

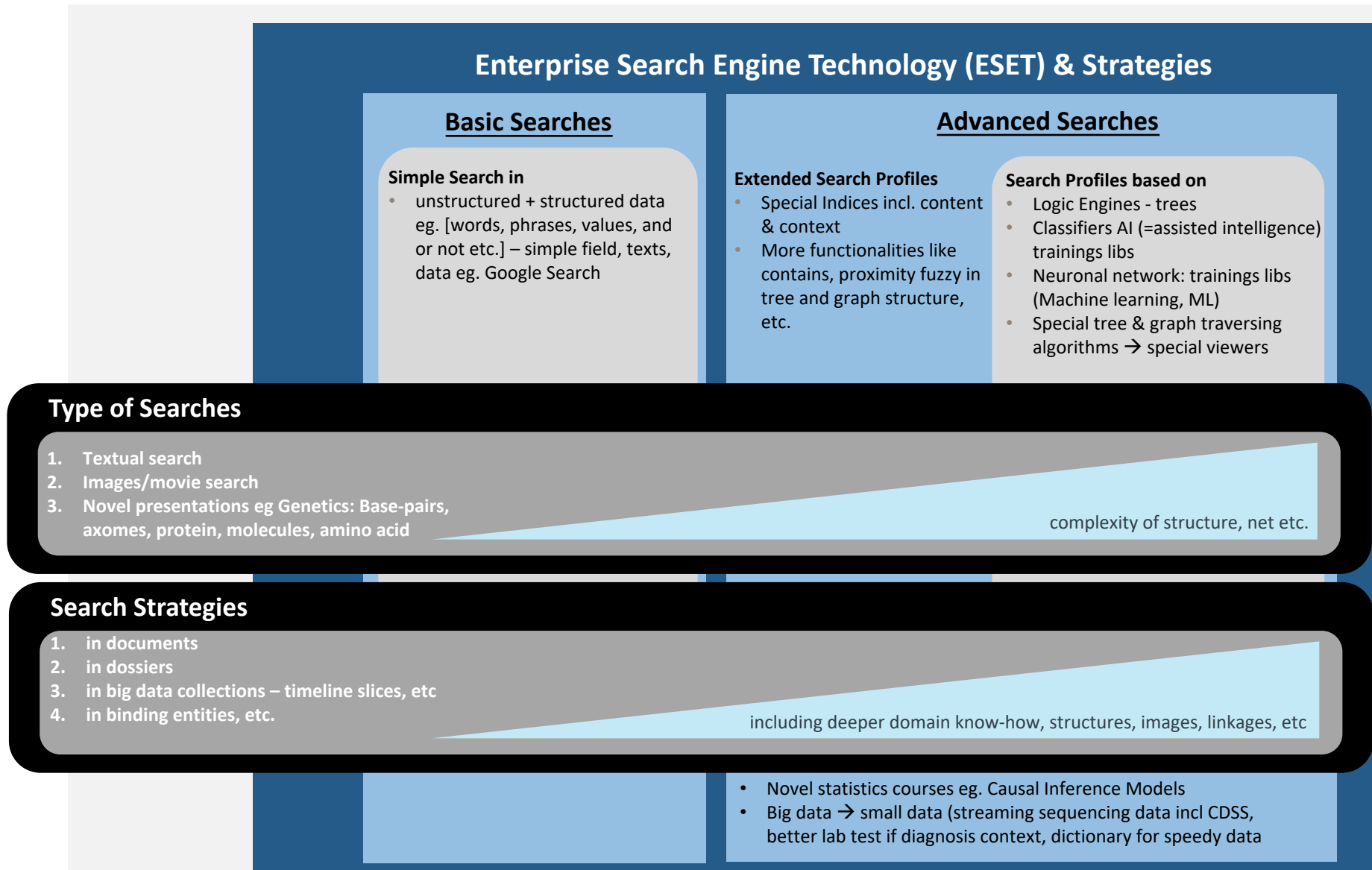
Introduction
Search
Profiles



ITERATA EXPLAINS

A CRITICAL APPROACH TOWARDS THE SUBJECT

Advanced Searches from BIG Data to SMALL Data



Cohort Identification

prior the search

- **Concrete** research question (PICO)
 - H0
 - H1
- Definition exclusion/inclusion criteria (age, time range, gender, diagnosis, side effects, medication, therapies...)

→ Set of **searchable** attributes

- Creating search profile

search request

- Running created search profile (e.g. 1'000)
- Adjusting search profile (add/remove attributed)
- Rerunning search profile (e.g. 800)

→ **feasible** # of patients in cohort

Cohort Validation

Validation

- **Manual** validation of patients based on inclusion/exclusion criteria
- Reasoning of inclusion/exclusion of patients (yes/no)

→ Final cohort extracted (e.g. 300)

Advanced validation

- Automated validation based on predefined classifier sets (machine learning methods)

Cohort Structuring

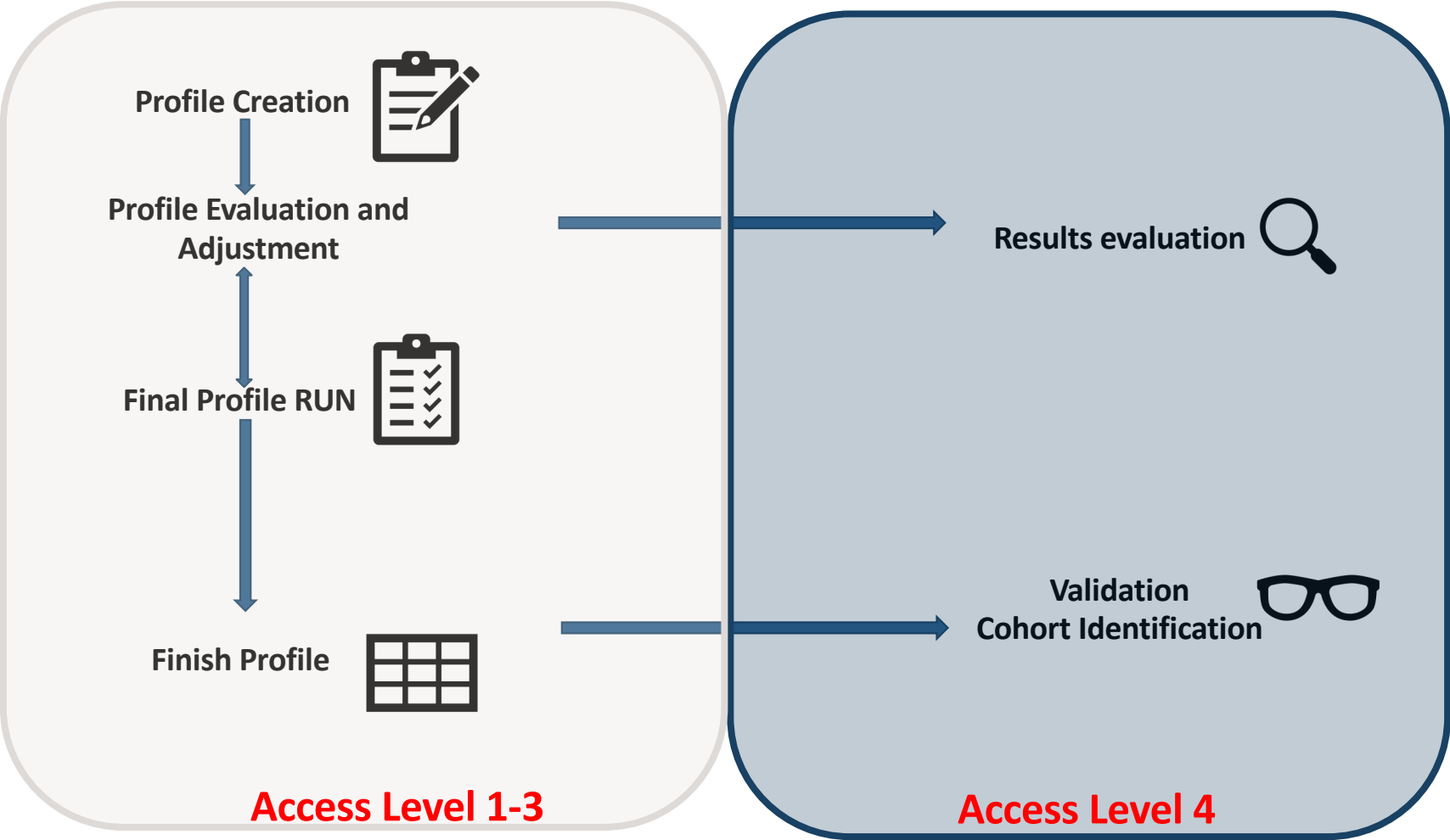
Structuring

- Search results **structured** in table form
 - Rows: patients
 - Column: search attributes

Statistics

- How to evaluate the search results
- #### → Statistic curriculum





Cohort Identification

Search Profiles



Prior the search - PICO

Development of question of interest that is as precisely as possible

- Based on this, extraction of keywords for search profile
- Determine hierarchical order of keywords

P: patient/population/ problem (age, gender, diagnosis, side effects)

I: intervention (treatment, medication, therapies, surgery, predictors)

C: comparison (alternative treatments, control group – for interventional studies)

O: outcome (treatment aim, outcome measures)

Additional:

T: Time (defined time frame)

S: study design (RCP, observational)

Search Profile Creation

	column A	column B	column C	column D
row 1	Vorhofflimmern			
row 2		Herzinsuffizienz		
row 3			Schlaganfall, TIA, Zerebrale transitorische Ischämie	
row 4				Hämatemesis
row 5		Herzinsuffizienz		
row 6			Hypertonie, Bluthochdruck~	Hämatemesis
row 7		Herzinsuffizienz		
row 8			Diabetes mellitus ~	
row 9			patientBirthdate=1946<>1955	
row 10		Schlaganfall		
row 11			Hämatemesis ~	
row 12	Biopsie			
row 13		Transplantation		
row 14			Nieren	
row 15		!Diabetes melitus		
row 16		Lebertransplantation		
row 17		*biopsie		
row 18		Lungen*		

- **AND** eg. («Vorhofflimmern AND Herzinsuffizienz») [same page]
- **AND & OR** eg. («Vorhofflimmern AND Herzinsuffizienz AND (Schlaganfall OR TIA OR Zerebrale transitorische Ischämie) AND Häatemesis») [same page]
- **Proximity search** eg. «Diabetes melitus~»
- **OR** eg. («Vorhofflimmern AND Herzinsuffizienz AND Diabetes mellitus») OR («Vorhofflimmern AND Herzinsuffizienz OR patientBirthdate=1946<>1955»)
- **Fuzzy + Factor** eg. «Hämatemesis~»
- **NOT** eg. «Diabetes melitus»
- **End-with** eg. *biopsie = Nierenbiopsie, Leberbiopsie, etc.
- **Start-with** eg. Lungen* = Lungenleiden, Lungenkrankheit, etc.
- **Substring** combination eg. *lungen* =Herzlungenmaschine

Semantic for Excel Tree check in:

- **AND:** An AND Linkage is represented in a row (horizontal merges). exact phrases between ""Each term is in the same row in the 1..n columes
- **OR:** An OR Linkage is represented in combination of more than one row (vertical merges)
- **OR** in cell: [item, item, ... , item] = [item OR item OR OR item]
- **NOT:** is represented as ![item or phrases]
- **End-with:** *[item]; **Start-with:** [item]*; **Substring:** *[item, phrases]*

All searches e.g. cell items or phrases « » will be executed as an exact term search, but with ~ at the end a **Proximity search** will be performed: [item item]~

- **Fuzzy:** Just per item could be weighted with a fuzzy factor, e.g. spelling mistakes "Aschpirin", "Vorhofflimmern"

Way of Searches - Result Set

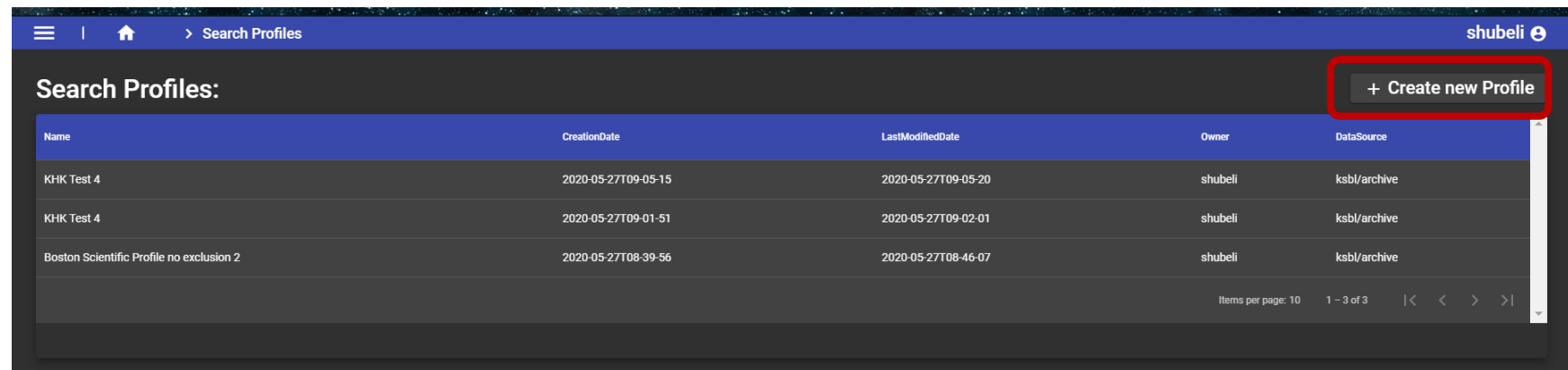
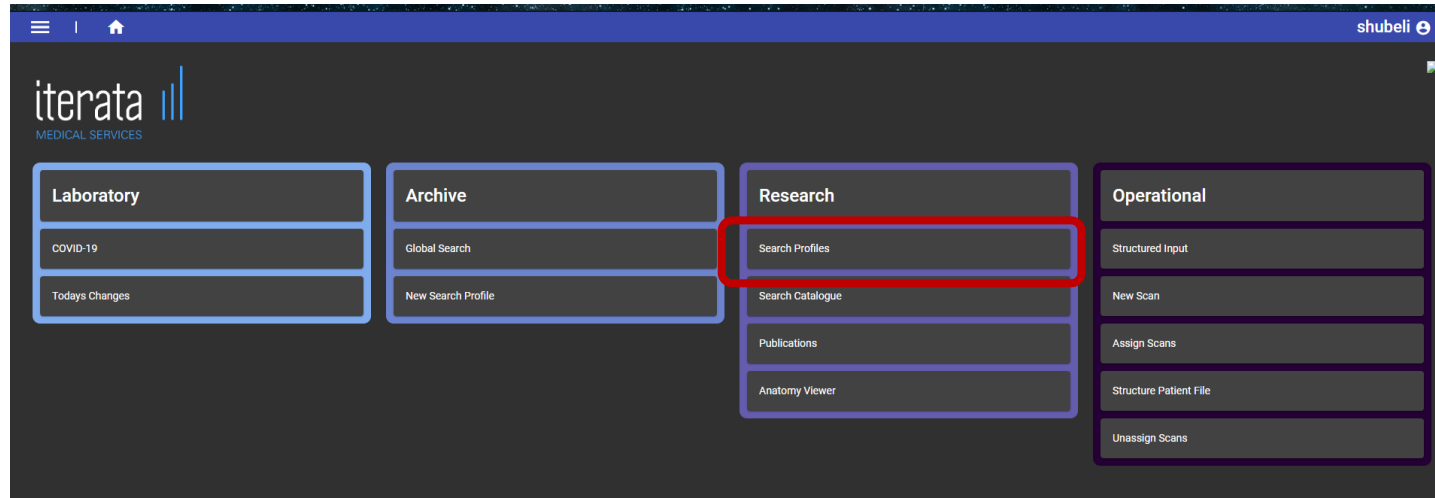
Way of Searches

1. In one report (all criteria's) - possible
2. In one patient dossier (throughout all concatenated reports) – possible for dedicated searches
3. All in Dictionary – referenced to all pat and docs – coming soon (graph database)

Result Set

- # Documents and #Patients per Combination (row)
- Over selected Row's → Freeze of potential cohort

Check in Profile



Check in Profile – Select Profile Definition

The screenshot displays the 'NewProfile' configuration interface. On the left, a sidebar shows profile metadata: Owner: shubeli, Created at: 2020-05-27T09-11-41, Last Modified: 2020-05-27T09-11-41, Last Evaluation: --, Data Source: ksbl/archive, and Status: created. A 'Finish Profile' button is at the bottom of this sidebar. The main area is titled 'No Definition available' and features an 'Upload Definition' button, which is highlighted with a red box. An 'Open' file explorer window is overlaid on the main area, showing the path 'This PC > Desktop > TestProfile'. The file list includes:

Name	Date modified	Type	Size
BostonScientific_CHA2DS	5/27/2020 10:45 AM	Microsoft Excel-Ar...	7
Iterata_Profil_Definition_BS-Profil-Reiter2...	5/27/2020 10:45 AM	Microsoft Excel-Ar...	15
KHK_SearchProfile_4	5/27/2020 10:56 AM	Microsoft Excel-Ar...	5
KHK_SearchProfile_3	5/27/2020 11:01 AM	Microsoft Excel-Ar...	10

The file 'KHK_SearchProfile_3' is selected and highlighted with a red box. The 'File name' field at the bottom of the explorer is empty, and the file type is set to 'All Files'.

RUN Profile → Result per each combination of items or phrases

The screenshot displays a web application interface for configuring and running search profiles. The profile is named "KHK Test 4" and is owned by "shubeli". The profile was created on 2020-05-27T09-05-15 and last modified on 2020-05-27T09-05-20. The data source is "ksbl/archive" and the status is "done". A "Finish Profile" button is visible.

The results table shows the following data:

Title	Count	GroupedCount
Koronare Herzkrankheit,KHK,1-Gefässerkrankung,Eingefässerkrankung,Einasterkrankung,1-Gefäss-KHK,1-Ast-Erkrankung,1-Ast-KHK,ko	274435	91387
dateTimeCreated=2015->2019 AND Koronare Herzkrankheit,KHK,1-Gefässerkrankung,Eingefässerkrankung,Einasterkrankung,1-Gefäss	84280	33274
documentTitle=*bericht* AND dateTimeCreated=2015->2019 AND Koronare Herzkrankheit,KHK,1-Gefässerkrankung,Eingefässerkrank	27343	8067
Koronare Herzkrankheit	48368	13630
KHK	180017	85099
1-Gefässerkrankung	8978	1855
Eingefässerkrankung	552	240
Einasterkrankung	461	215
1-Gefäss-KHK	386	195
1-Ast-Erkrankung	4854	1262

The interface includes a navigation bar with "Search Profiles" and "Configure Profile" options, and a user profile "shubeli". The results table has a "Run" button highlighted with a red arrow and a red box around the highlighted row.

Cohort Validation

Search Profiles



Validate Result of Profile on Level 4

KHK Test 4

Owner: shubeli
Created at: 2020-05-27T09-05-15
Last Modified: 2020-05-27T09-05-20
Last Evaluation: 2020-05-27T09-05-20

Data Source: ksbj/archive

Status: done

Group	GroupedTotal
13636cab854e0e1d61ba45b1a0685ea	67
a364486b41fae5cb62850a588b5d400	49
3830e2d13350ab9db7113047dd5ca4e	42
84ce0b6f0d50a4b1a1d1a9b160abe26	37
64603e404959f90bab4615f8474bf9ee	34
1802369c6ffa61babb5db66cb0c0fcb8	33
629c79bd2b845aa4466ff16ddcbaf4f2	32
3fb75073baed9663a133511f38f75bec	32
9e281be0d3de1b603c5c9c0e3989a3d	31
b1048a75f75ebd7e5971d0be63e033	30

1 - 10 of 8067 | Items per page: 10

Finish Profile

Gender: F
Birthdate: 1941-10-03
Total Documents: 67
Last Document:

Show Search Documents

Search

Document	Date
Med_Austrittsbericht - ...	2018-12-12T15:17:04
Med_Kurzaustrittsberi...	2018-12-03T14:20:41
Med_Kurzaustrittsberi...	2018-12-03T14:20:41
NFS_Austrittsbericht a...	2018-11-12T17:51:39
NFS_Austrittsbericht a...	2018-11-12T17:51:39
NFS_Austrittsbericht a...	2018-10-06T16:06:06
NFS_Austrittsbericht a...	2018-10-06T16:06:06
Chir_Sprechstundenbe...	2018-07-03T14:06:42

Diagnosen

- Ambulant erworbene Pneumonie rechts basal ml/bei
- qSOFA-Kriterien: 0/3
- Röntgen Thorax: Verdacht auf flues/beginnendes Infiltrat rechts basal, am ehesten dem Mittellappen entsprechend.
- Sputum 11/2018: Positiv auf *Haemophilus influenzae* und *E. coli*
- Blutkulturen (4/4) 11/2018: Steril
- Legionellen-/Pneumokokken-Antigen 11/2018: Negativ
- Chronisch obstruktive Lungenerkrankung (COPD) GOLD II, Risikogruppe B und Verdacht auf allergisches Asthma (Verdacht auf Asthma-COPD-Overlap-Syndrom)
- Status nach Tabakkonsum bis 2006, kumulativ 50 py
- mittelschwere, zurzeit nicht-reversible, obstruktive Ventilationsstörung, schwergradig eingeschränkter Gasaustausch, Verbesserung im Vergleich zum 12/2017
- anamnestisch Allergie auf Tierhaare, Pollen, Gesamt IgE erhöht auf 238 kU/l, spez. IgE diskret auf Katze, negativ auf Hund, Staub, Pilze, Pollen, grenzwertige Bluteosinophilie
- Status nach Infekt mit Influenza B und Hospitalisation 27.12.17 - 01.01.18
- Koronare 2-Gefässerkrankung**
- cvRF: Arterielle Hypertonie, Dyslipidämie, Status nach 50 py, sistiert 2007
- antero-apikale Ischämie (MPS 01.10.12), LVEF 69 %
- Koronarangiographie vom 11/2012: RCX-RM 75 % Stenose->< PTCA/2x DES, kleine ACD 60 % Stenose.
- TTE 12/2016: Normale systolische LV-Funktion, EF 64 %, keine regionalen Wandbewegungsstörungen, kein relevantes Vitium, normale Rechtsherzabschnitte.
- 24-Stunden-Blutdruckmessung 12/2016: Ordentliche Aufzeichnungsqualität. Im Mittel leicht erhöhte systolische Blutdruckwerte tagsüber und nachts. Im 24-Stunden-Durchschnitt 133/75 mmHg. Normale diastolische Blutdruckwerte. Eingeschränkte nächtliche Blutdruckabsenkung.
- Koronarangiographie 1/2017: RCX: Gutes Langzeitresultat nach DES-Implantation 2012, ACD 60 % Stenose ohne relevante Progredienz zur Voruntersuchung.
- Normozytäre hyperchrome Anämie, ED 30.11.18
- am ehesten bei Folsäuremangel
- Vitamin B12-Mangel, ED 30.11.18

Kantonsspital Baselland
4101 Brudermolz T +41 (0)61 438 21 55 F +41 (0)61 438 30 70 www.ksbj.ch/medizin
medizin.brudermolz@ksbj.ch

- 2 -

Austrittsbericht vom 12.12.2018 / [REDACTED]

- Reizsymptomatik Wurzel S1 llnksseitig
- Substituierte Hypothyreose
- Grosses subkutanes Hämatom bei Status nach Sublay-Hernienplastik bei Narbenhernie am 23.05.18 ml/bei
- Status nach Sterilisation und Appendektomie via Pfannenstielinzision 1973
- Status nach vaginaler Hysterektomie bei Litoreumtum und Zystoepile 1982

Cohort Structuring

Search Profiles



What's next

Correlation \neq Causation

AND

Data \neq Knowledge

→ Data validation, critically question findings, put results into domain-specific context!

5 principles to build on to

- **Human role:** stand on the power of computers, don't compete with them (**Use-cases, research question**)
- **Mainstream:** Establish a mainstream, core computational thinking curriculum for all (**Curriculum**)
- **Realistic:** Use the real-world as your guide for what to learn: concepts, strategies and toolsets (**Data quality, evaluation of data**)
- **Toolset:** Priorities breadth of computational applications over the details of their implementation (**Search strategies, ML (Classifier), LogicEngine**)
- **Urgency:** Implement computational thinking education – society's key preparation for the AI age.

Key essence: Applying the available tools considering scientific methods, critically question common methods and don't use them just because it has always been done that way.