs .
    ?ct_acq_cncs d2r:00080100 "113819" .
    ?ct_acq_cncs d2r:00080102 "DCM" .
    ?ct_acq_item d2r:0040A730 ?ct_dose_cont_seq .
    ?ct_dose_cont_seq d2r:item ?ct_dose_item .
    ?ct_dose_item d2r:0040A043 ?ct_dose_cncs .
    ?ct_dose_cncs d2r:00080100 "113829" .
    ?ct_dose_cncs d2r:00080102 "DCM" .
    ?ct_dose_item d2r:0040A730 ?ctdivol_cont_seq .
    ?ctdivol_cont_seq d2r:item ?ctdivol_item .
    ?ctdivol_item d2r:0040A043 ?ctdivol_cncs .
    ?ctdivol_cncs d2r:00080100 "113830" .
    ?ctdivol_cncs d2r:00080102 "DCM" .
    ?ctdivol_item d2r:0040A300 ?ctdivol_mv_seq .
    ?ctdivol_mv_seq d2r:004008EA ?ctdivol_mu_cs .
    ?ctdivol_mu_cs d2r:00080100 "mGy" .
    ?ctdivol_mu_cs d2r:00080102 "UCUM" .
    ?ctdivol_mv_seq d2r:0040A30A ?ctdivol .
}
```



"What are the $CTDI_{vol}$ measurements of a given Structured Report?"

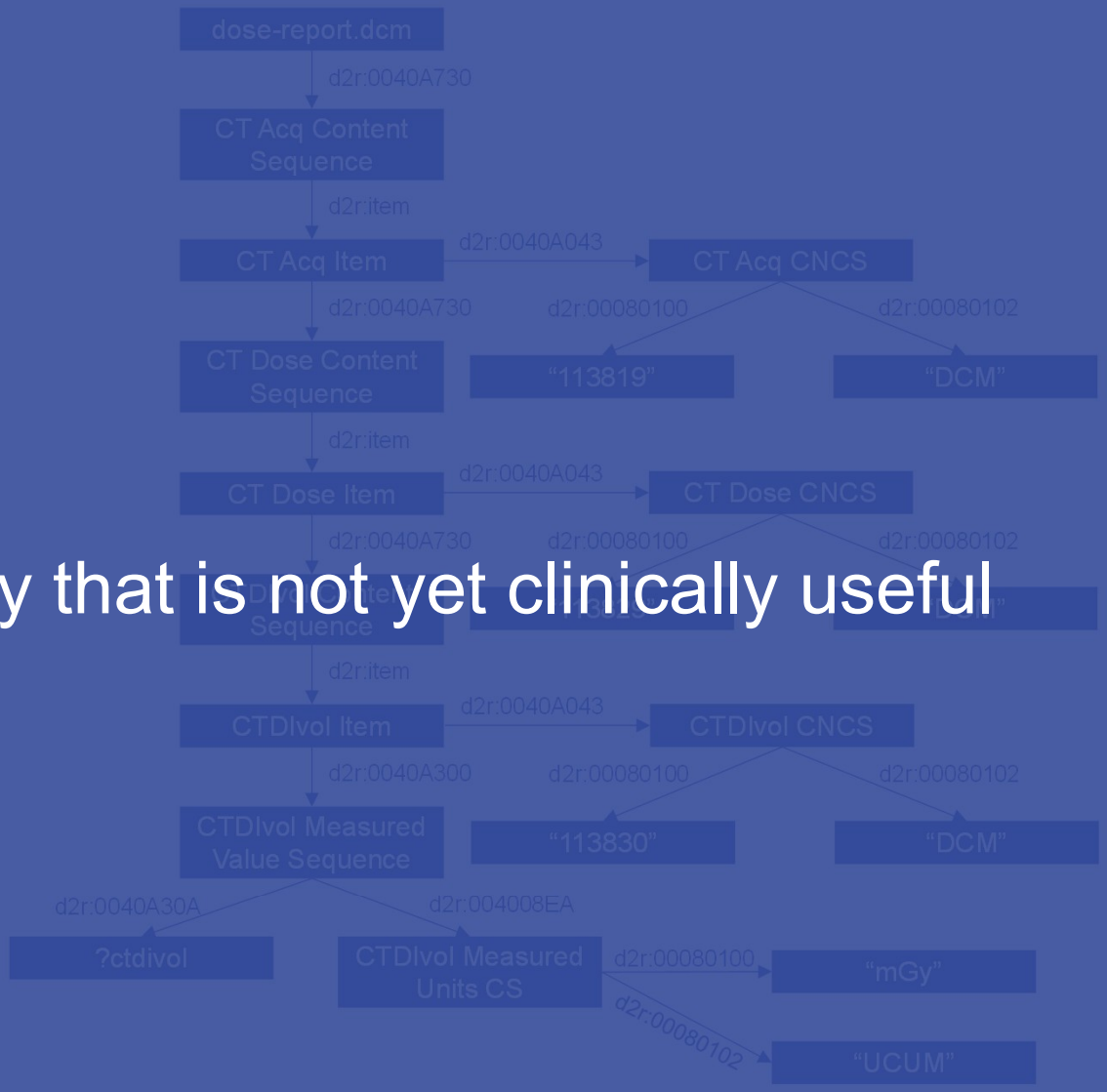
```
SELECT * WHERE {
    ?dose_report d2r:0040A730 ?ct_acq_cont_seq .
    ?ct_acq_cont_seq d2r:item ?ct_acq_item .
    ?ct_acq_item d2r:0040A043 ?ct_acq_cncs .
    ?ct_acq_cncs d2r:00080100 "113819" .
    ?ct_acq_cncs d2r:00080102 "DCM" .
    ?ct_acq_item d2r:0040A730 ?ct_dose_cont_seq .
    ?ct_dose_cont_seq d2r:item ?ct_dose_item .
    ?ct_dose_item d2r:0040A043 ?ct_dose_cncs .
    ?ct_dose_cncs d2r:00080100 "113829" .
    ?ct_dose_cncs d2r:00080102 "DCM" .
    ?ct_dose_item d2r:0040A730 ?ctdivol_cont_seq .
    ?ctdivol_cont_seq d2r:item ?ctdivol_item .
    ?ctdivol_item d2r:0040A043 ?ctdivol_cncs .
    ?ctdivol_cncs d2r:00080100 "113830" .
    ?ctdivol_cncs d2r:00080102 "DCM" .
    ?ctdivol_item d2r:0040A300 ?ctdivol_mv_seq .
    ?ctdivol_mv_seq d2r:004008EA ?ctdivol_mu_cs .
    ?ctdivol_mu_cs d2r:00080100 "mGy" .
    ?ctdivol_mu_cs d2r:00080102 "UCUM" .
    ?ctdivol_mv_seq d2r:0040A30A ?ctdivol .
}
```



"What are the $CTDI_{vol}$ measurements of a given Structured Report?"

```
SELECT * WHERE {
    ?dose_report d2r:0040A730 ?ct_acq_cont_seq .
    ?ct_acq_cont_seq d2r:item ?ct_acq_item .
    ?ct_acq_item d2r:0040A043 ?ct_acq_cncs .
    ?ct_acq_cncs d2r:00080100 "113819" .
    ?ct_acq_cncs d2r:00080102 "DCM" .
    ?ct_acq_item d2r:0040A730 ?ct_dose_cont_seq .
    ?ct_dose_cont_seq d2r:item ?ct_dose_item .
    ?ct_dose_item d2r:0040A043 ?ct_dose_cncs .
    ?ct_dose_cncs d2r:00080100 "113829" .
    ?ct_dose_item d2r:0040A730 ?ctdivol_cont_seq .
    ?ctdivol_cont_seq d2r:item ?ctdivol_item .
    ?ctdivol_item d2r:0040A043 ?ctdivol_cncs .
    ?ctdivol_cncs d2r:00080100 "113830" .
    ?ctdivol_cncs d2r:00080102 "DCM" .
    ?ctdivol_item d2r:0040A300 ?ctdivol_mv_seq .
    ?ctdivol_mv_seq d2r:004008EA ?ctdivol_mu_cs .
    ?ctdivol_mu_cs d2r:00080100 "mGy" .
    ?ctdivol_mu_cs d2r:00080102 "UCUM" .
    ?ctdivol_mv_seq d2r:0040A30A ?ctdivol .
}
```

20 triple patterns for a query that is not yet clinically useful



"What are the $CTDI_{vol}$ measurements of a given Structured Report?"

```
SELECT * WHERE {
    ?dose_report d2r:0040A730 ?ct_acq_cont_seq .
    ?ct_acq_cont_seq d2r:item ?ct_acq_item .
    ?ct_acq_item d2r:0040A043 ?ct_acq_cncs .
    ?ct_acq_cncs d2r:00080100 "113819" .
    ?ct_acq_cncs d2r:00080102 "DCM" .
    ?ct_acq_item d2r:0040A730 ?ct_dose_cont_seq .
    ?ct_dose_cont_seq d2r:item ?ct_dose_item .
    ?ct_dose_item d2r:0040A043 ?ct_dose_cncs .
    ?ct_dose_cncs d2r:00080100 "113830" .
    ?ct_dose_cncs d2r:00080102 "DCM" .
    ?ct_dose_item d2r:0040A730 ?ctdivol_cont_seq .
    ?ctdivol_cont_seq d2r:item ?ctdivol_item .
    ?ctdivol_item d2r:0040A043 ?ctdivol_cncs .
    ?ctdivol_cncs d2r:00080100 "113830" .
    ?ctdivol_cncs d2r:00080102 "DCM" .
    ?ctdivol_item d2r:0040A300 ?ctdivol_mv_seq .
    ?ctdivol_mv_seq d2r:004008EA ?ctdivol_mu_cs .
    ?ctdivol_mu_cs d2r:00080100 "mGy" .
    ?ctdivol_mu_cs d2r:00080102 "UCUM" .
    ?ctdivol_mv_seq d2r:0040A30A ?ctdivol .
}
```

20 triple patterns for a query that is not yet clinically useful

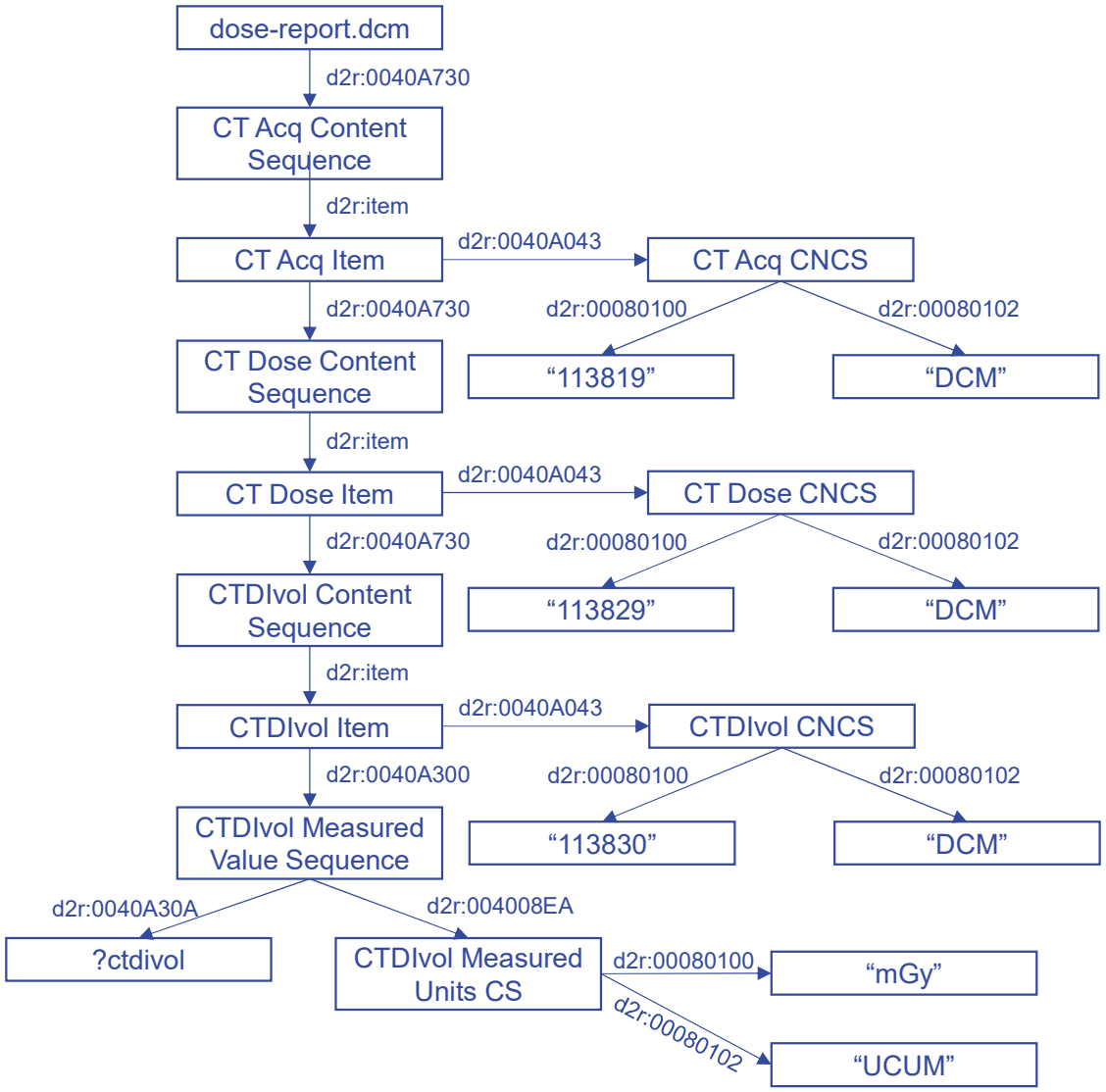
More complex queries can easily reach **over 100 triples**

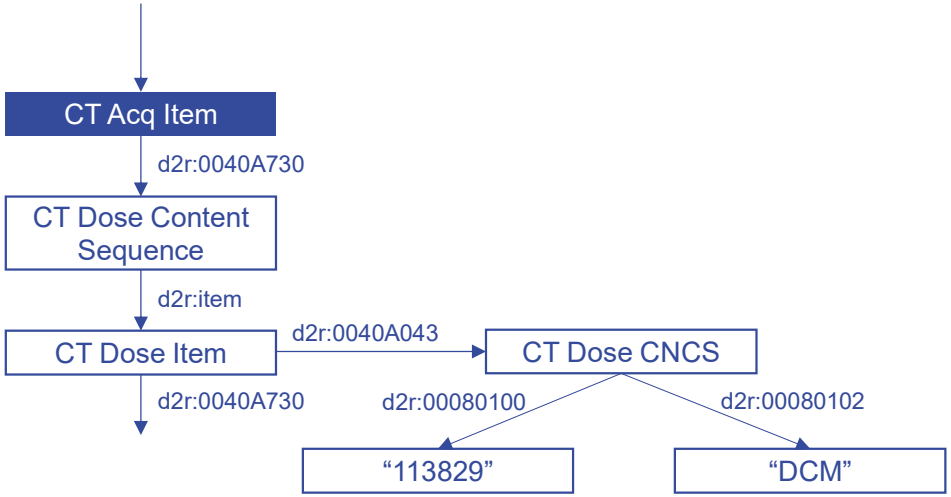
Stage 2 – *construct*

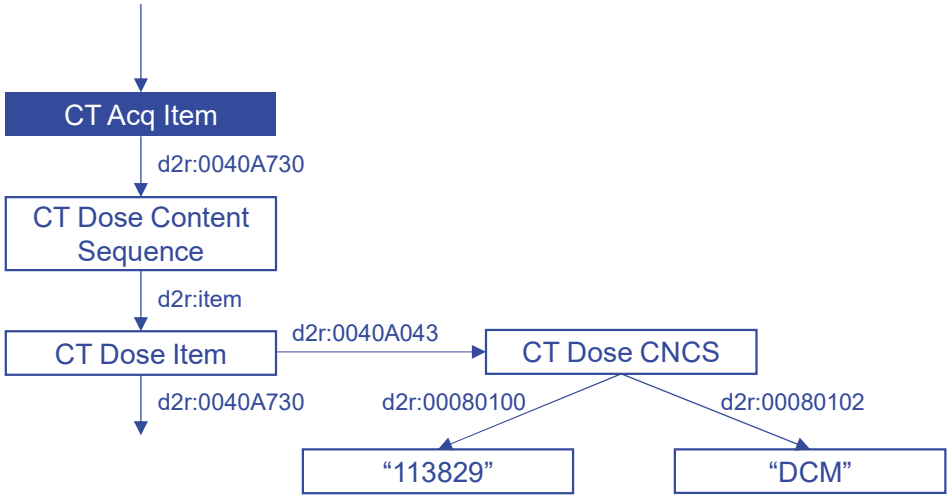
What is the problem?

What is the problem?

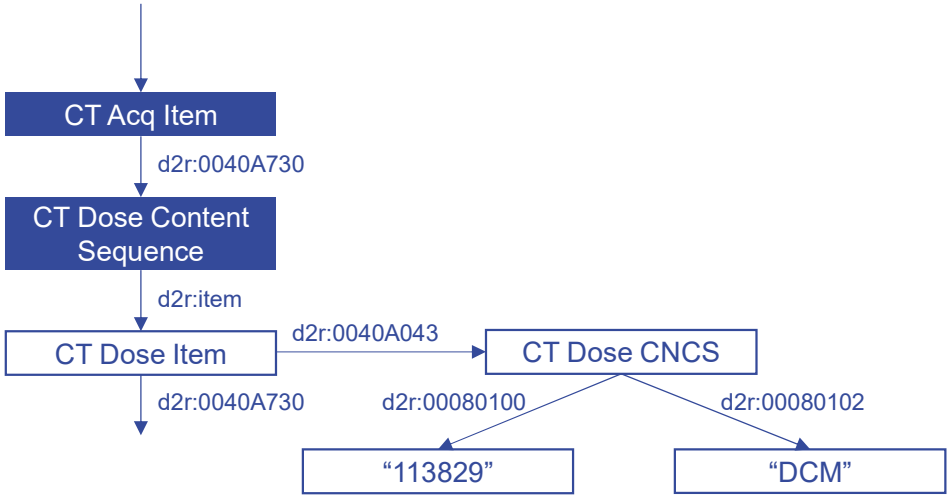
Too many hops required to jump from one concept to another


dose-report.dcm
d2r:0040A730
CT Acq Content Sequence
d2r:item
CT Acq Item
d2r:0040A043
CT Acq CNCS
d2r:0040A730
d2r:00080100
d2r:00080102
CT Dose Content Sequence
“113819”
“DCM”
d2r:item
CT Dose Item
d2r:0040A043
CT Dose CNCS
d2r:0040A730
d2r:00080100
d2r:00080102
CTDIvol Content Sequence
“113829”
“DCM”
d2r:item
CTDIvol Item
d2r:0040A043
CTDIvol CNCS
d2r:0040A300
d2r:00080100
d2r:00080102
CTDIvol Measured Value Sequence
“113830”
“DCM”
d2r:0040A30A
d2r:004008EA
?ctdivol
CTDIvol Measured Units CS
d2r:00080100
“mGy”
d2r:00080102
“UCUM”

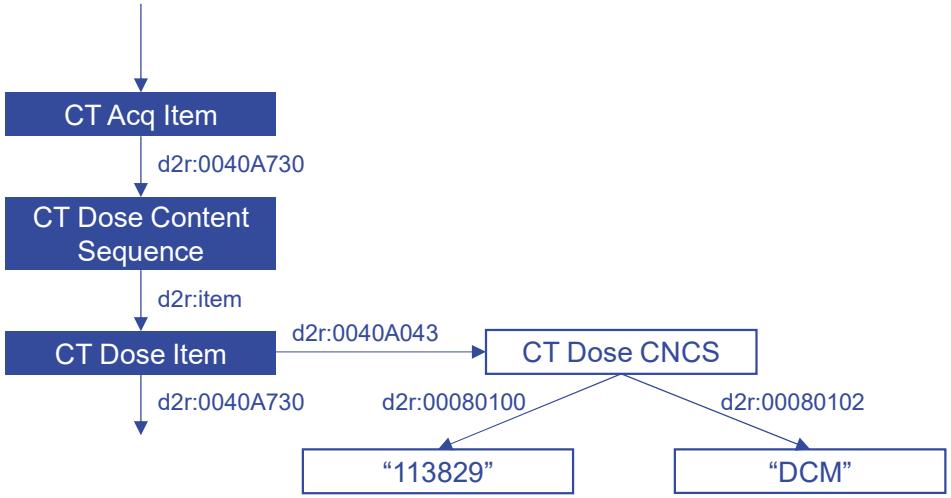

CT Acq Item
d2r:0040A730
CT Dose Content Sequence
d2r:item
CT Dose Item
d2r:0040A043
CT Dose CNCS
d2r:0040A730
d2r:00080100
d2r:00080102
“113829”
“DCM”



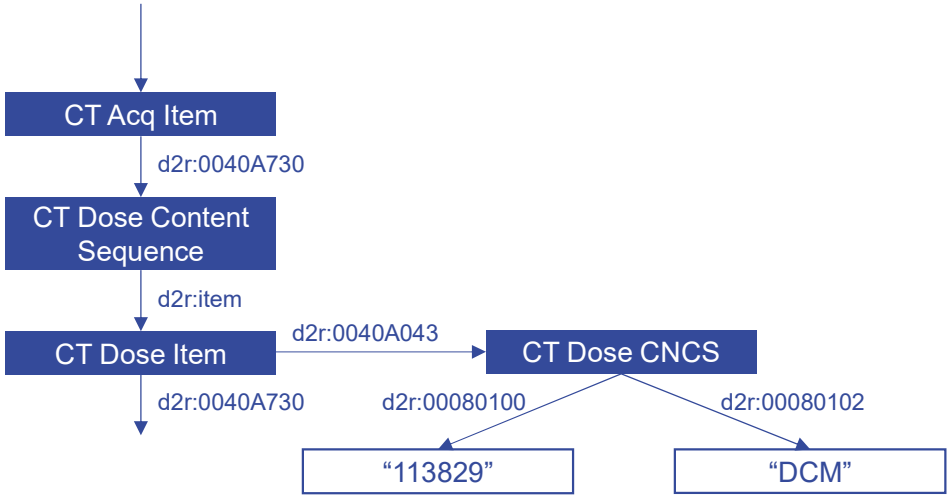
Hop Counter: 0



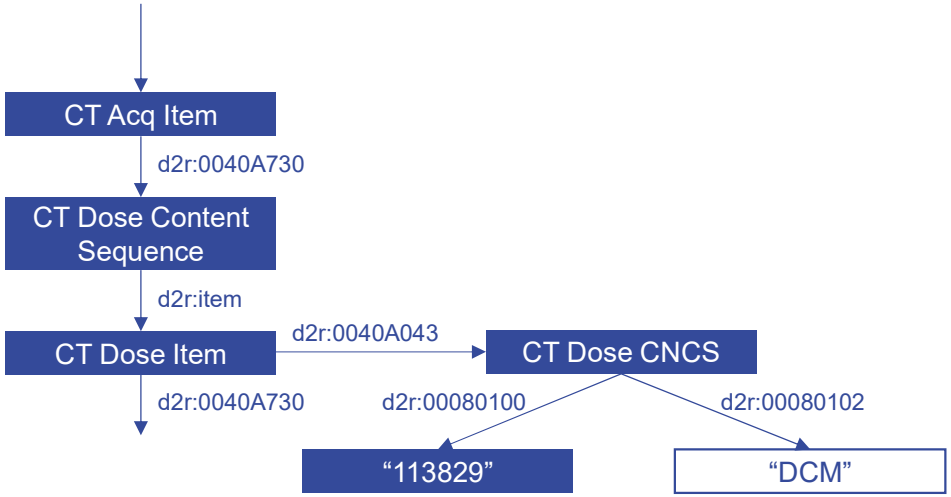
Hop Counter: 1



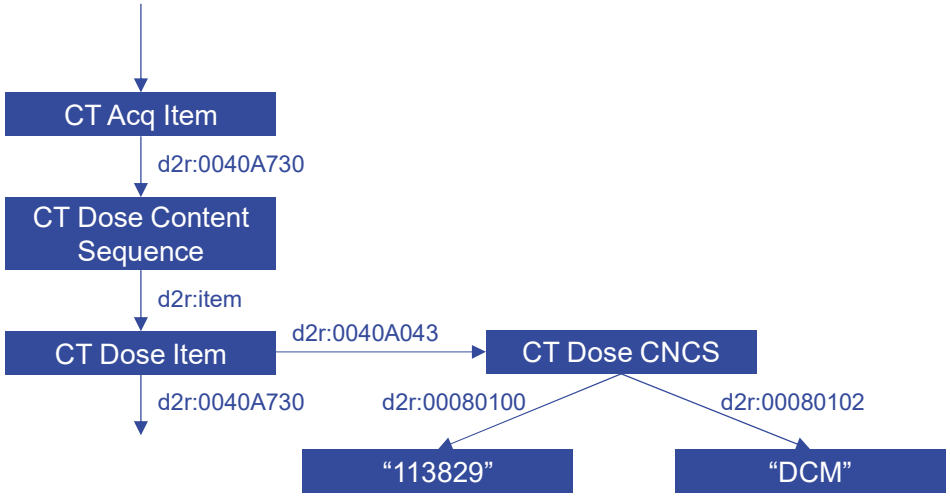
Hop Counter: 2



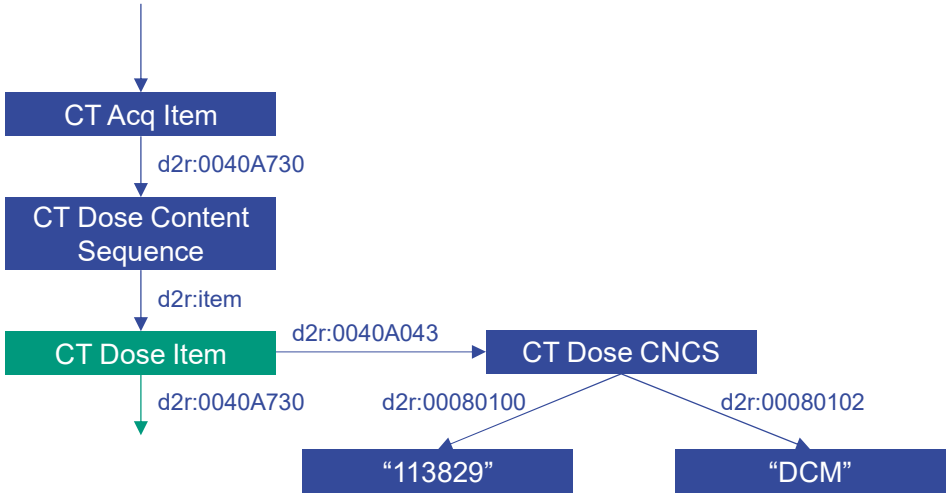
Hop Counter: 3



Hop Counter: 4

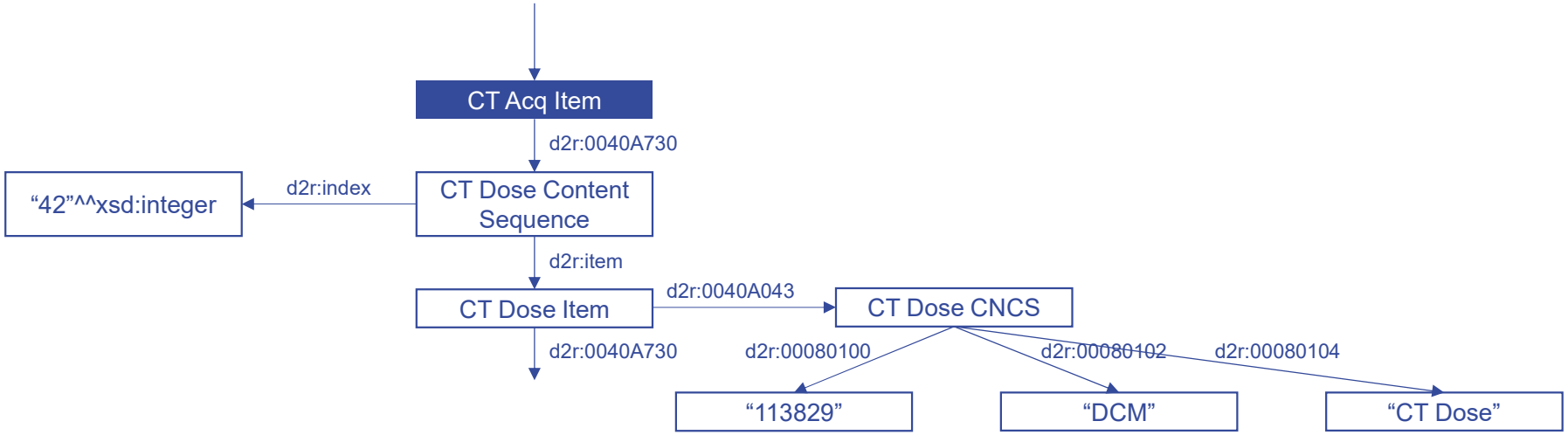


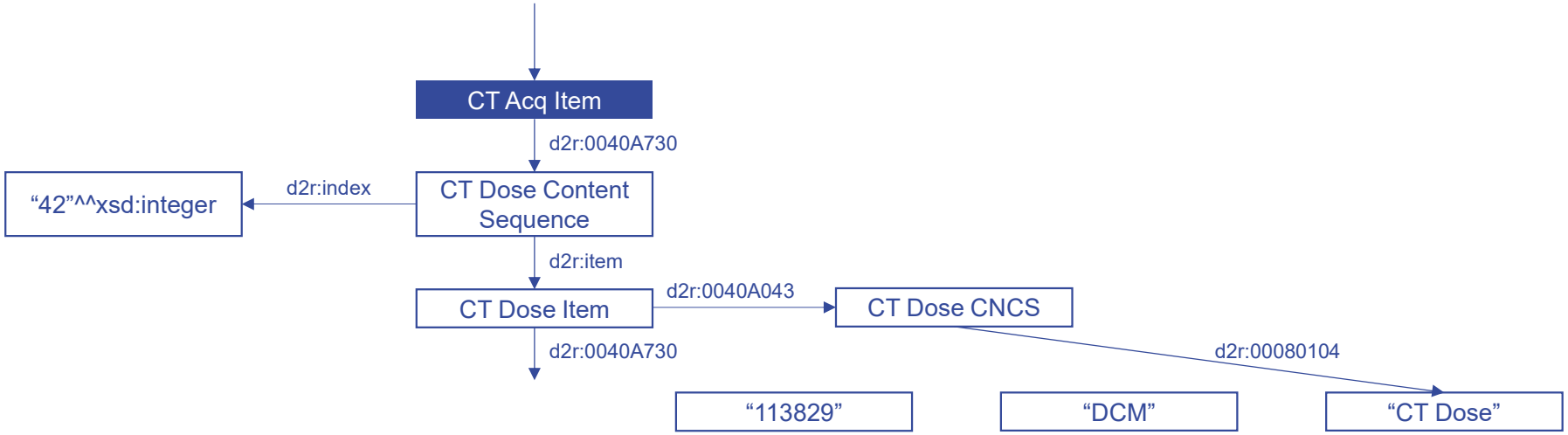
Hop Counter: 5

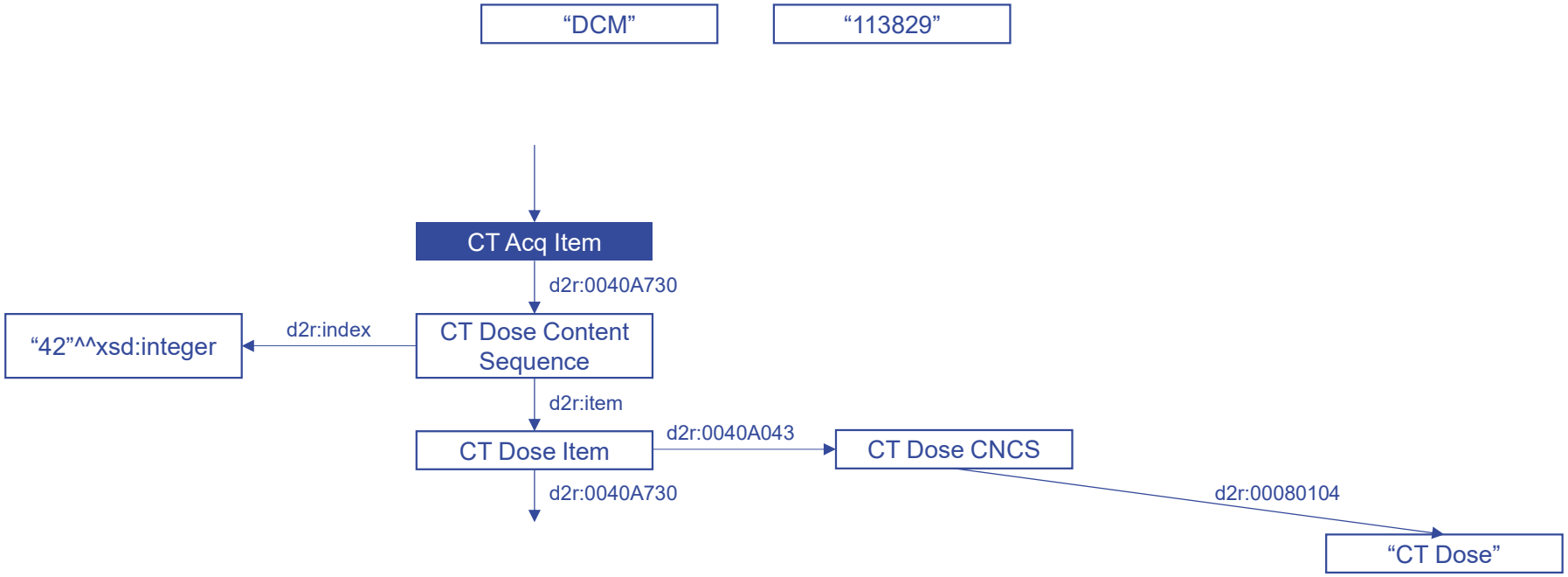


Hop Counter: 5

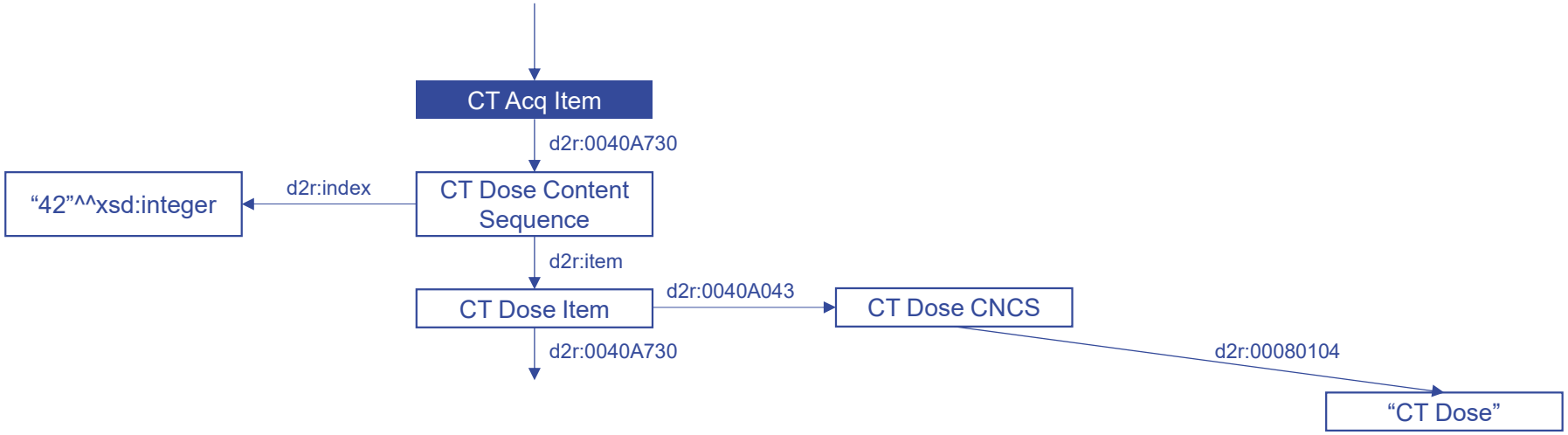
Idea: maximize information gain per hop

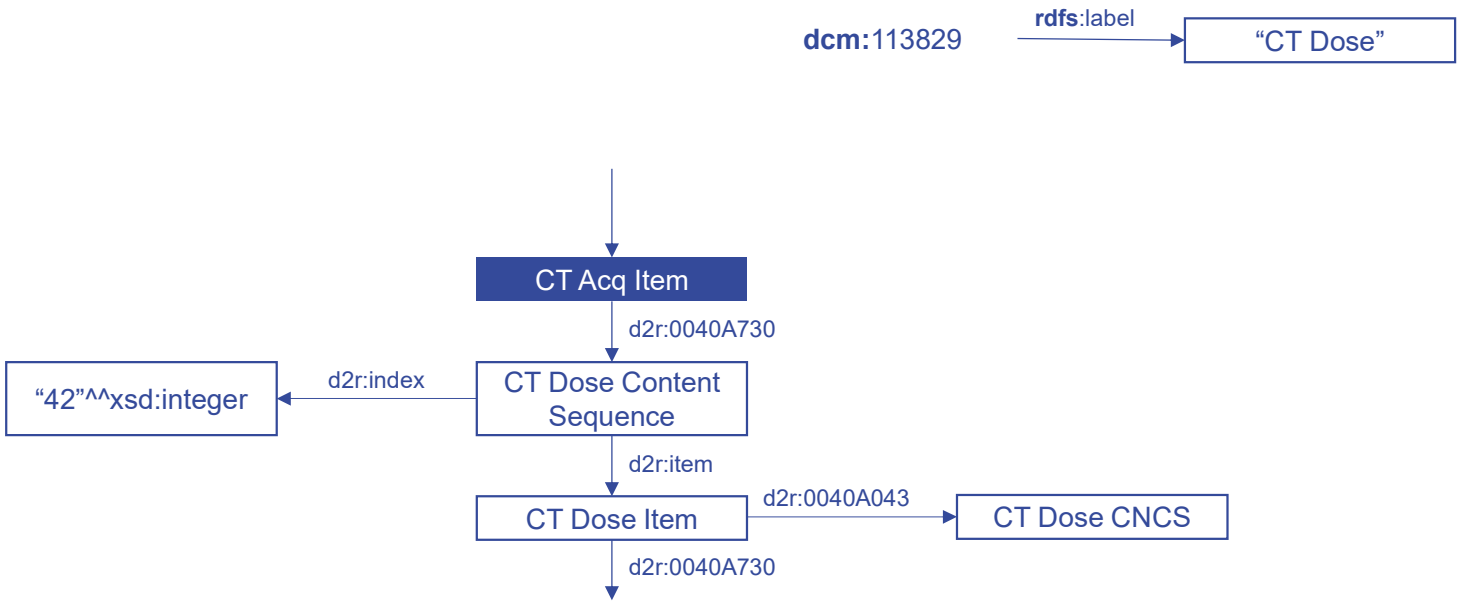

CT Acq Item
d2r:0040A730
CT Dose Content Sequence
d2r:index
"42"^^xsd:integer
d2r:item
CT Dose Item
d2r:0040A043
CT Dose CNCS
d2r:0040A730
d2r:00080100
d2r:00080102
d2r:00080104
"113829"
"DCM"
"CT Dose"

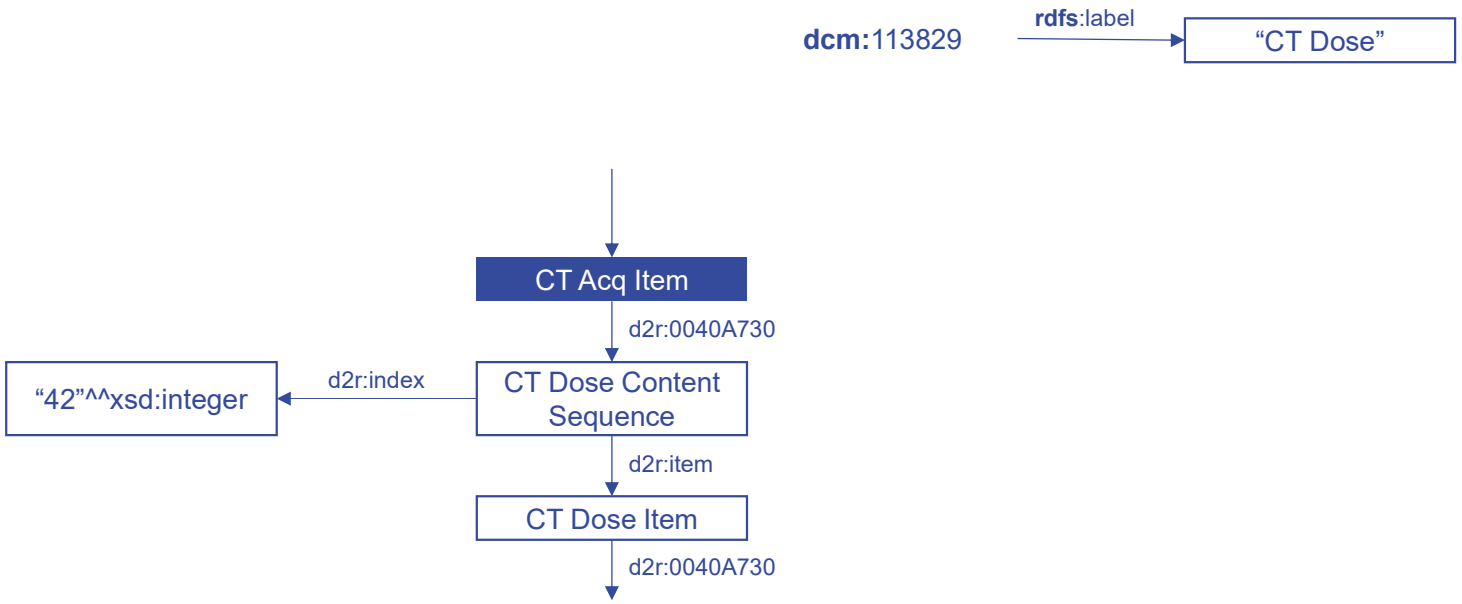

CT Acq Item
d2r:0040A730
CT Dose Content Sequence
d2r:index
"42"^^xsd:integer
d2r:item
CT Dose Item
d2r:0040A043
CT Dose CNCS
d2r:0040A730
d2r:00080104
"113829"
"DCM"
"CT Dose"

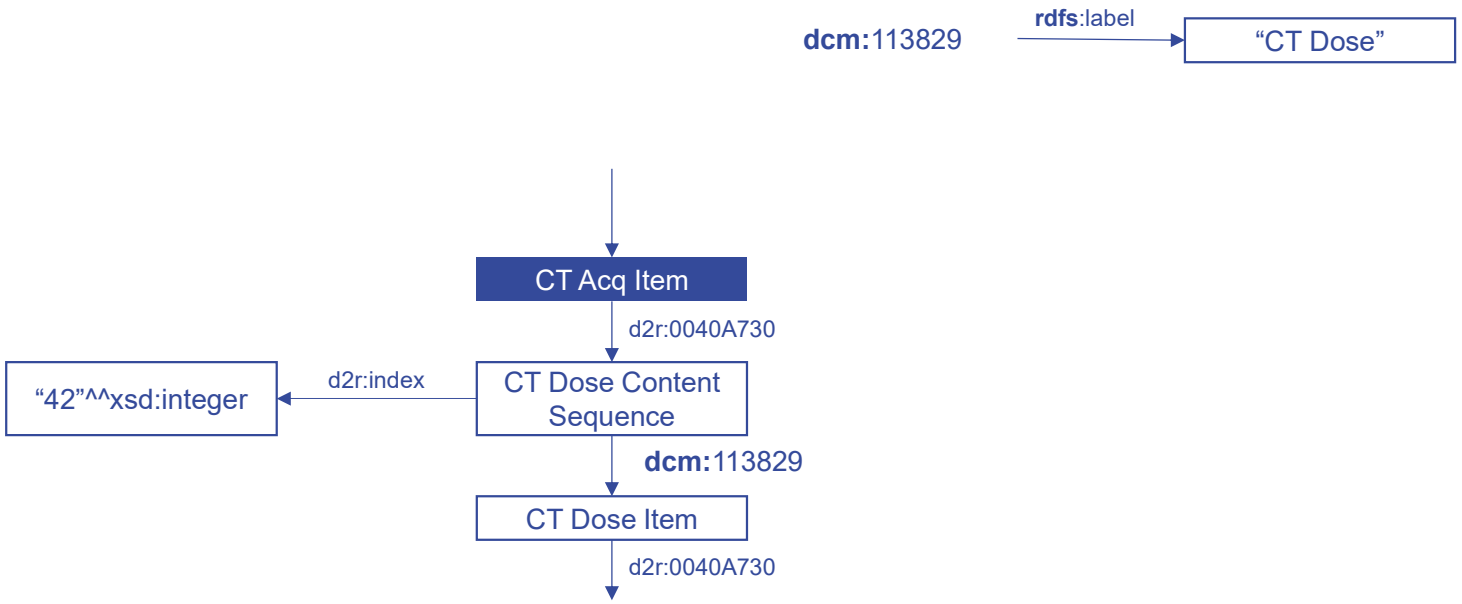

“DCM”
“113829”
CT Acq Item
d2r:0040A730
“42”^^xsd:integer
d2r:index
CT Dose Content Sequence
d2r:item
CT Dose Item
d2r:0040A043
CT Dose CNCS
d2r:0040A730
d2r:00080104
“CT Dose”

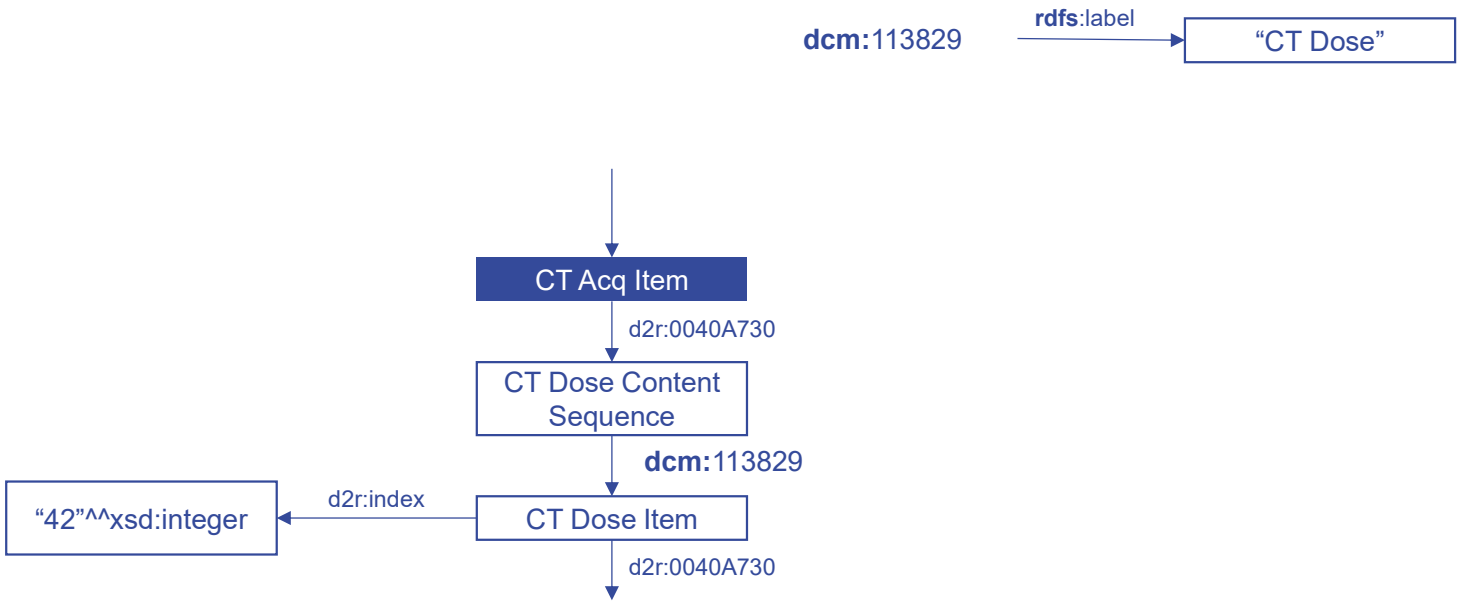
dcm:113829

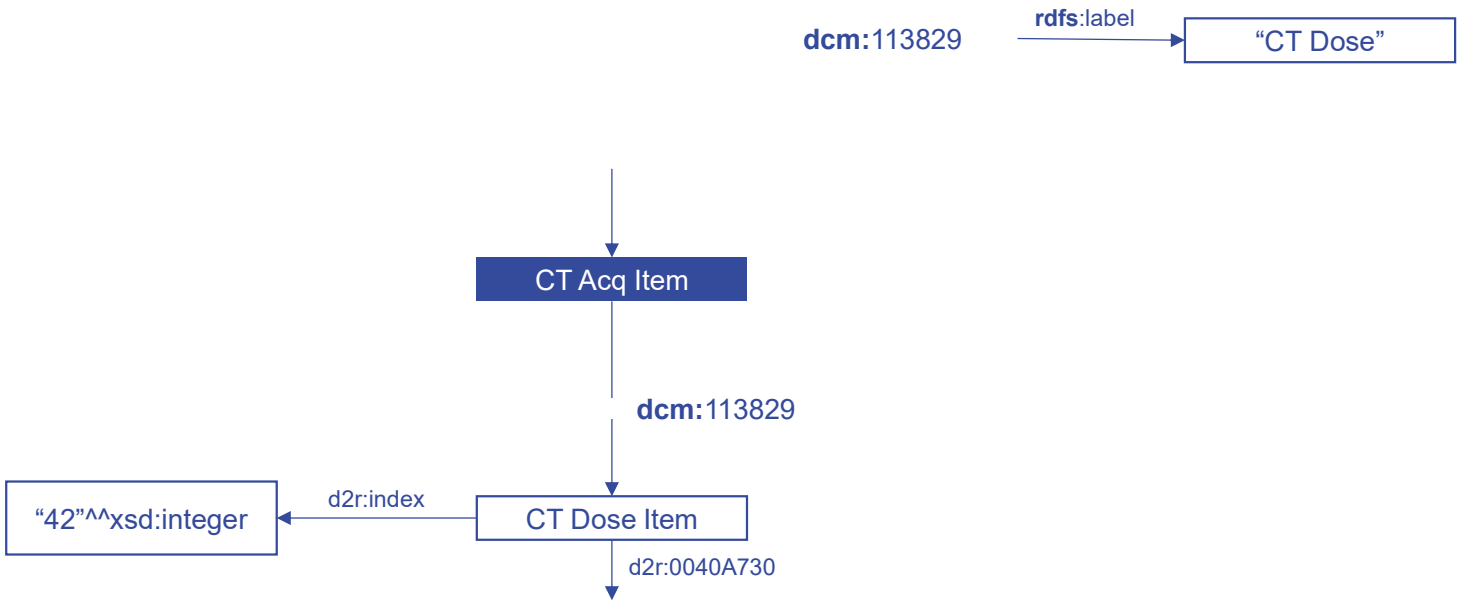

CT Acq Item
d2r:0040A730
CT Dose Content Sequence
d2r:index
"42"^^xsd:integer
d2r:item
CT Dose Item
d2r:0040A043
CT Dose CNCS
d2r:0040A730
d2r:00080104
"CT Dose"


dcm:113829
rdfs:label
“CT Dose”
CT Acq Item
d2r:0040A730
“42”^^xsd:integer
d2r:index
CT Dose Content Sequence
d2r:item
CT Dose Item
d2r:0040A043
CT Dose CNCS
d2r:0040A730

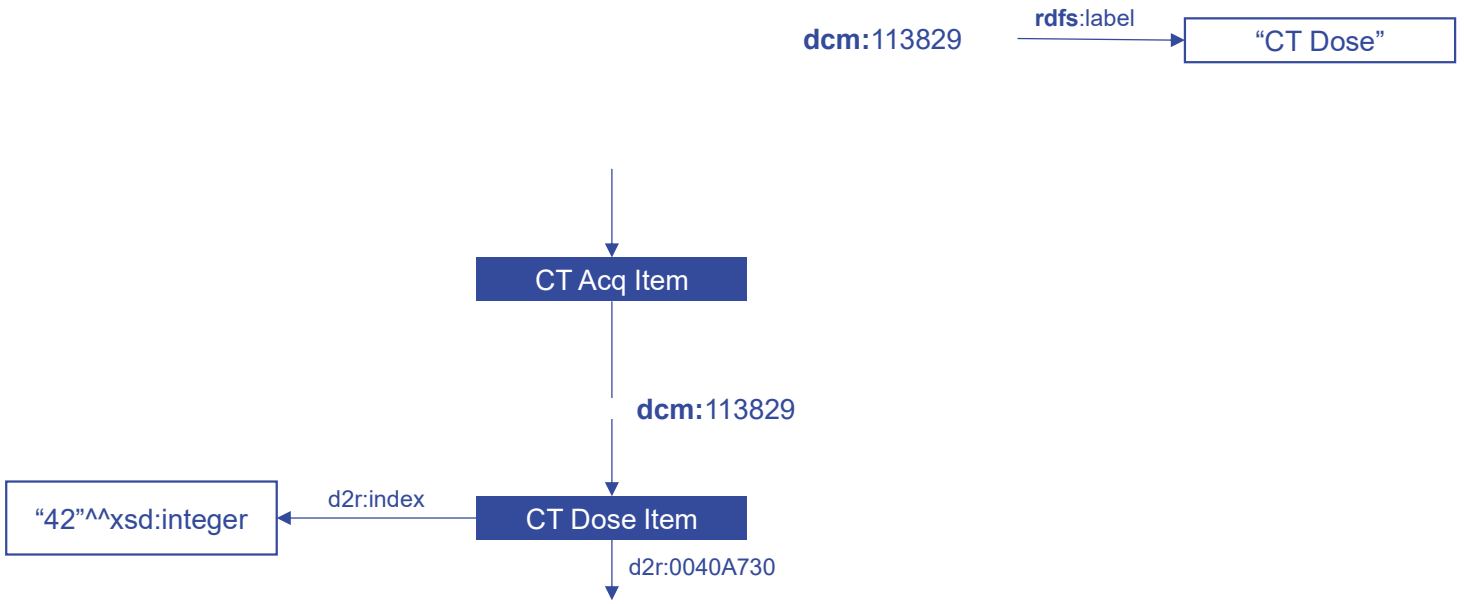

dcm:113829
rdfs:label
“CT Dose”
CT Acq Item
d2r:0040A730
“42”^^xsd:integer
d2r:index
CT Dose Content Sequence
d2r:item
CT Dose Item
d2r:0040A730


dcm:113829
rdfs:label
“CT Dose”
CT Acq Item
d2r:0040A730
“42”^^xsd:integer
d2r:index
CT Dose Content Sequence
dcm:113829
CT Dose Item
d2r:0040A730

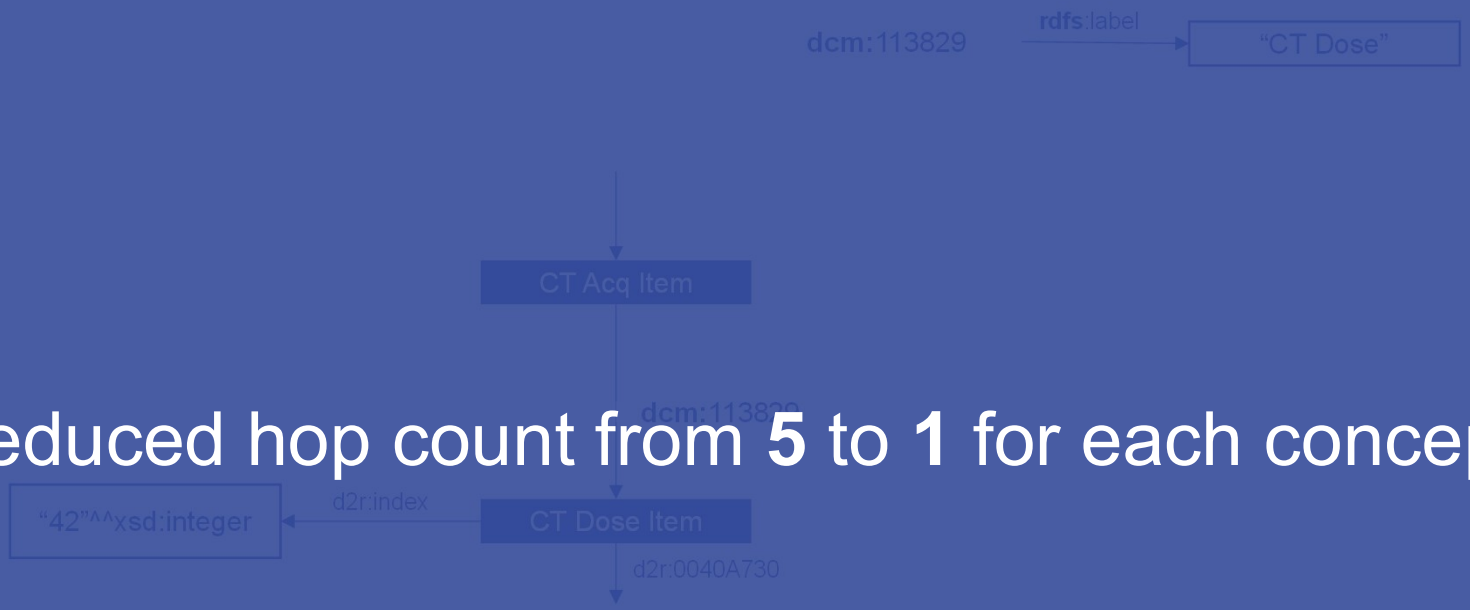

dcm:113829
rdfs:label
“CT Dose”
CT Acq Item
d2r:0040A730
CT Dose Content Sequence
dcm:113829
“42”^^xsd:integer
d2r:index
CT Dose Item
d2r:0040A730



Hop Counter: 0



Hop Counter: 1



Reduced hop count from **5** to **1** for each concept traversal

```
SELECT * WHERE {
    ?dose_report d2r:0040A730 ?ct_acq_cont_seq .
    ?ct_acq_cont_seq d2r:item ?ct_acq_item .
    ?ct_acq_item d2r:0040A043 ?ct_acq_cncs .
    ?ct_acq_cncs d2r:00080100 "113819" .
    ?ct_acq_cncs d2r:00080102 "DCM" .
    ?ct_acq_item d2r:0040A730 ?ct_dose_cont_seq .
    ?ct_dose_cont_seq d2r:item ?ct_dose_item .
    ?ct_dose_item d2r:0040A043 ?ct_dose_cncs .
    ?ct_dose_cncs d2r:00080100 "113829" .
    ?ct_dose_cncs d2r:00080102 "DCM" .
    ?ct_dose_item d2r:0040A730 ?ctdivol_cont_seq .
    ?ctdivol_cont_seq d2r:item ?ctdivol_item .
    ?ctdivol_item d2r:0040A043 ?ctdivol_cncs .
    ?ctdivol_cncs d2r:00080100 "113830" .
    ?ctdivol_cncs d2r:00080102 "DCM" .
    ?ctdivol_item d2r:0040A300 ?ctdivol_mv_seq .
    ?ctdivol_mv_seq d2r:004008EA ?ctdivol_mu_cs .
    ?ctdivol_mu_cs d2r:00080100 "mGy" .
    ?ctdivol_mu_cs d2r:00080102 "UCUM" .
    ?ctdivol_mv_seq d2r:0040A30A ?ctdivol .
}
```

```
SELECT * WHERE {
    ?dose_report dcm:113819 ?ct_acq .
    ?ct_acq dcm:113829 ?ct_dose .
    ?ct_dose dcm:113830 ?ctdivol .
    ?ctdivol qudt:unit ucum:mGy .
    ?ctdivol qudt:numericValue ?ctdivol_value .
}
```

Optimizations

Optimizations

- Leverage the fact that DICOM files do not depend on each other ➔ Data parallelism via Rayon
- Trade I/O boundedness for CPU cycles ➔ On-the-fly TTL output gzip compression
- Insufficient memory for CONSTRUCT queries on all 7 million DICOM files ➔ Split QLever instances
- Suboptimal memory allocation with default musl allocator ➔ Jemalloc

Evaluation

Technical performance and clinical utility

Evaluation
Workstation Specifications

- AMD Ryzen 9 5900X 12-core processor (24 threads)
- 128 GB RAM
- 3.6 TB NVMe SSD
- Ubuntu 24.04 LTS

Evaluation
Datasets

Public benchmark dataset

- Obtained from the National Cancer Institute Imaging Data Commons
- 19,870 Structured Reports (2.2 GB)
  - → derived from 40 GB DICOM data of NSCLC-Radiomics collection (mostly CT scans)
  - → 240 MB Structured Reports (1,987 files)
  - → replicated 10-fold
- Purpose: evaluation of *convert* stage
  - Conversion performance vs. *dcm2rdf*

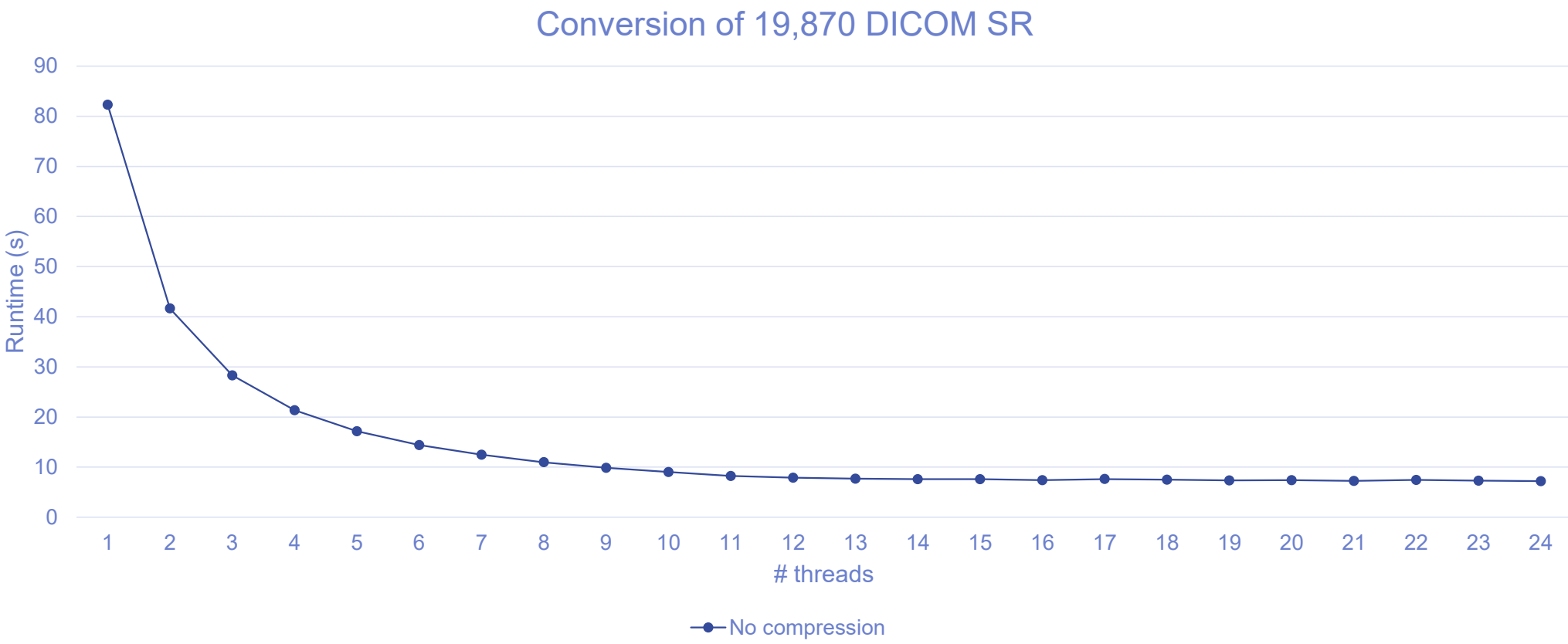
Institutional dataset

- Obtained from the University Medical Center Freiburg
- 7,139,423 Structured Reports (357 GB) of all kinds
- Purpose: evaluation of *construct* stage
  - Reduction of triple patterns and query runtime
  - Clinical utility analyses

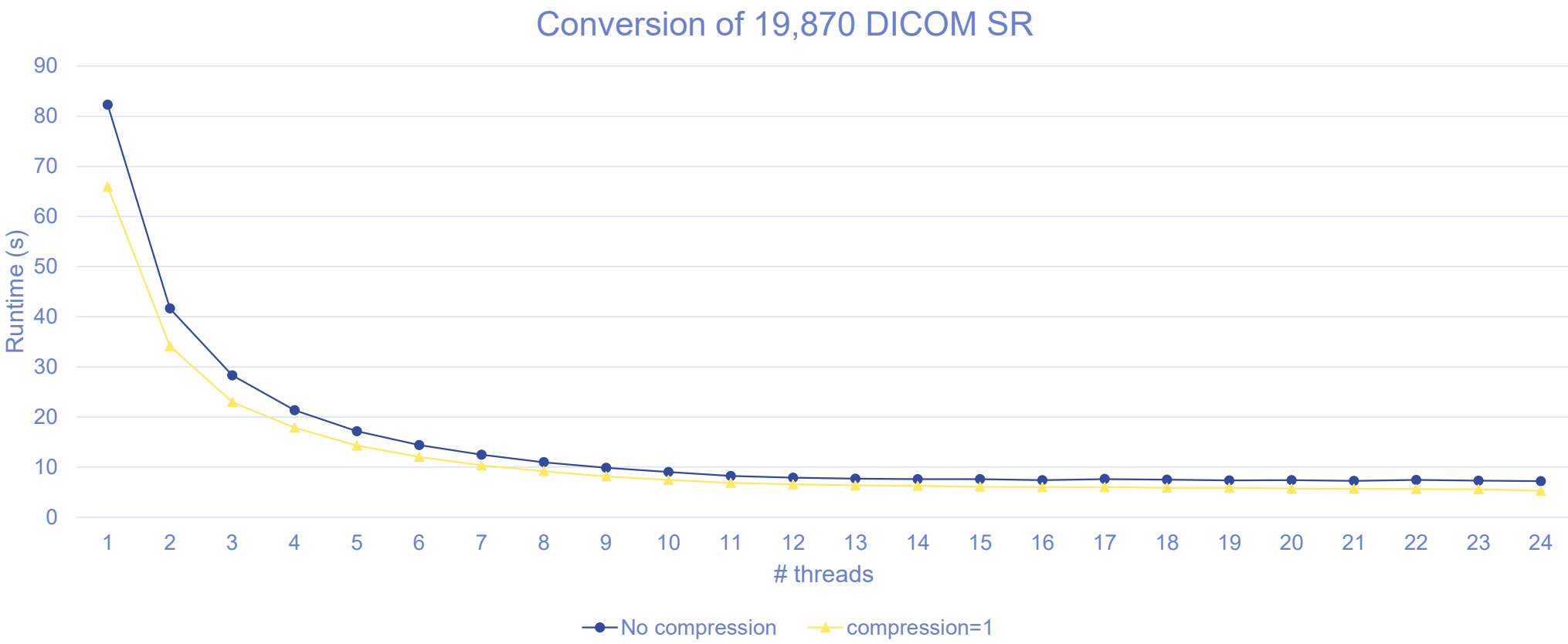
Evaluation
General Results

- Processed **7 million DICOM files** via ***convert*** stage, yielding **over 7 billion raw triples**
- Constructed semantic triples via ***construct*** stage, **reducing number of triples down to a billion**

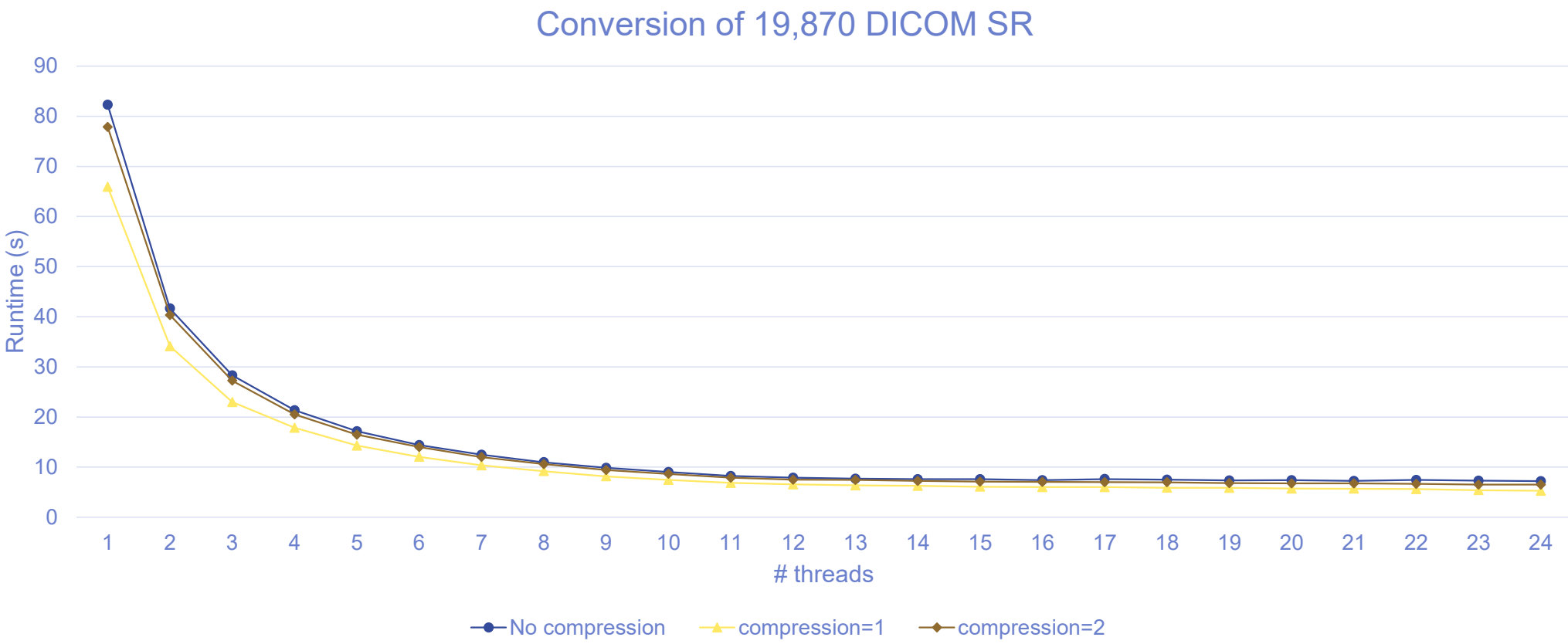
Evaluation
convert: Runtime



Evaluation
convert: Runtime

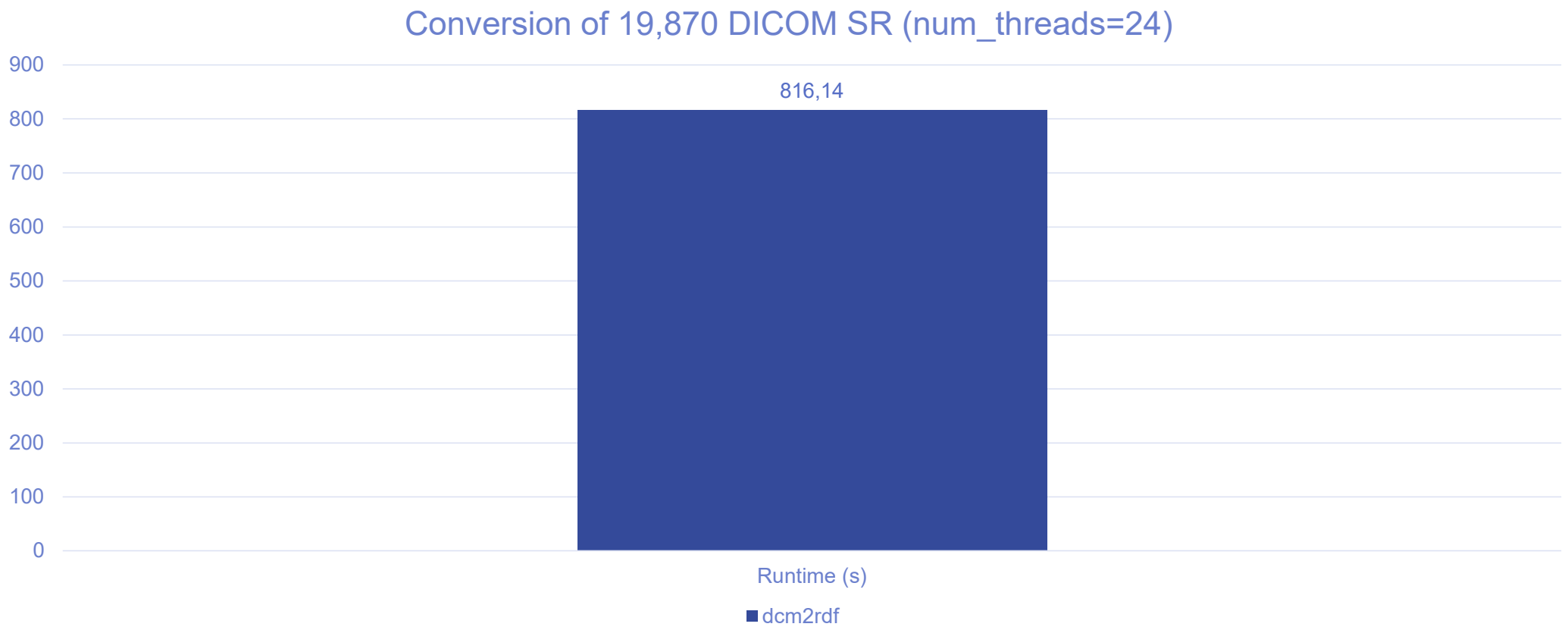


Evaluation
convert: Runtime



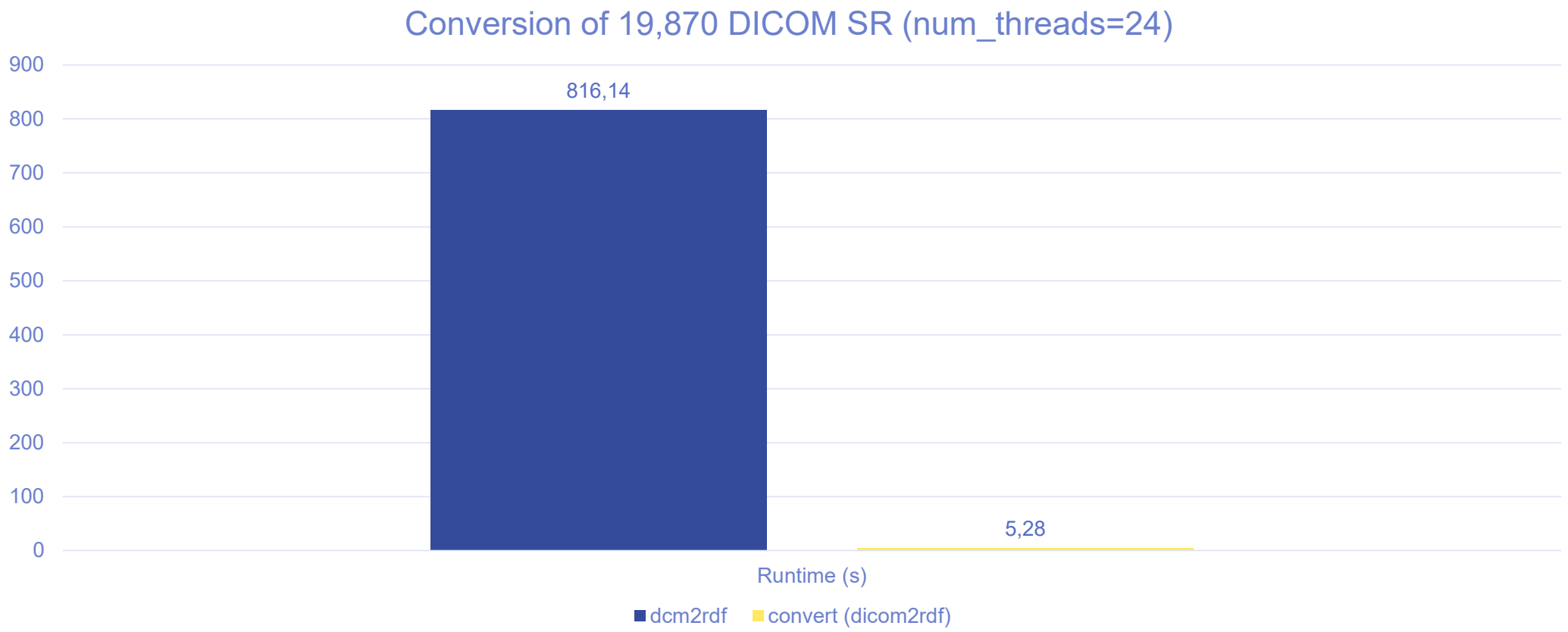
Evaluation

convert: Runtime vs. dcm2rdf



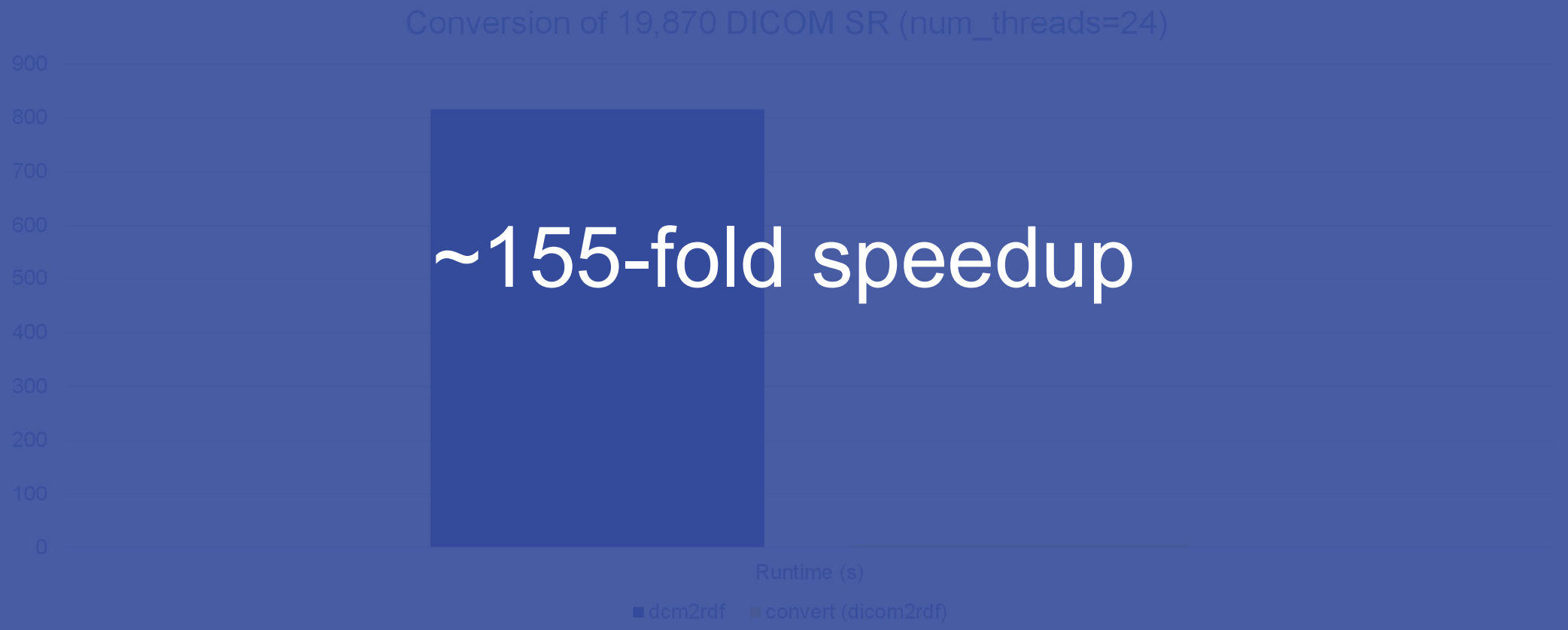
Evaluation

convert: Runtime vs. dcm2rdf



Evaluation

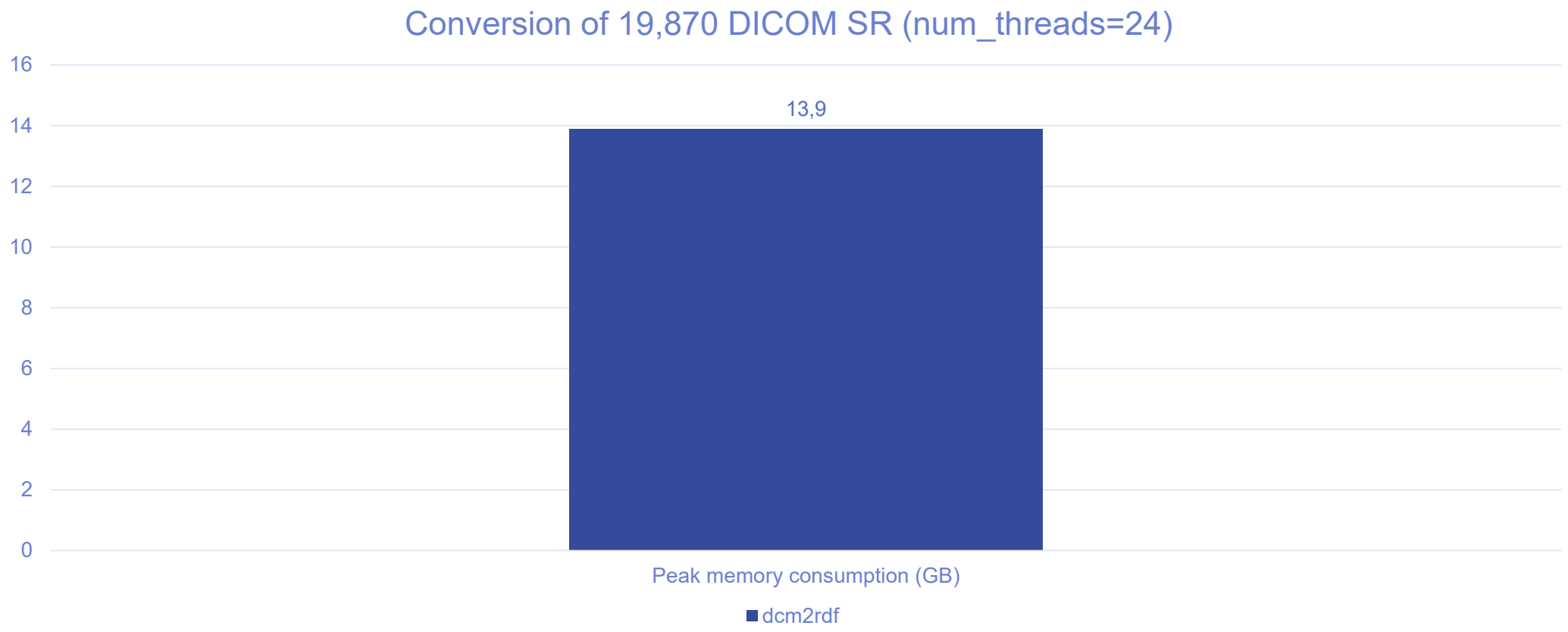
convert: Runtime vs. dcm2rdf



~155-fold speedup

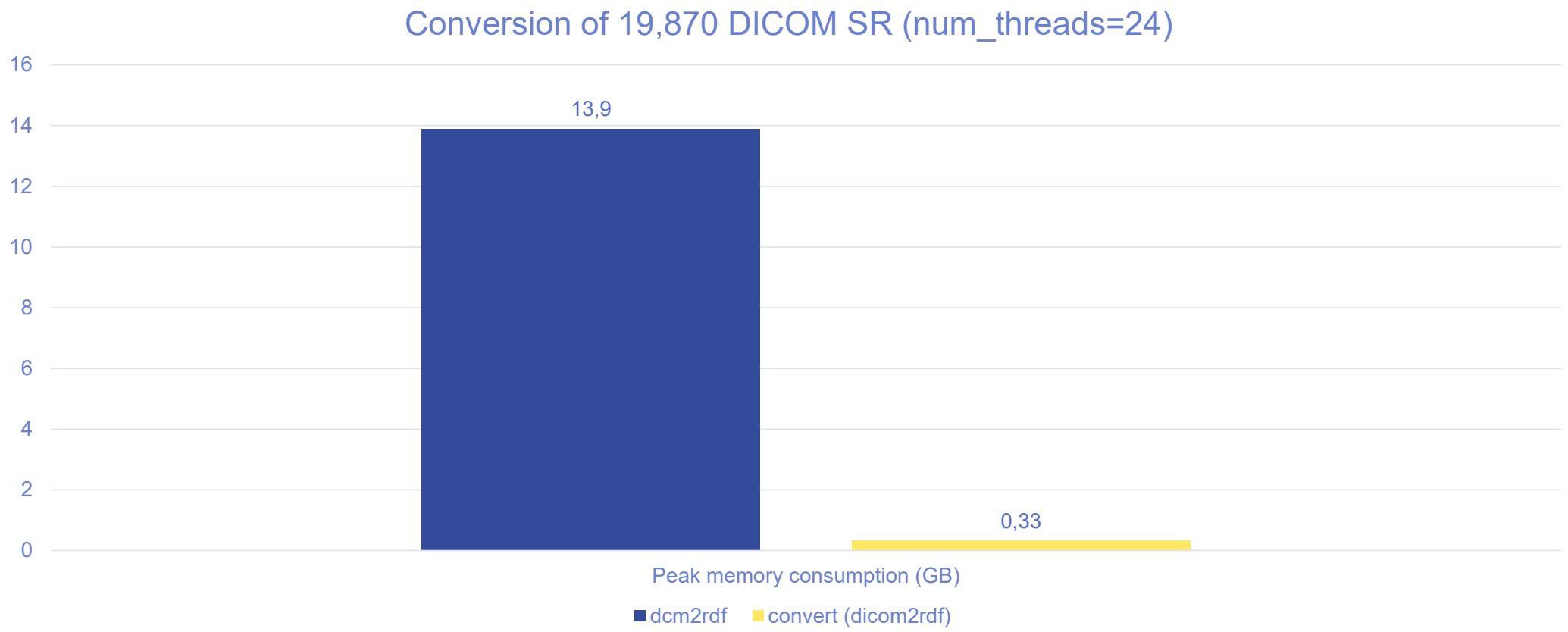
Evaluation

convert: Peak memory consumption vs. dcm2rdf



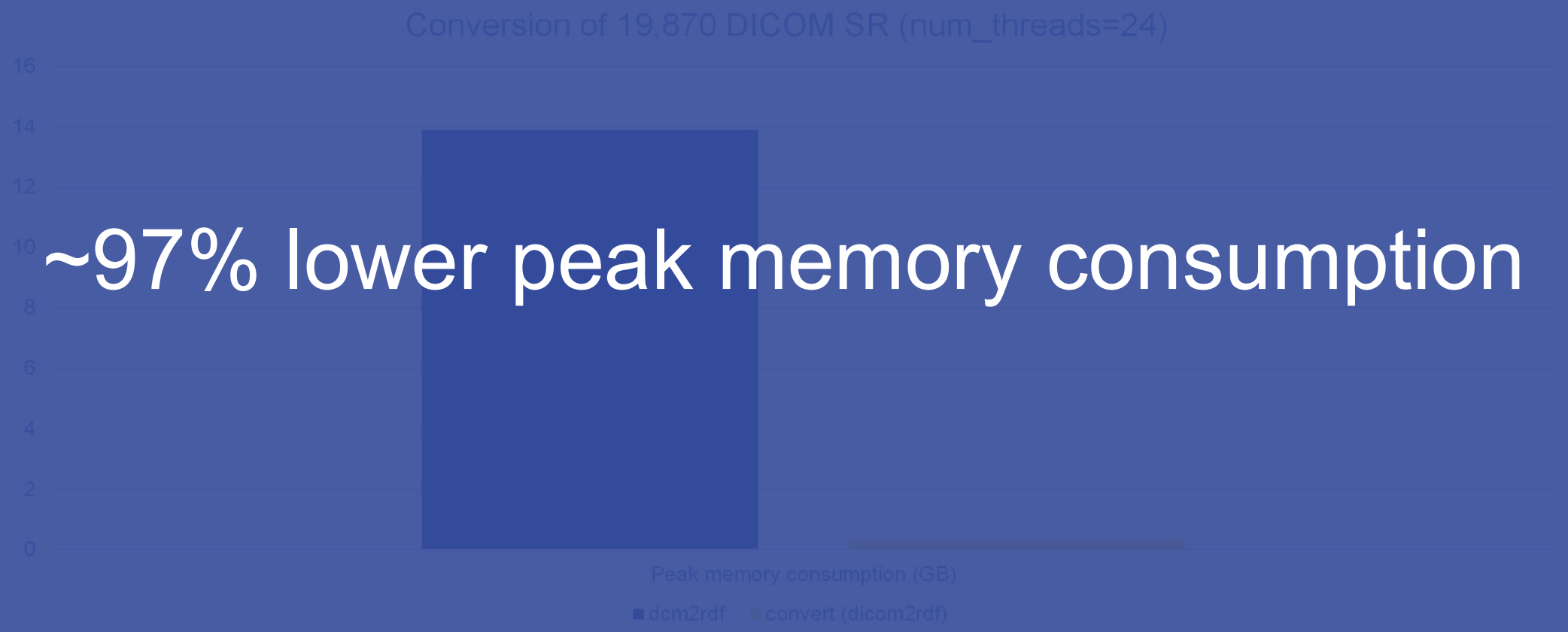
Evaluation

convert: Peak memory consumption vs. dcm2rdf



Evaluation

convert: Peak memory consumption vs. dcm2rdf



~97% lower peak memory consumption

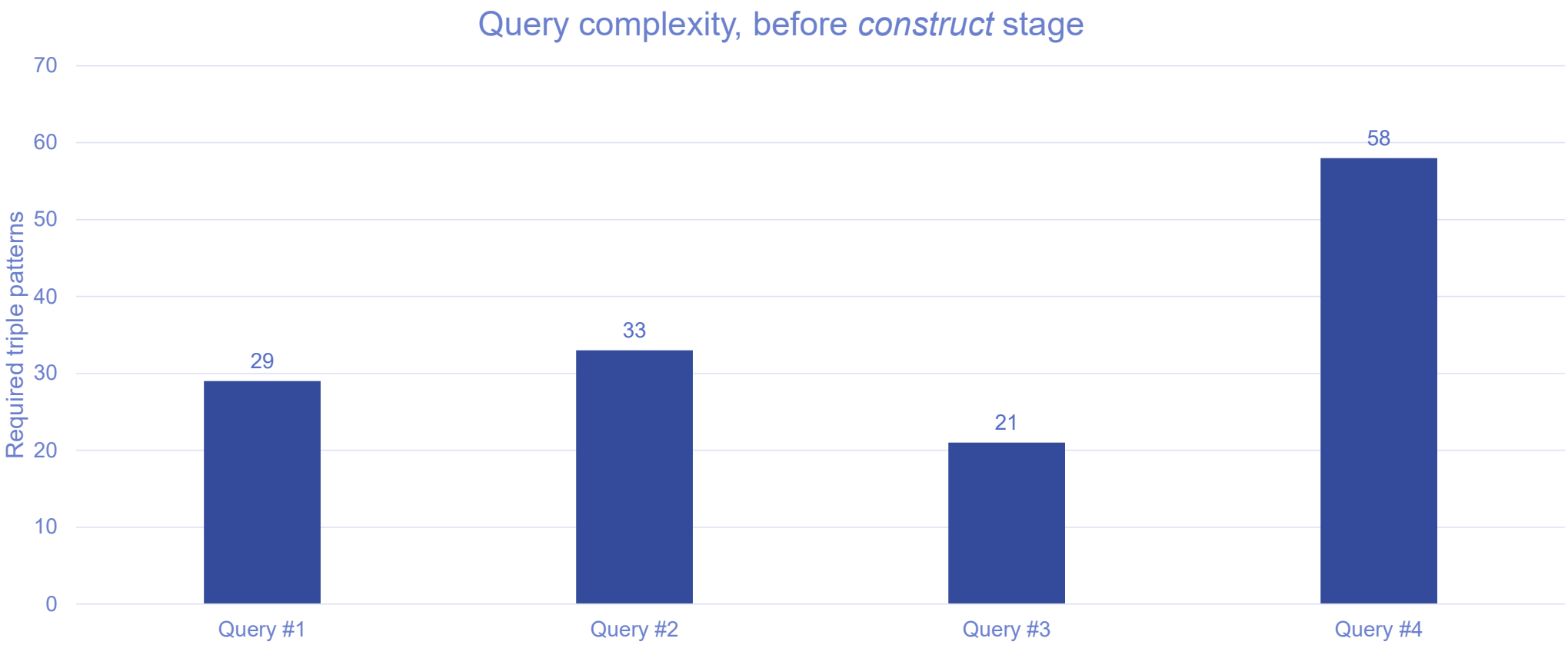
Evaluation

construct: Query complexity and runtime

- Query #1: *“CTDIvol by scanner model, acquisition protocol, patient height, patient weight”*
- Query #2: *“Impact of excluding non-diagnostic dose measurements for top 10 protocol names”*
- Query #3: *“Lung-RADS assessment distribution by gender”*
- Query #4: *“Mean hip-knee angle (HKA) deviation by BMI range and gender”*

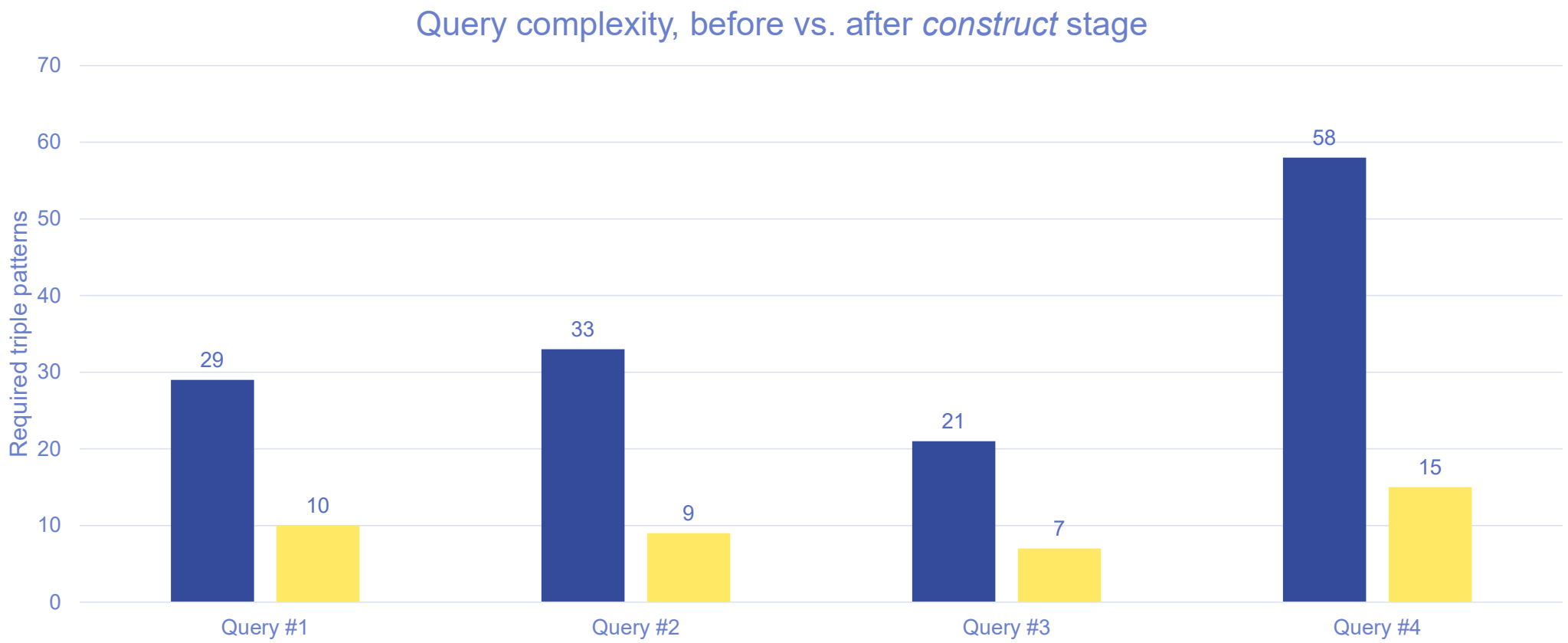
Evaluation

construct: Query complexity



Evaluation

construct: Query complexity



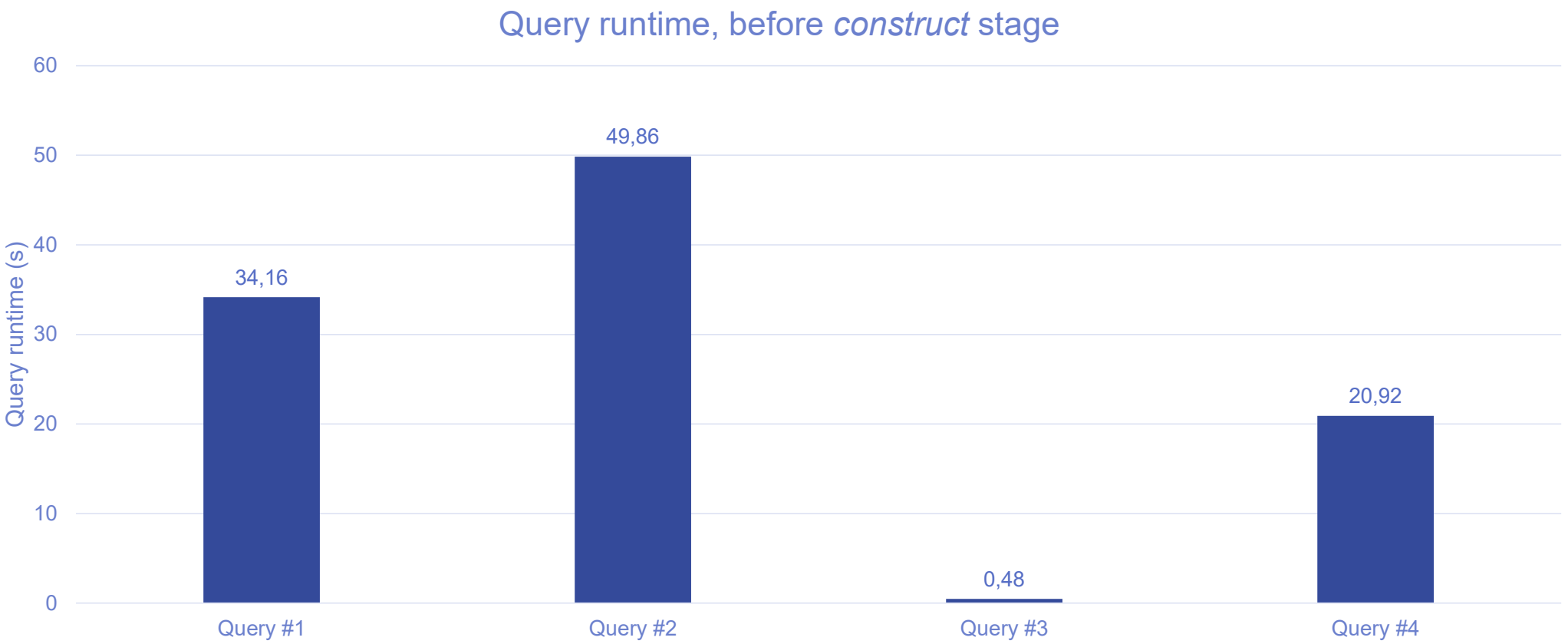
Evaluation

construct: Query complexity

Up to 74% reduction in query complexity

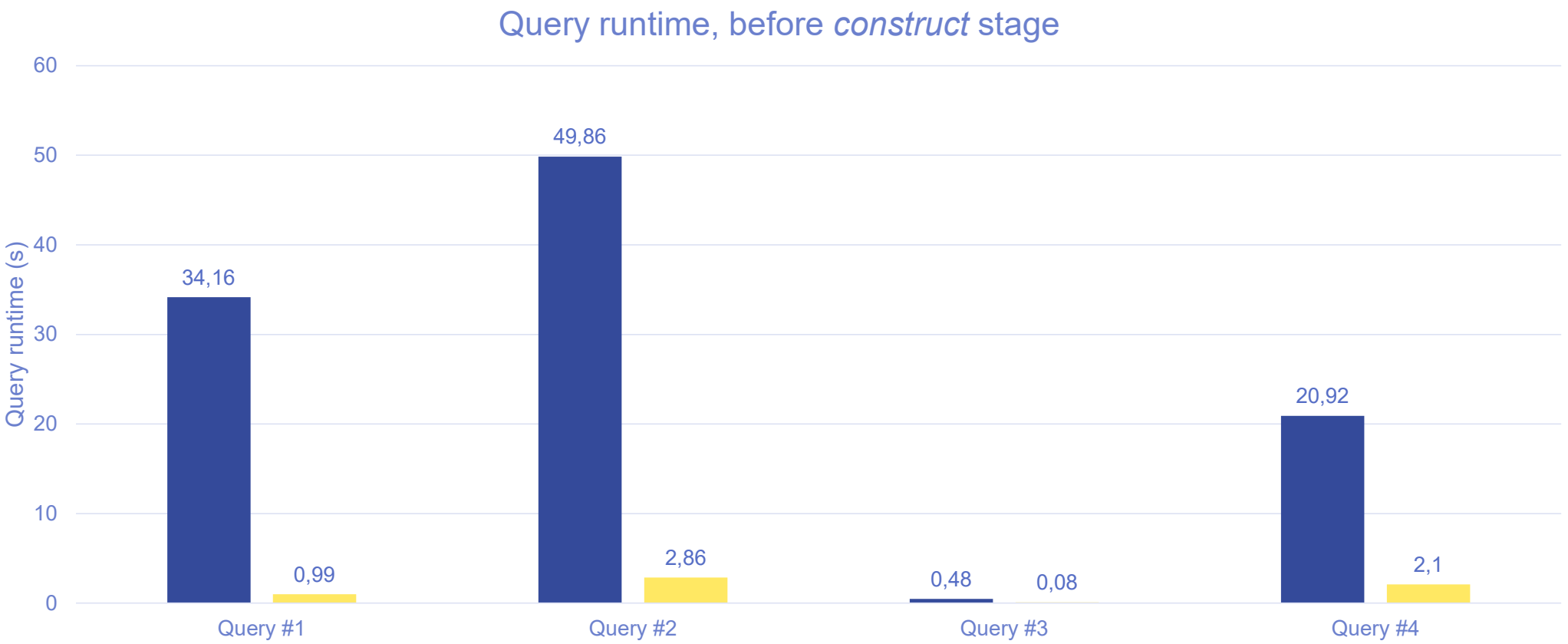
Evaluation

construct: Query runtime



Evaluation

construct: Query runtime



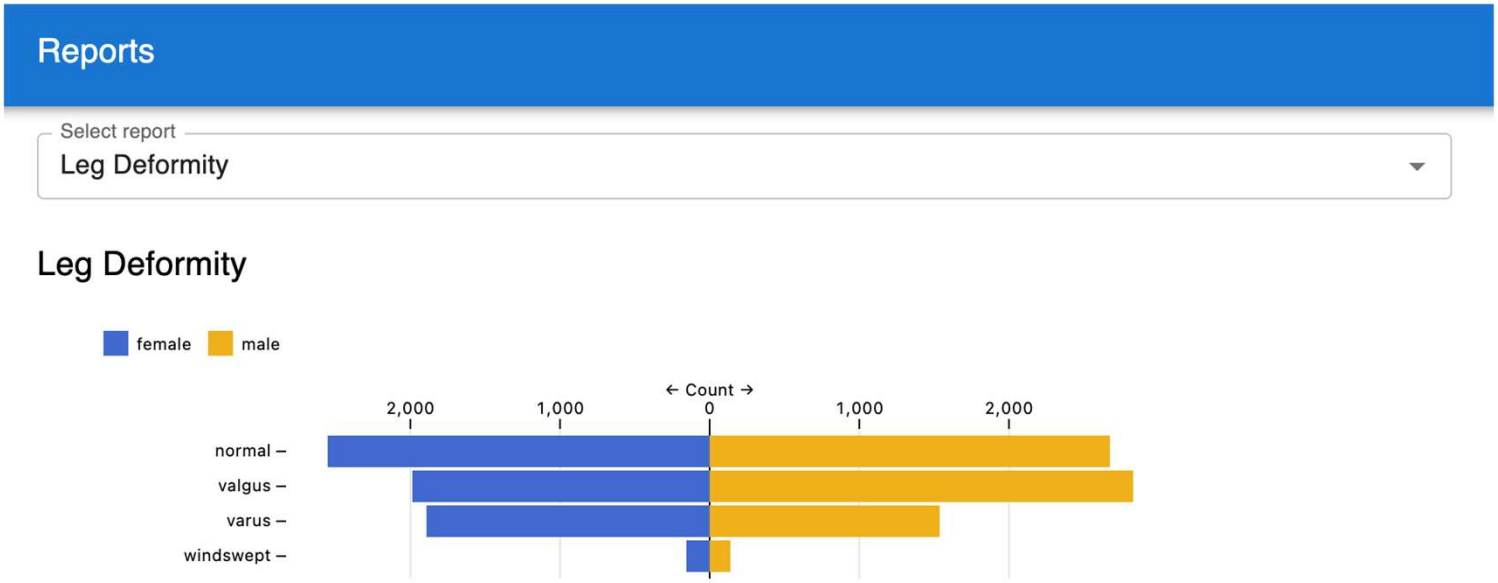
Evaluation

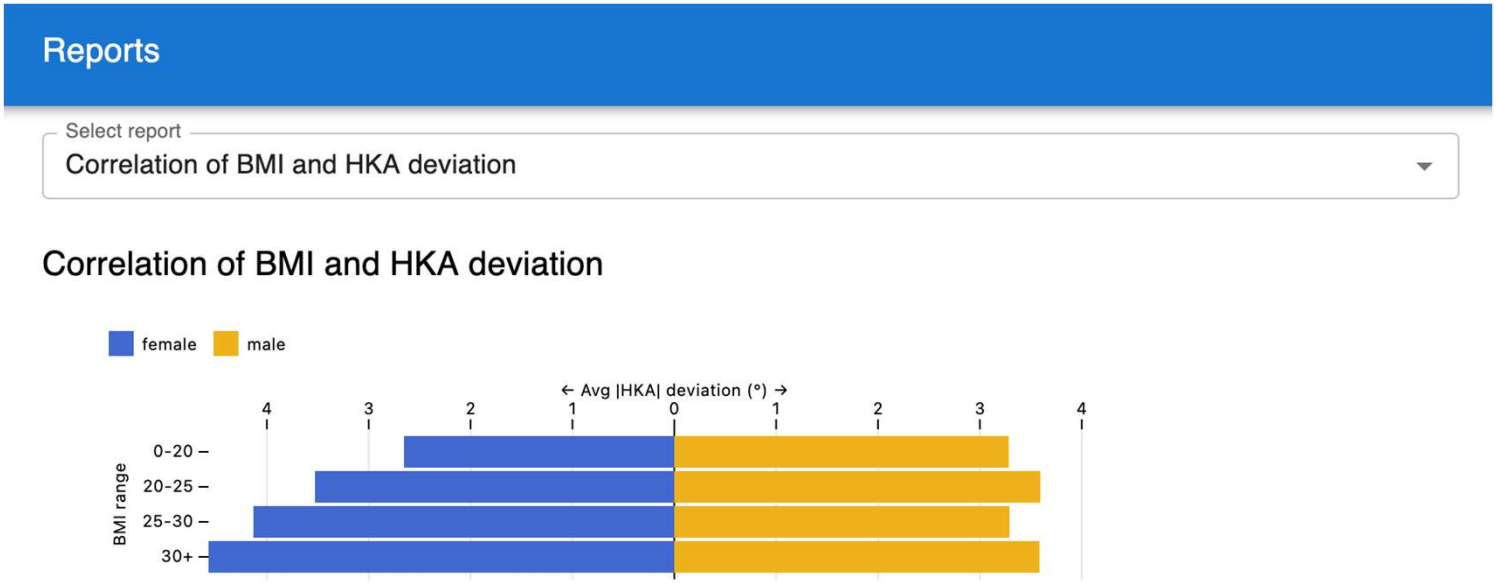
construct: Query runtime

Up to 97% reduction in query runtime

Evaluation
Clinical Utility

Evaluation
Clinical Utility




Reports
Select report
Correlation of BMI and HKA deviation
Correlation of BMI and HKA deviation
female
male
← Avg |HKA| deviation (°) →
4
3
2
1
0
1
2
3
4
BMI range
0-20
20-25
25-30
30+

Conclusion

Conclusion

- Representing **DICOM as knowledge graphs** is **feasible at institutional scale**
- ***convert*** stage is approx. **155 faster** with **97% less peak memory consumption** than comparable solution
- ***construct*** stage **decreases query complexity** by up to **97%** and **runtime** by up to **74%**

Future Work

Future Work

- construct: Iterative construction of semantic triples
- construct: More extractor queries for remaining VRs
- Explore data beyond radiation dose reports (mammography, ultrasound, …)
- Transition from retrospective to prospective system
- Explore LLMs as natural language interface for translating clinical question to SPARQL queries

Thank you for listening! Any Questions?

References

Bundesamt für Strahlenschutz, “What does the ALARA principle mean?.” Accessed: Jan. 17, 2026. [Online]. Available: https://www.bfs.de/SharedDocs/FAQs/BfS/EN/ion/ion/alara.html

World Health Organization, “Ionizing radiation and health effects.” Accessed: Jan. 17, 2026. [Online]. Available: https://www.who.int/news room/fact sheets/detail/ionizing radiation and health effects

R. Smith Bindman et al., “Projected Lifetime Cancer Risks From Current Computed Tomography Imaging,” JAMA Internal Medicine, vol. 185, no. 6, pp. 710–719, 2025, doi: 10.1001/jamainternmed.2025.0505.

“Council Directive 2013/59/Euratom of 5 December 2013 laying down basic safety standards for protection against the dangers arising from exposure to ionising radiation.” [Online]. Available: https://eur lex.europa.eu/eli/dir/2013/59/oj

References

The DICOM Standard authors, “The DICOM Standard.” Accessed: Jan. 18, 2026. [Online]. Available: https://dicomstandard.org/

I. K. Sebelego, S. Acho, B. van der Merwe, and W. I. D. Rae, “FACTORS INFLUENCING SIZE SPECIFIC DOSE ESTIMATES OF SELECTED COMPUTED TOMOGRAPHY PROTOCOLS AT TWO CLINICAL PRACTICES IN SOUTH AFRICA,” Radiation Protection Dosimetry, vol. 199, no. 7, pp. 588–602, 2023, doi: 10.1093/rpd/ncad059.

[7] E. Z. Dalah et al., “Challenges Associated with Effective Implementation of CT Dose Check Standards and Radiation Monitoring Index in Computed Tomography: Healthcare Sector Experience,” Healthcare, vol. 10, no. 10, 2022, doi: 10.3390/healthcare10101970.

[8] F. Ria et al., “Expanding the Concept of Diagnostic Reference Levels to Noise and Dose Reference Levels in CT,” AJR American Journal of Roentgenology, vol. 213, no. 4, pp. 889–894, 2019, doi: 10.2214/AJR.18.21030.