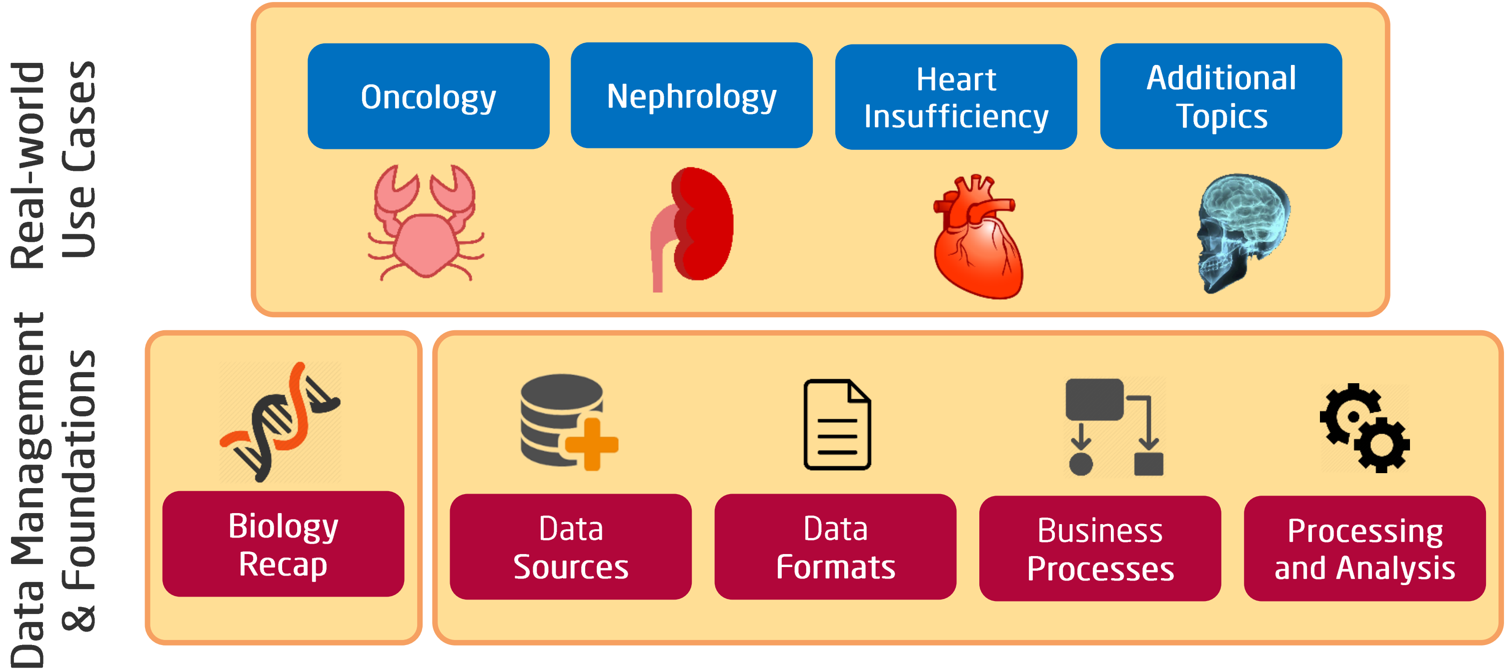
IT Systems Engineering | Universität Potsdam

Use Case Nephrology

Harry Freitas da Cruz

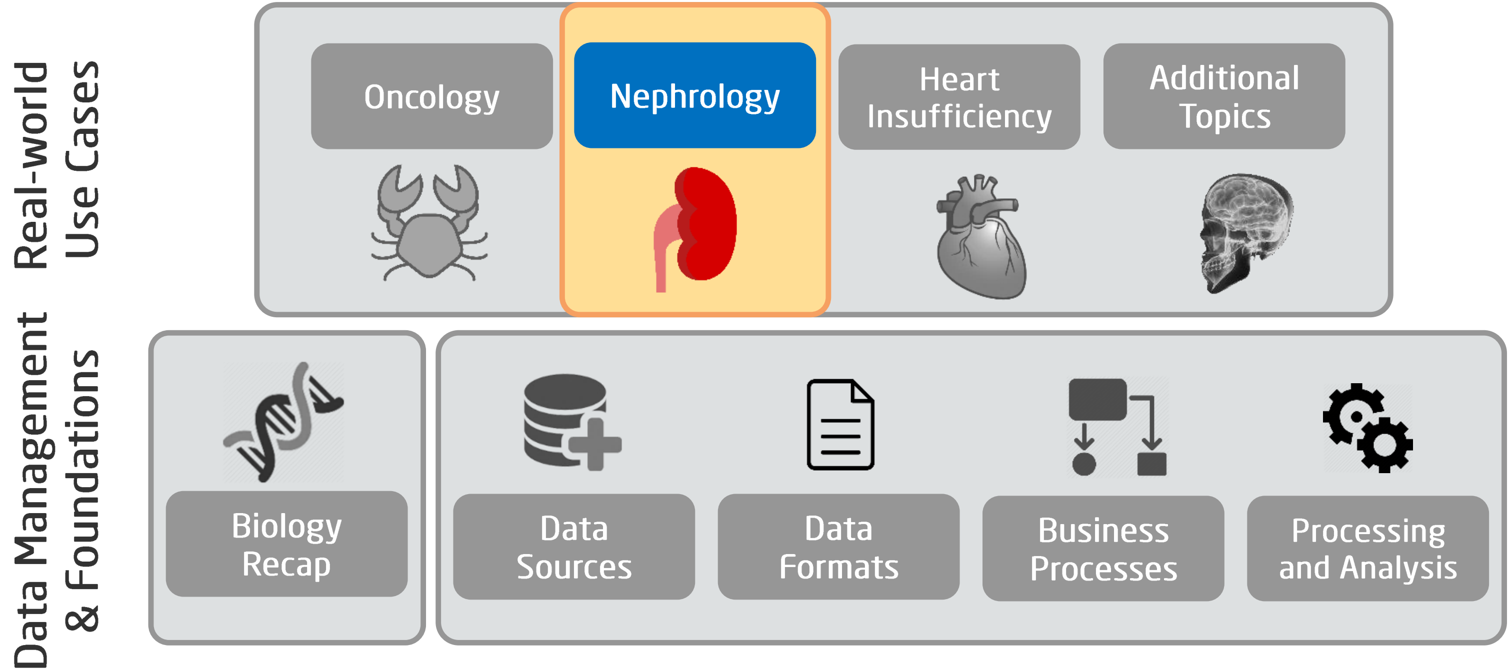
Data Management for Digital Health
Summer 2017

Where are we?



Real-world Use Cases
Oncology
Nephrology
Heart Insufficiency
Additional Topics
Data Management & Foundations
Biology Recap
Data Sources
Data Formats
Business Processes
Processing and Analysis

Where are we?


Real-world Use Cases
Oncology
Nephrology
Heart Insufficiency
Additional Topics
Data Management & Foundations
Biology Recap
Data Sources
Data Formats
Business Processes
Processing and Analysis

Recap: Data Sources and Formats

- Genomic data formats
  - Sequencing: FASTQ
  - Alignment: FASTA
  - Variant calling: VCF
  - Annotation: GTF
- Medical data formats
  - HL7
  - IHE
  - DICOM
  - FHIR
  - CDISC ODM

```
@SRR831012.1 HWI-ST155_(
NGAGATGAAGCACTGTAGCTTG(
```

```
##fileformat=VCFv4.1
##fileDate=20140930
##source=23andme2vcf.pl
##reference=file://23and
##FORMAT=<ID=GT,Number=1
#CHROM   POS ID   REF ALT
```

IHE
Integrating the Healthcare Enterprise

Stiftung Münch, Studie zur elektronischen Patientenakte im Ausland (2015)

Agenda

- The basics: Nephrology primer
- When things go awry: Kidney diseases
- Use case: prediction of acute kidney injury (AKI)
- Machine learning in Nephrology

Big data, big problems

Hasso Plattner Institut

"You really need to know something about medicine. If statistics lie, then Big Data can lie in a very, **very big way**"[1].

[1] MIT editors Business Report: Data-driven Health Care. MIT Technol Rev. 2014;117:1-19

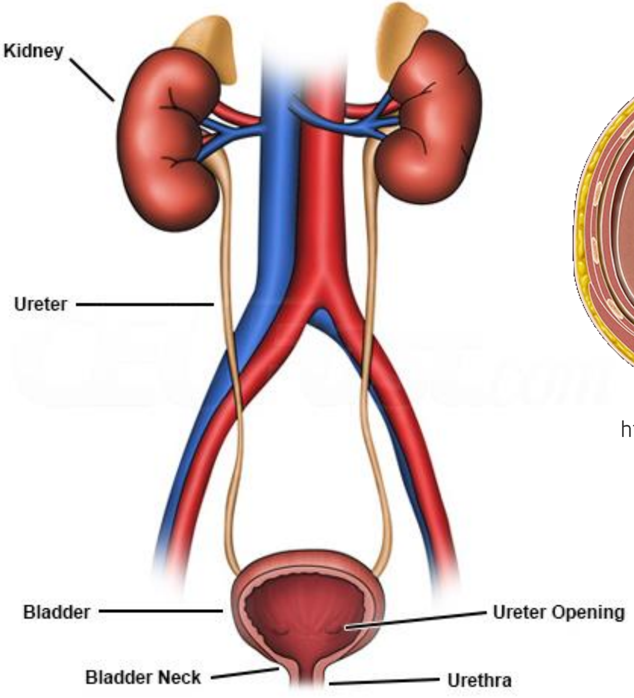
Nephrology Primer
Summer is Coming!

[1] http://landschloss-fasanerie.com/wp-content/uploads/2016/05/grillen-1.jpg
[2] http://www.duden.de/_media_/full/B/Biergarten-201100279875.jpg

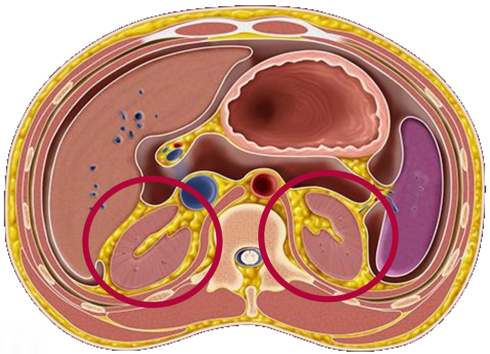
Nephrology Primer
Kidneys as Filters?

Nephrology Primer
The Urinary System

- Regulating water volume and pH levels
- Influencing red cell production
- Mediating blood pressure
- Helping you in danger: adrenaline (i.a.)
- Removing **waste**



https://ceufast.com/course/urinary-tract-infections-the-unappreciated-giant

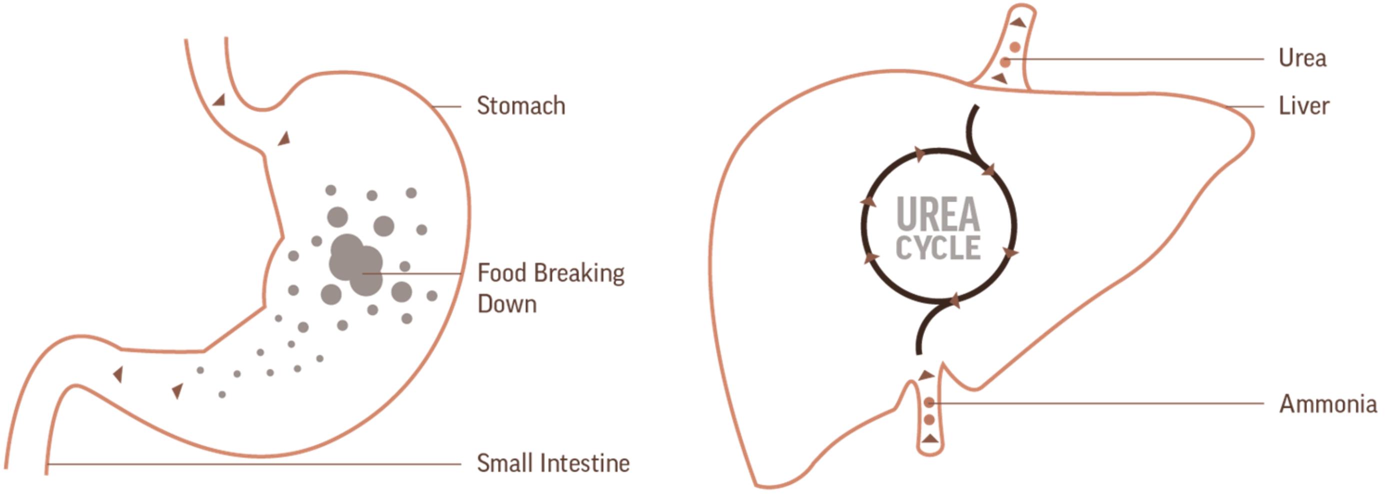


https://web.stanford.edu/dept/radiology/radiologysite/site170.html

https://www.urmc.rochester.edu/encyclopedia/content.aspx?ContentTypeID=85&ContentID=P01468

Nephrology Primer
Source of the „Waste“: Urea Cycle

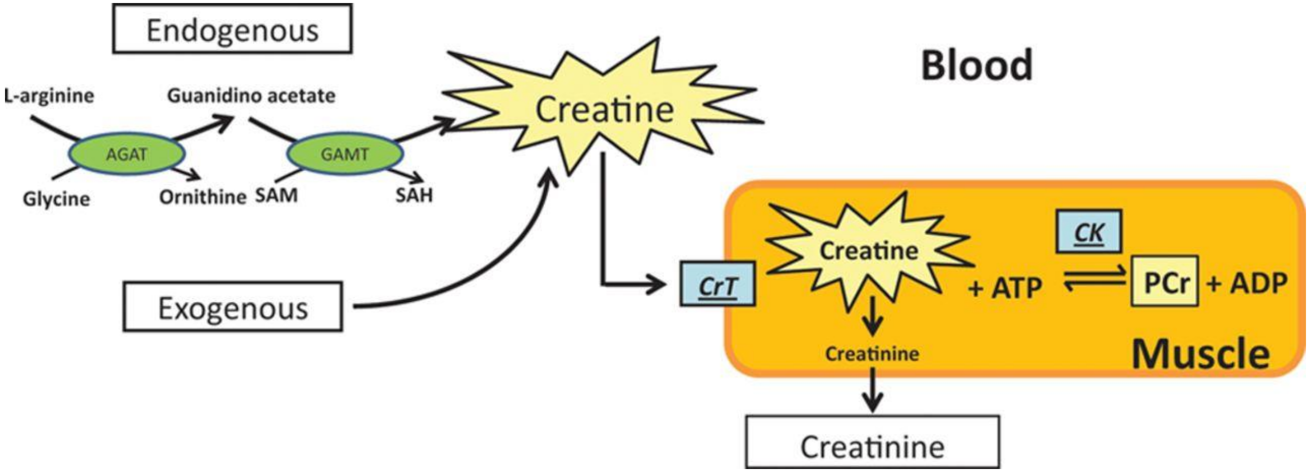
- Amino acids‘ metabolism results in toxic ammonia: NH_3
- Ammonia is metabolized in the liver into non-toxic urea: $CO(NH_2)_2$



https://www.ravicti.com/Content/images/RAVICTI_Graphic_Elements_fi_Protein_and_Cycle.png

Nephrology Primer
Source of the „Waste“: Creatinine

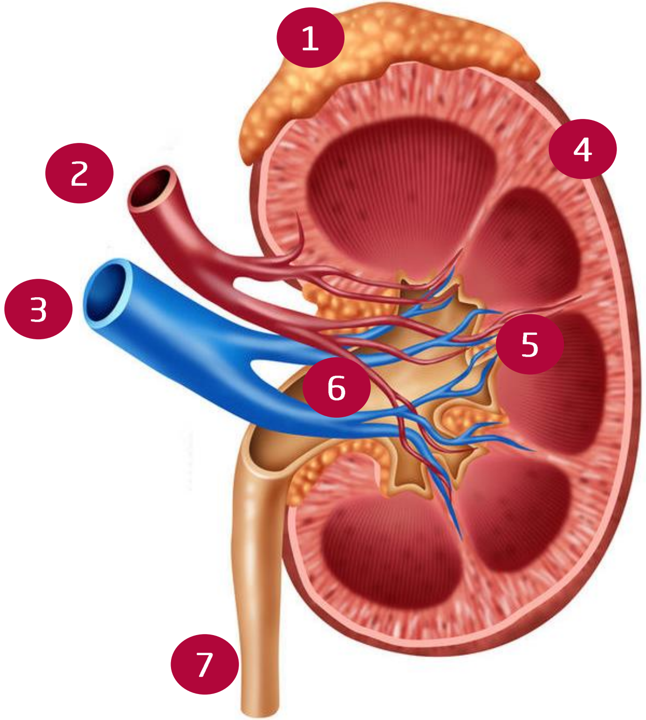
- Creatine kinase (CK) reaction: energy for the muscles as Phosphocreatine (PCr)
- Creatinine is the result of muscle contractions and brain effort
- Normal values depend on age, gender, and other factors: 0.6 to 1.2 mg/dl



Heinrich Taegtmeyer, Joanne S. Ingwall. Creatine–A Dispensable Metabolite? Circulation Research. 2013

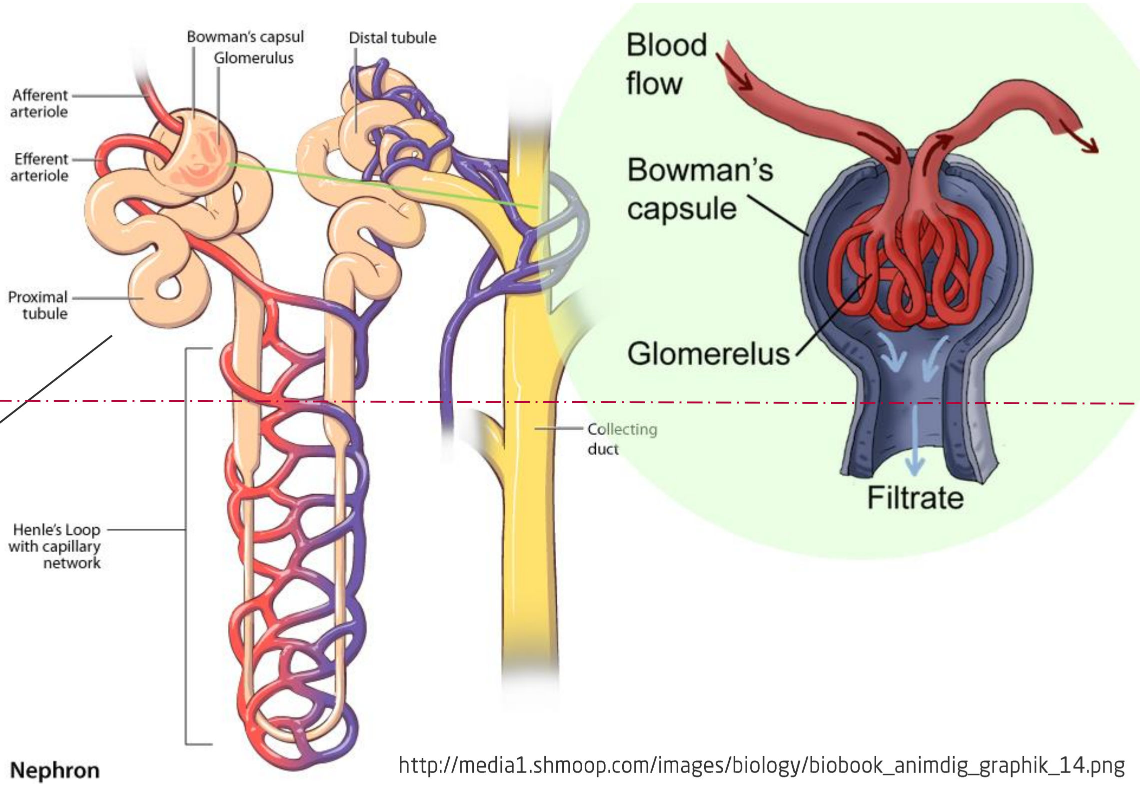
Nephrology Primer
The Kidney

- Adrenal gland (1)
- Renal artery (2)
- Renal vein (3)
- Renal cortex (4)
- Renal medulla (5)
- Renal pelvis (6)
- Urethra (7)



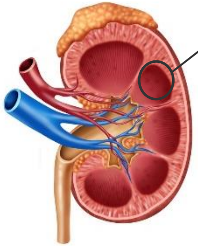
https://edc2.healthtap.com/ht-staging/user_answer/reference_image/3694/large/Kidney.jpeg

Nephrology Primer
The Filtration Unit: Nephron



Renal cortex

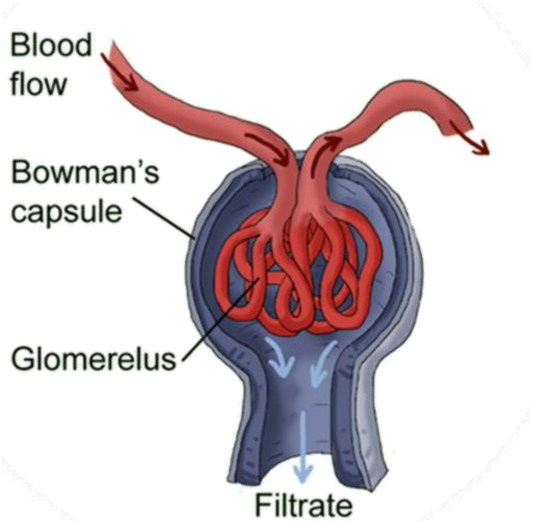
Renal medulla



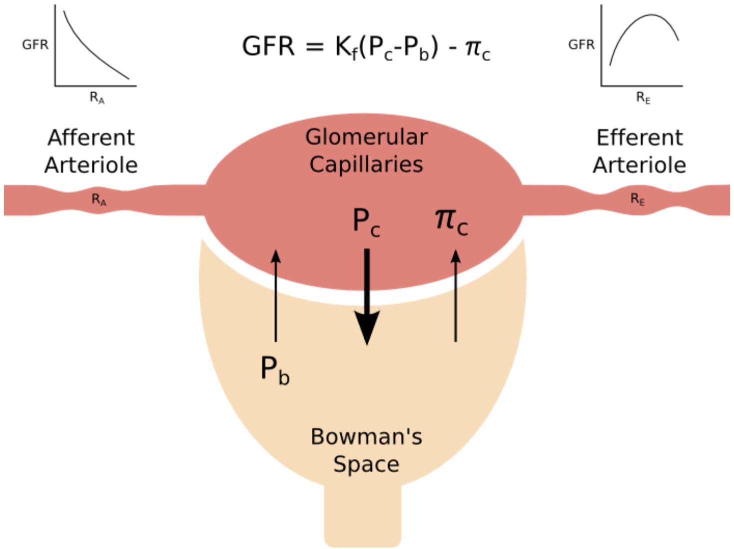
http://media1.shmoop.com/images/biology/biobook_animdig_graphik_14.png

Nephrology Primer
Glomerular Filtration Rate

- Volume of fluid filtered per unit of time (ml/min)
- GFR is influenced by capillary hydrostatic pressure



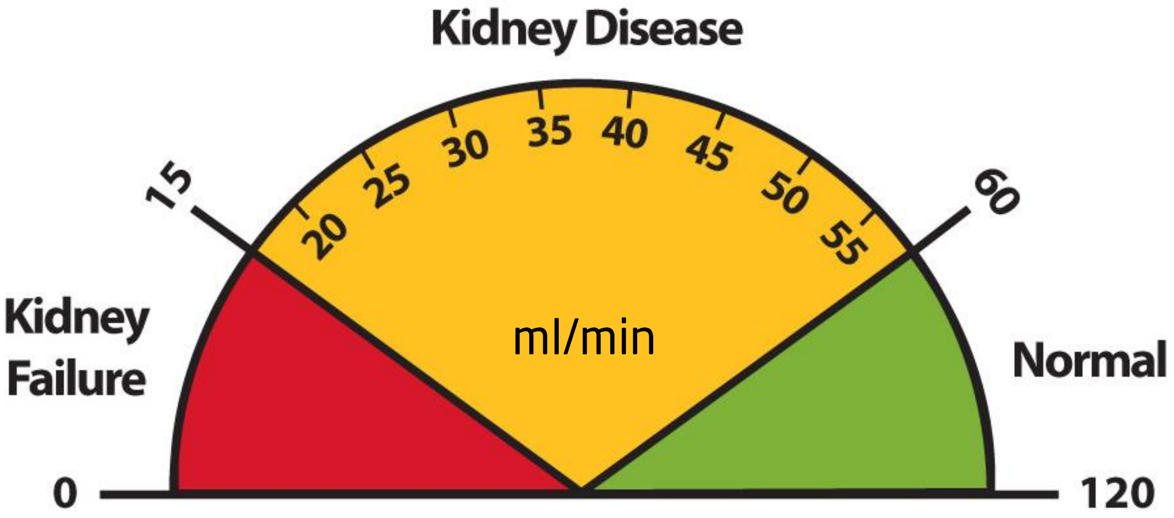
http://media1.shmoop.com/images/biology/biobook_animdig_graphik_14.png



http://www.pathwaymedicine.org/glomerular-filtration-rate

Nephrology Primer
Glomerular Filtration Rate

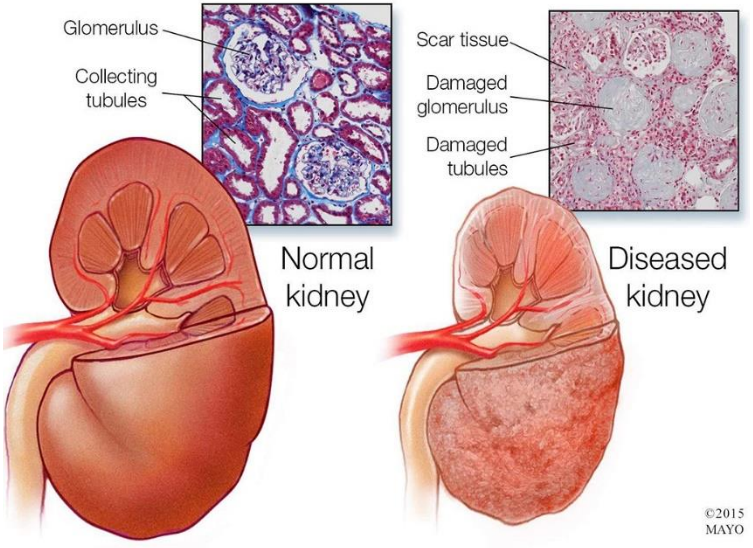
- Important measure of kidney function
- Variations within normal range *may* indicate future disease



National Institute of Diabetes and Digestive and Kidney Diseases, Understanding GFR (2012)

Kidney Disease(s)

- 17th among the causes of deaths globally[1]
- 9th death cause in the US[2]
- Some of the most common:
  - Chronic kidney disease (CKD)
  - Acute kidney injury (AKI)
  - Diabetic nephropathy
  - Glomerulonephritis
  - Kidney stones
  - …. and 290+ more[3].



http://www.youffyhealth.com/blogimage/1457956539.jpg

[1] GBD 2015 Mortality and Causes of Death Collaborators. Global, regional, and national life expectancy, all-cause mortality, and cause-specific mortality for 249 causes of death, 1980–2015: a systematic analysis for the Global Burden of Disease Study 2015. Lancet. 2016; 388: 1459–1544
[2] http://www.cbsnews.com/news/the-leading-causes-of-death-in-the-us/
[3] http://www.kidney.nyc/types-of-kidney-disease/

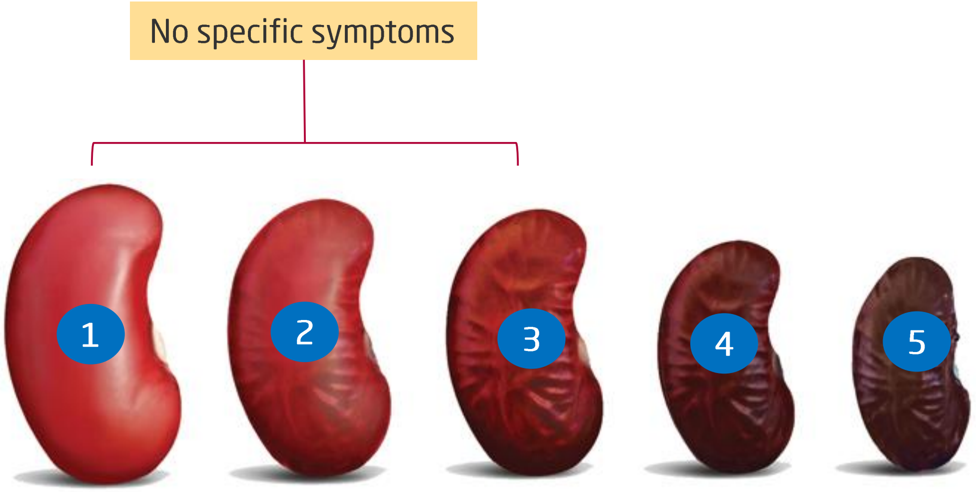
Kidney Disease(s)
Chronic Kidney Disease (CKD)

- Prolonged, sustained loss of renal function (> 3 months)
- GFR lower than 60 ml/min
- Risk factors
  - Diabetes, hypertension, heart disease, obsesity
  - Age, gender, African-American descent

http://sunlightpharmacy.com/wp-content/uploads/2017/03/CKD.jpg

Kidney Disease(s)
Chronic Kidney Disease (CKD): “Silent Killer”

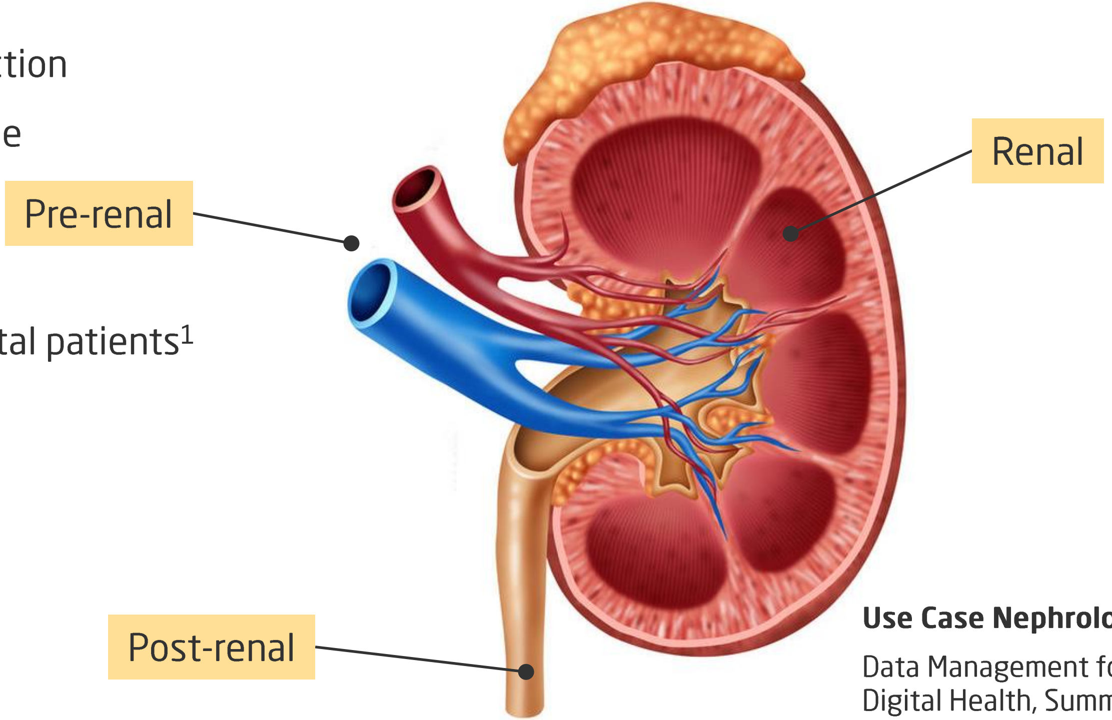
- CKD stages: 1 – 5, according to severity (GFR values)
- No definite symptoms until advanced stage of the disease
- Symptoms: fatigue, confusion, itching, etc.



http://sunlightpharmacy.com/wp-content/uploads/2017/03/CKD.jpg

Kidney Disease(s)
Acute Kidney Injury (AKI)

- Sudden and severe drop of renal function
- Increased levels of urea and creatinine
- May or may not be reversible
- Leads to poor patient outcomes
- Affects between 7 and 18% of hospital patients[1]
- Etiology
  - Pre-renal
  - Renal
  - Post-renal

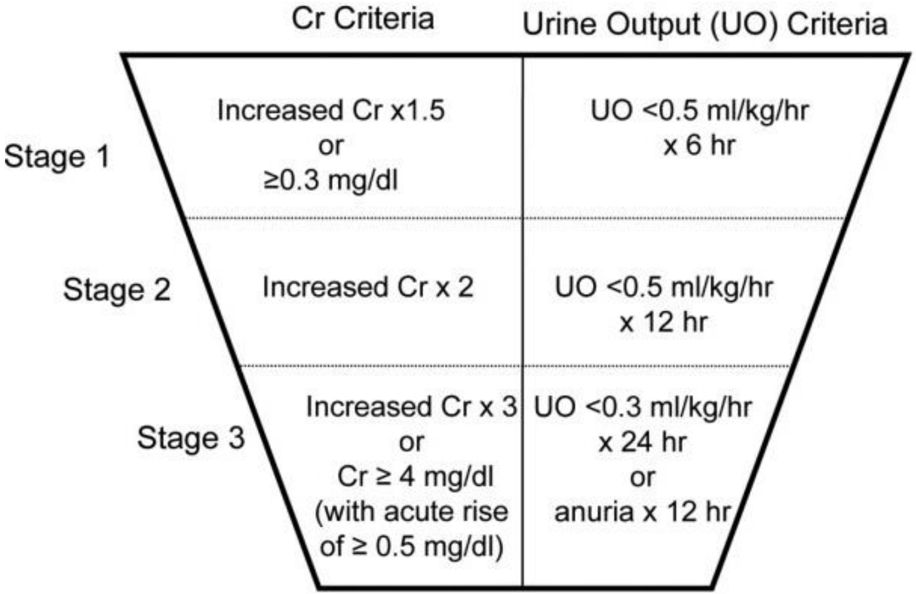


https://edc2.healthtap.com/ht-staging/user_answer/reference_image/3694/large/Kidney.jpeg

Lewington AJ, Cerdá J, Mehta RL. Raising Awareness of Acute Kidney Injury: A Global Perspective of a Silent Killer. Kidney international. 2013;84(3):457-467. doi:10.1038/ki.2013.153.

Kidney Disease(s)
Acute Kidney Injury: Staging Systems

- Define severity of AKI[1]
- Associated with patient outcomes
- Serve as reference for patient care and triage
- RIFLE (Risk, Injury, Failure, Loss, End-stage)
- AKIN (Acute Kidney Injury Network)
  - Stage 1, 1.5x SCr ⇧
  - Stage 2, 2x SCr ⇧
  - Stage 3, 3x SCr ⇧



Cruz DN, Ricci Z, Ronco C. Clinical review: RIFLE and AKIN--time for reappraisal. (2009)

[1] http://www.nature.com/nrneph/journal/v7/n4/fig_tab/nrneph.2011.14_T1.html

Kidney Disease(s)
Acute Kidney Injury (AKI)

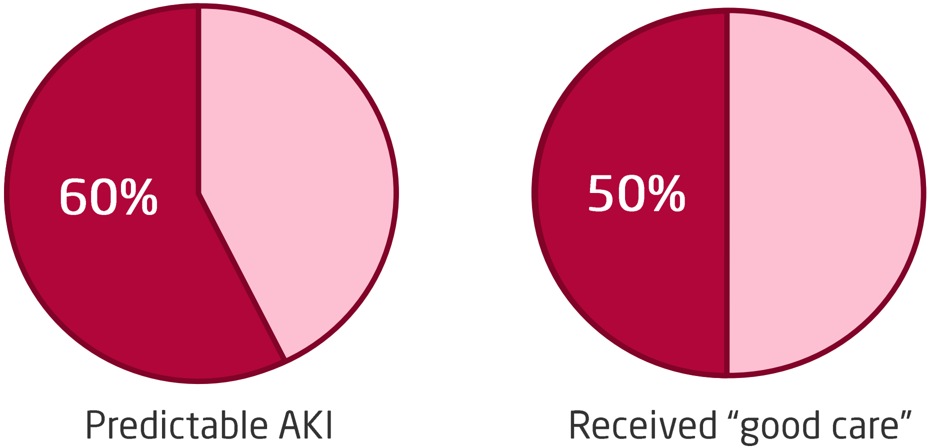
- Currently in Germany[1]
  - 70.000 patients / 2,5 Mio. EUR p.a.
  - 100.000 patients by 2020
- Kidney disease is asymptomatic (silent)
- Severe implications for patients
- Higher risk of mortality
- Very high medical costs for dialysis

Source: Anna Frodesiak, CC0

[1] http://www.aerzteblatt.de/nachrichten/41258/Zahl-der-Dialysepatienten-steigt

Kidney Diseases
Acute Kidney Injury (AKI)

- Key stats for AKI[1]
  - 50% deemed to receive "good care"
  - 60% of post-admission AKI predictable

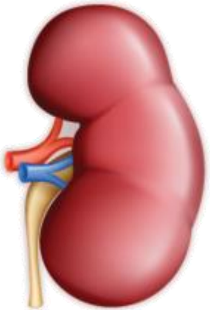


[1] National Confidential Enquiry into Patient Outcome and Death, UK (NCEPOD, 2009)

Kidney Diseases

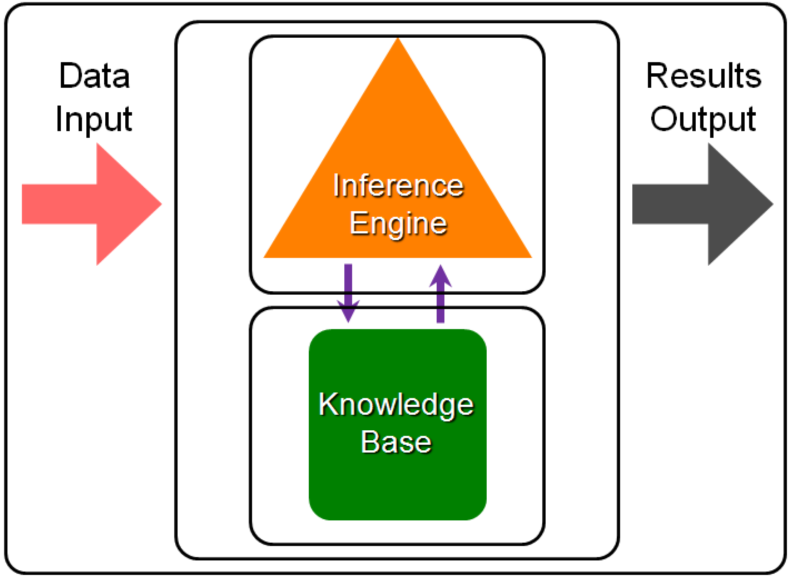
- Can we help physicians diagnose AKI **before** its onset?

Source: techcrunch.com



Use Case Nephrology
Clinical Decision Support Systems

- Data input
  - EMR data
- Inference engine
  - Rule-based
  - Decision trees (forest)
  - Neural networks
  - Bayesian networks
- Output
  - Diagnosis, alerts, etc.

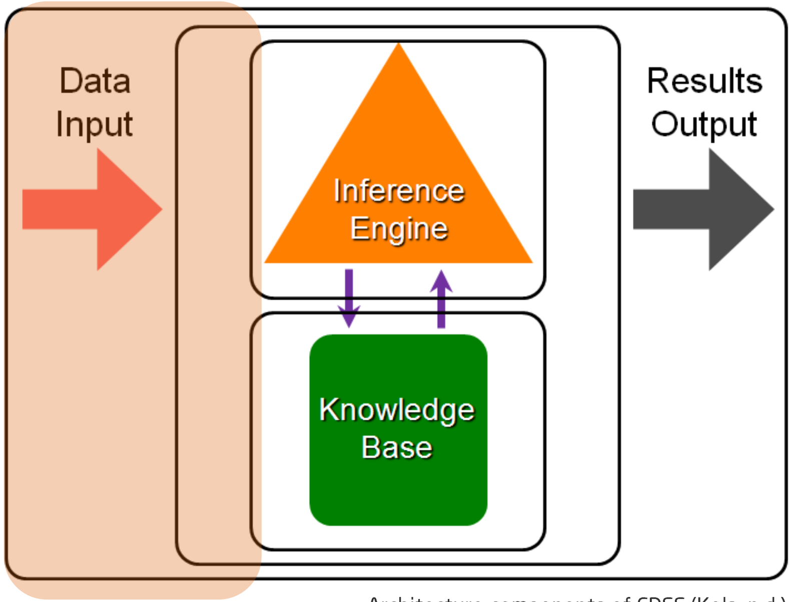


Architecture components of CDSS (Kola, n.d.)

Wilfred Bonney (2011). Impacts and Risks of Adopting Clinical Decision Support Systems, Efficient Decision Support Systems – Practice and Challenges in Biomedical Related DomainProf. Chiang Jao (Ed.), InTech, DOI: 10.5772/16265.

Use Case Nephrology[1]
Predicting Risk of AKI: Model Input

- Demographics
  - Gender, age, ethinicity
- Comorbidities
  - Heart disease, hypertension
  - Liver disease, pulmonary issues
  - CKD, diabetes, obesity, etc.
- Lab values
  - Glomerular filtration rate
  - Serum creatinine

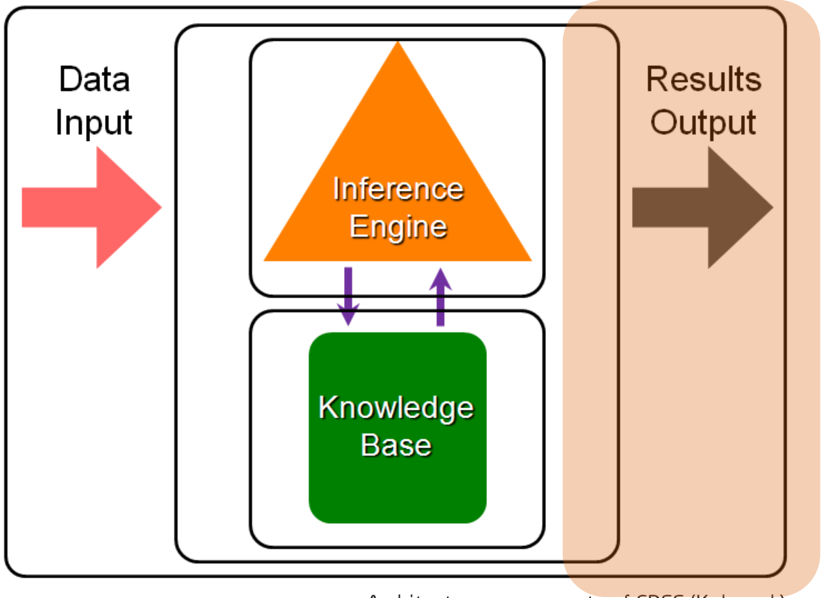


Architecture components of CDSS (Kola, n.d.)

[1] Cruz, Harry; Grasnick, B.; Dinger, H.; Bier, F.; Meinel, C. Early detection of acute kidney injury with Bayesian networks. 7th International Symposium on Semantic Mining in Biomedicine (2016)

Use Case Nephrology
Predicting Risk of AKI: Model Output

- Probabilities for:
  - Presence of AKI (yes or no)
  - Onset of renal failure (yes or no)
  - Risk, Injury, Failure, Loss
    - RIFLE[1] guideline
  - AKI Stage 1, 2 or 3
    - AKIN[1] guideline

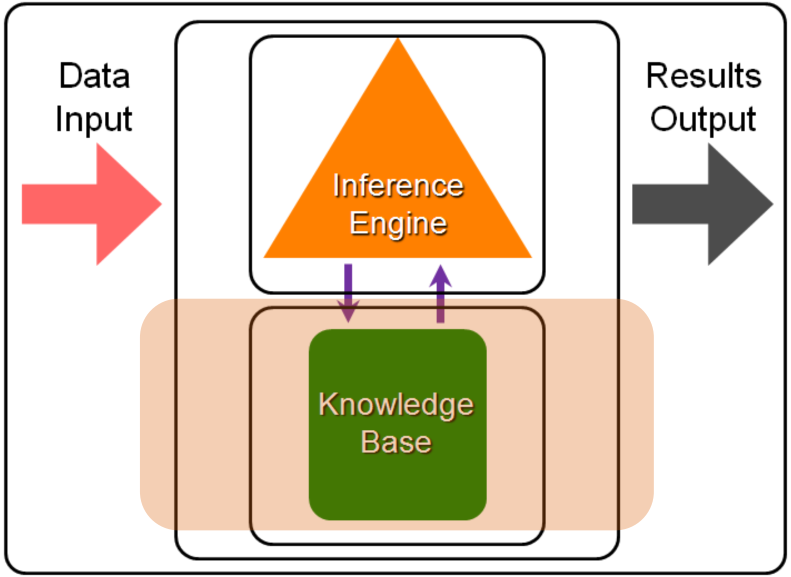


Architecture components of CDSS (Kola, n.d.)

[1] Cruz, Harry; Grasnick, B.; Dinger, H.; Bier, F.; Meinel, C. Early detection of acute kidney injury with Bayesian networks. 7th International Symposium on Semantic Mining in Biomedicine (2016)

Use Case Nephrology
Predicting Risk of AKI: Knowledge Base

- MIMIC II Database
  - Open database (MIT)
  - Beth Israel Deaconess Hospital
- ICU clinical data
  - 48,000 patients in total
  - ~5,000 AKI patients
- Available data
  - Bedside monitoring
  - Lab tests, orders
  - Demographics

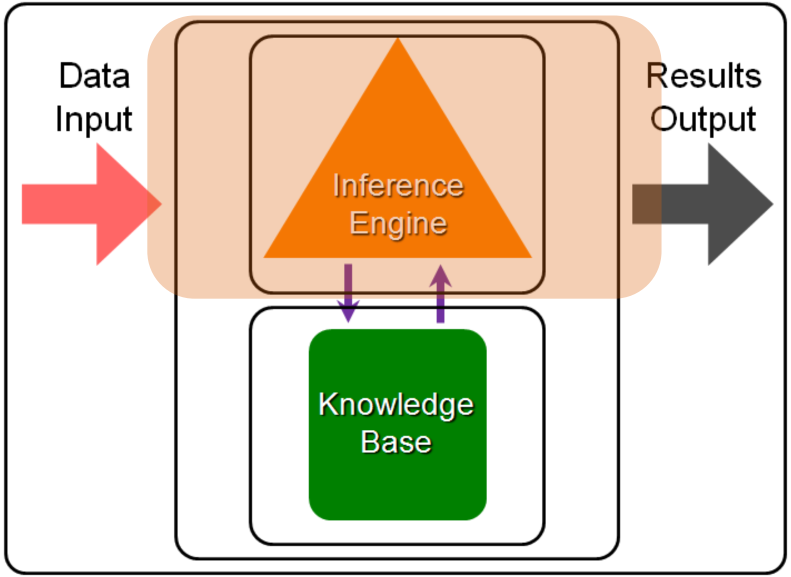


Architecture components of CDSS (Kola, n.d.)

[1] Cruz, Harry; Grasnick, B.; Dinger, H.; Bier, F.; Meinel, C. Early detection of acute kidney injury with Bayesian networks. 7th International Symposium on Semantic Mining in Biomedicine (2016)

Use Case Nephrology
Predicting Risk of AKI: Inference Engine

- Bayesian model
- Probabilistic relationships among variables of interest
- Causal relationships
- Combination of prior knowledge (causal) and data (probabilistic)

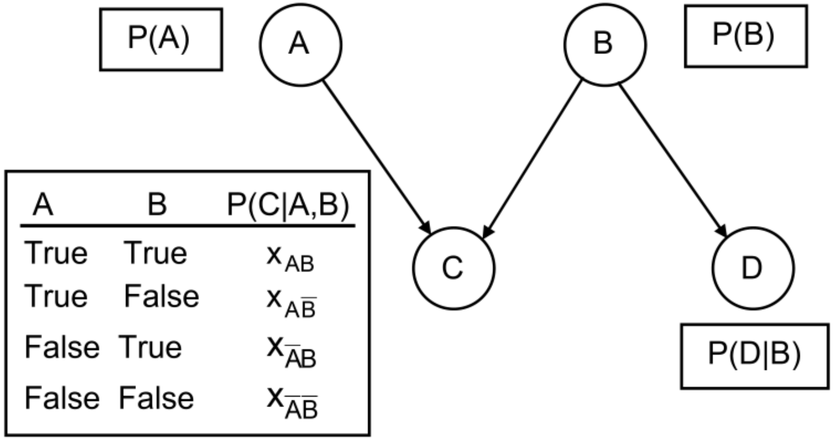


Architecture components of CDSS (Kola, n.d.)

[1] Cruz, Harry; Grasnick, B.; Dinger, H.; Bier, F.; Meinel, C. Early detection of acute kidney injury with Bayesian networks. 7th International Symposium on Semantic Mining in Biomedicine (2016)

Use Case Nephrology
Predicting Risk of AKI: Inference Engine

- Bayesian model
- Probabilistic relationships among variables of interest
- Causal relationships
- Combination of prior knowledge (causal) and data (probabilistic)

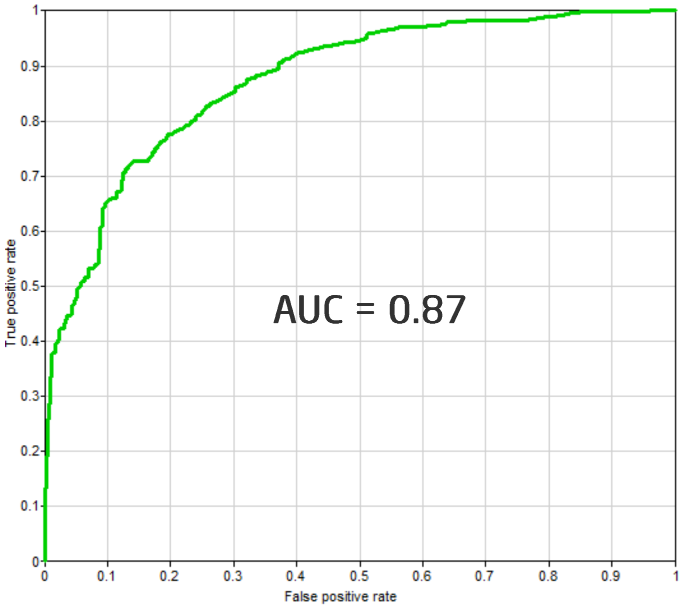


Emanuela Barbini, Pietro Manzi and Paolo Barbini (2013).

[1] Cruz, Harry; Grasnick, B.; Dinger, H.; Bier, F.; Meinel, C. Early detection of acute kidney injury with Bayesian networks. 7th International Symposium on Semantic Mining in Biomedicine (2016)

Use Case Nephrology
Predicting Risk of AKI: Inference Engine

- Bayesian model
- Probabilistic relationships among variables of interest
- Causal relationships
- Combination of prior knowledge (causal) and data (probabilistic)

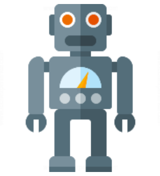


ROC curve for AKI onset [1]

[1] Cruz, Harry; Grasnick, B.; Dinger, H.; Bier, F.; Meinel, C. Early detection of acute kidney injury with Bayesian networks. 7th International Symposium on Semantic Mining in Biomedicine (2016)

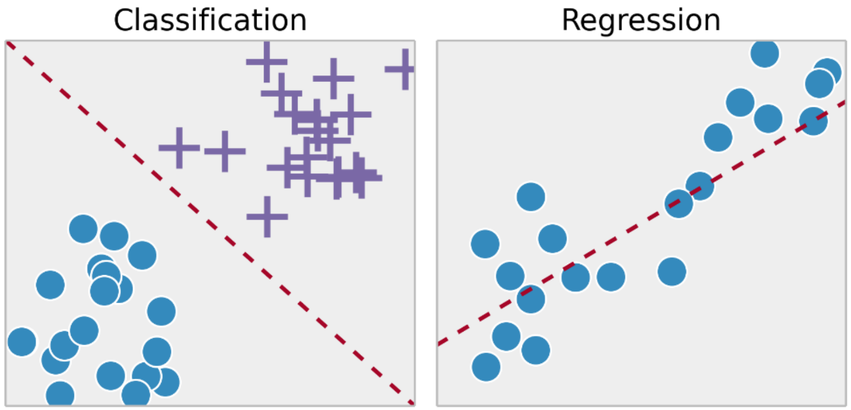
Machine Learning for Nephrology Literature Review

Publication	Method	Topic	Tool
Ramya, S. Radha, N. (2016)	Back-Propagation Neural Network; Radial Basis Function and Random Forest	Chronic Kidney Disease (CKD)	
Cruz, H., Grasnick, B., Dinger H. et al. (2016)	Bayesian networks	Acute Kidney Injury	Genie, Weka
Vijayarani, S., & Dhayanand, S. (2015).	Support Vector Machines and Neural networks	Predict kidney diseases (Acute Nephritic Syndrome, Chronic Kidney disease, Acute Renal Failure, Chronic Glomerulonephritis)	MATLAB
Vijayarani, S., & Dhayanand, S. (2015).	Naïve Bayes and Support Vector Machine	Predict kidney diseases (Acute Nephritic Syndrome, Chronic Kidney disease, Acute Renal Failure, Chronic Glomerulonephritis)	MATLAB
Sinha, P., & Sinha, P. (2015).	K-Nearest Neighbors and Support Vector Machines	Chronic kidney disease prediction	MATLAB
Baby, P. S., & Vital, P. (2015).	AD Trees, J48, K-means, Random Forest, Naive Bayes	Populational risk factors for Kidney disease	Weka, Orange
Lakshmi, K. R., Nagesh, Y., & Veerakrishna, M. (2014).	Artificial Neural Networks, Decision tree and Logical Regression	Dialysis survival	Tanagra tool
Greco, R., Papalia, T., Lofaro, D., Maestripieri, S., Mancuso, D., & Bonofiglio, R. (2010).	Decision trees	Transplant follow-up	Not mentioned
Koyuncugil, A. S., & Ozgulbas, N. (2010).	Association rule	Donor matching	Not mentioned
Bellazzi, R., Larizza, C., Magni, P., & Bellazzi, R. (2005).	Association rule discovery and temporal rule discovery are applied to the time series	Dialysis quality	Not mentioned
Kusiak, A., Dixon, B., & Shah, S. (2005).	Rough-set (RS) algorithm, Decision trees	Survival time dialysis	Not mentioned
Shah, S., Kusiak, A., & Dixon, B. (2003).	Rough-set (RS) algorithm	Survival time dialysis	Not mentioned
Agar, J. W., & Webb, G. I. (1992)	DLG	Renal biopsy	Not mentioned

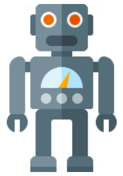


Machine Learning for Nephrology
Common Tasks

- Classification
  - Onset of CKD and AKI
  - Renal biopsy
  - Transplant follow-up
- Regression
  - Dialysis survivability
- Non-supervised
  - Risk factors for CKD
  - Donor matching

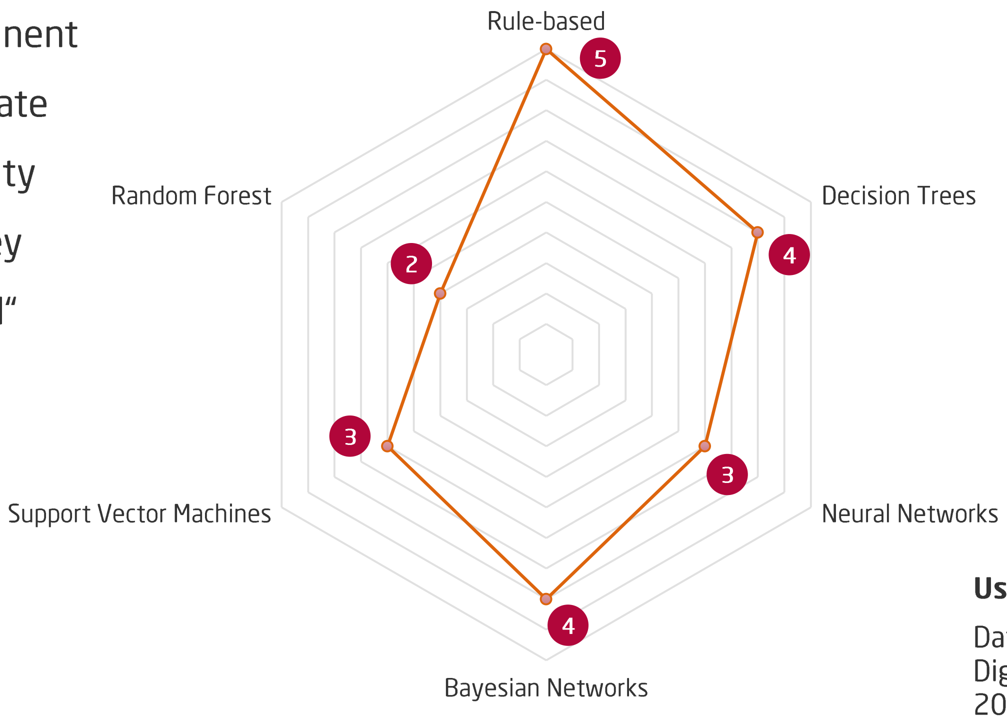


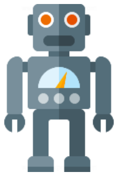
http://ipython-books.github.io/images/ml.png



Machine Learning for Nephrology
Algorithms Utilized

- Rule-based methods are prominent
- Tree-based approaches dominate
- SVM and NN limited applicability
- Interpretability of models is key
- Most models are „bench-based“
- MATLAB appears frequently





Machine Learning for Nephrology
Single-blind Study on AKI alerts

Automated, electronic alerts for acute kidney injury: a single-blind, parallel-group, randomised controlled trial

F Perry Wilson, Michael Shashaty, Jeffrey Testani, Iram Aqeel, Yuliya Borovskiy, Susan S Ellenberg, Harold I Feldman, Hilda Fernandez, Yevgeniy Gitelman, Jennie Lin, Dan Negoianu, Chirag R Parikh, Peter P Reese, Richard Urbani, Barry Fuchs

Lancet 2015; 385: 1966–74

Published Online February 26, 2015

http://dx.doi.org/10.1016/S0140-6736(15)60266-5

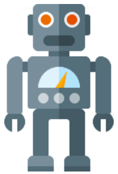
See Comment page 1924

Yale University School of Medicine, Program of Applied Translational Research

Summary

Background Acute kidney injury often goes unrecognised in its early stages when effective treatment options might be available. We aimed to determine whether an automated electronic alert for acute kidney injury would reduce the severity of such injury and improve clinical outcomes in patients in hospital.

Methods In this investigator-masked, parallel-group, randomised controlled trial, patients were recruited from the hospital of the University of Pennsylvania in Philadelphia, PA, USA. Eligible participants were adults aged 18 years or older who were in hospital with stage 1 or greater acute kidney injury as defined by Kidney Disease Improving Global Outcomes creatinine-based criteria. Exclusion criteria were initial hospital creatinine 4·0 mg/dL (to convert to



Interpretation: **did not** improve clinical outcomes among patients in that hospital

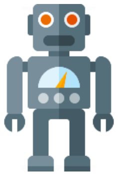
Use Case Nephrology

Data Management for Digital Health, Summer 2017

Wilson FP et al. Automated, electronic alerts for acute kidney injury: a single-blind, parallel-group, randomised controlled trial. Lancet (2015)

Machine Learning for Nephrology
There is still hope!

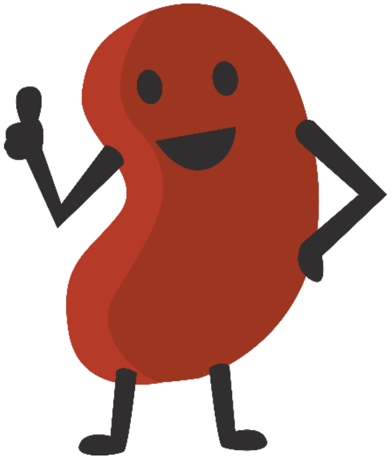
> These findings should inform the adoption of electronic alerting in the future. Specifically, future trials should examine novel diagnostic algorithms for acute kidney injury that might improve detection of individuals likely to progress to clinically meaningful endpoints. Also, studies that provide more direction regarding interventions and process measures could provide valuable intermediate



Wilson FP et al. Automated, electronic alerts for acute kidney injury: a single-blind, parallel-group, randomised controlled trial. Lancet (2015)

How to Avoid Kidney Disease?

- Appropriate hydration (drink enough water!)
- Prevent diabetes and hypertension: healthy diet
- Exercise more (never enough said!)
- Quit smoking (damages blood vessels)
- Beware of vitamins and supplements
- If under risk, get screened



http://cartoonsmix.com/cartoons/happy-kidney-cartoon.html

http://www.fmh.org/News-Calendar/News/2016/March/Understanding-Kidney-Disease-A-Silent-Killer.aspx

What to Take Home?

- „You need to know something about Nephrology“
- Applying Bayesian networks to AKI
- Machine learning is used in different tasks
- Validating the models is essential
- Drink more water!

What's Next?

- Health care analytics
- Clinical prediction models

To Know More

- Berner ES. Clinical decision support systems: State of the Art. AHRQ Publication No. 09-0069-EF. Rockville, Maryland: Agency for Healthcare Research and Quality. June 2009.
- Shortliffe, E. H., & Cimino, J. J. (Eds.). (2014). Biomedical Informatics. London: Springer London. doi:10.1007/978-1-4471-4474-8
- Hastie, T., Tibshirani, R., & Friedman, J. (2009). The Elements of Statistical Learning. Springer Series in Statistics (Vol. 1). Springer. http://doi.org/10.1007/b94608
- Clifford, G., Scott, D., & Villarroel, M. (2009). User guide and documentation for the MIMIC II database. MIMIC-II Database Version, 0500. Retrieved from http://physionet.nlm.nih.gov/mimic2/UserGuide/UserGuide.pdf
- Lehman, L., Moody, G., Heldt, T., & Kyaw, T. H. (2011). Multiparameter Intelligent Monitoring in Intensive Care II (MIMIC- II): A public-access intensive care unit database. Critical Care, 39(February 2010), 952-960. http://doi.org/10.1097/CCM.0b013e31820a92c6.Multiparameter
- Bates, D., Kuperman, G., Wang, S., Gandhi, T., Kittler, A., Volk, L., … Middleton, B. (2003). Ten Commandments for Effective Clinical Decision Support: Making the Practice of Evidence-based Medicine a Reality. Journal of the American Medical Informatics Association, 10(6), 523-530. http://doi.org/10.1197/jamia.M1370