Clinical Prediction Models - Current State of Knowledge

Klemens Budde
Charité Universitätsmedizin Berlin

NEPHROCAGE
German-Canadian consortium on AI for improved kidney transplantation outcome
2nd International NephroCAGE Symposium, Aug 16, 2022

Supported by: Federal Ministry for Economic Affairs and Climate Action on the basis of a decision by the German Bundestag

Why do we need prediction models?

- key elements for the **individual risk assessment:**
  - Predicting patients at risk for graft loss, complications (rejection, infection, malignancy) and death,
  - Estimating time to event: graft loss, complications, death
  - Aim is to avoid event or early detection
  - Importance for managing Immunosuppression, medication, psychosocial issues, preparation & planning for dialysis, re-transplant, etc.
- Use as surrogate endpoints for **group comparisons in clinical trials**

Supported by: Federal Ministry for Economic Affairs and Climate Action on the basis of a decision by the German Bundestag

How do we manage patients now?

- **Rely on clinical judgement with known risk factors**
  - Donor factors such as age, sex, comorbidities, medication
  - Recipient age, sex, weight, cause of death/living donor, comorbidities etc.
  - Transplant factors such as HLA mismatch, ischemia times, delayed graft function
  - Immunosuppression, co-medication
  - eGFR, GFR slope
  - Proteinuria
  - HLA antibodies
- **Manage complications, when they occur**

NEPHROCAGE
German-Canadian consortium on AI for improved kidney transplantation outcome
2nd International NephroCAGE Symposium, Aug 16, 2022

Supported by: Federal Ministry for Economic Affairs and Climate Action on the basis of a decision by the German Bundestag

Independent risk factors for graft loss in the ALERT-trial

Clincal consequences:
Stopp Smoking
treat
hypertension
diabetes
and lipids
look at
proteinuria
+ creatinine

TABLE 4. Independent risk factors for renal graft loss, renal graft loss or doubling of serum creatinine, and renal graft loss or doubling of serum creatinine or death (Cox multivariate analysis)

Risk factor	Hazard ratio (95% CI)	*P* value
Renal graft loss		
Proteinuria (per 1-g/24-hr increase)	1.76 (1.55–1.98)	<0.0001
Serum creatinine (per 100-μM increase)	4.37 (3.47–5.52)	<0.0001
Time since last transplant (yr)	1.06 (1.01–1.11)	0.027
Renal graft loss or doubling of serum creatinine		
Proteinuria (per 1-g/24 hr increase)	1.79 (1.60–2.01)	<0.0001
Scaled creatinine (per 100-μM increase)	3.43 (2.74–4.29)	<0.0001
Smoking	1.77 (1.21–2.60)	0.003
Pulse pressure (per 10 mm Hg)	1.12 (1.02–1.23)	0.021
Treatment for rejection	1.35 (0.96–1.89)	0.086
Renal graft loss or doubling of serum creatinine or death		
Proteinuria (per 1-g/24 hr increase)	1.59 (1.43–1.7)	<0.0001
Serum creatinine (per 100-μM increase)	2.74 (2.24–3.35)	<0.0001
HDL (mM)	0.74 (0.54–1.03)	0.0747
Age (yr)	1.03 (1.01–1.04)	<0.0001
Pulse pressure (per 10 mm Hg)	1.09 (1.01–1.17)	0.03
Diabetes mellitus	1.78 (1.32–2.41)	0.0002
Smoking	1.56 (1.14–2.13)	0.0053
Time since last RTx (yr)	1.05 (1.01–1.09)	0.0097
No. of RTx	1.51 (1.18–1.95)	0.0011

RTx, Renal transplantation.

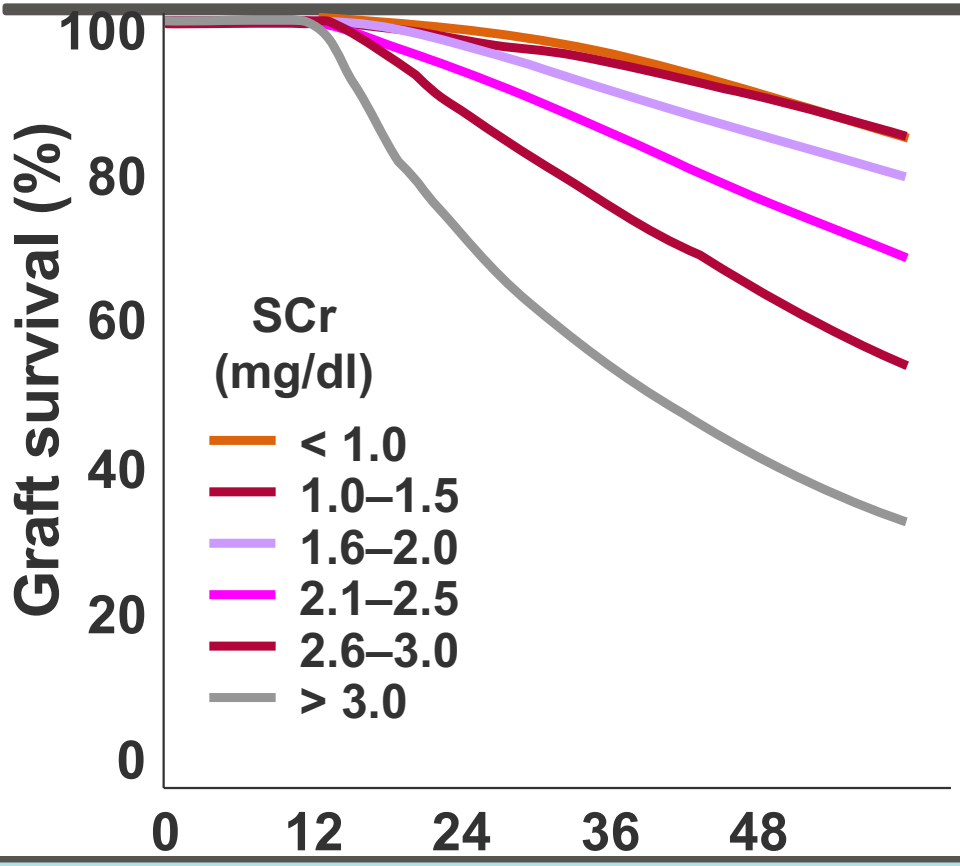
Graft loss:
Proteinuria
creatinine
time since Tx
+ doubling creatinine
Smoking
pulse pressure
rejection
+ death:
HDL
age
diabetes
No. of RTx

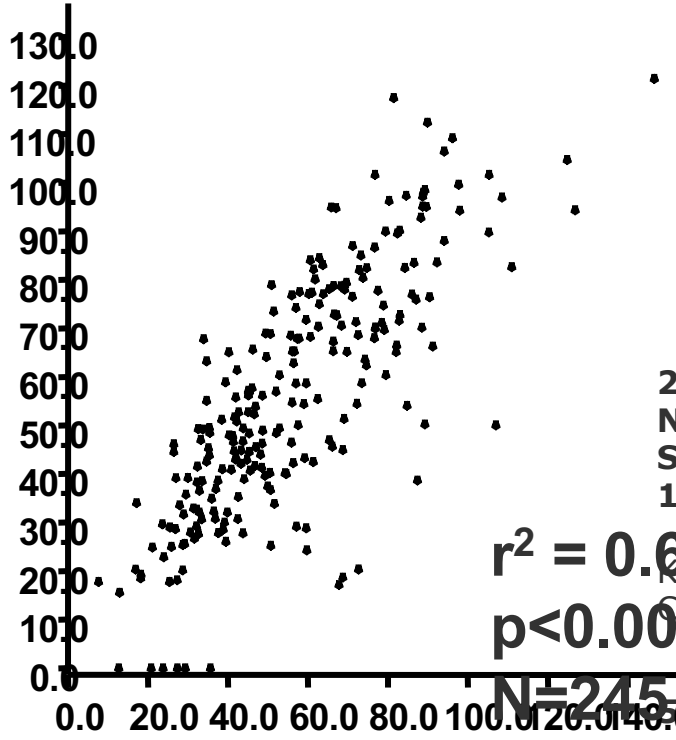
2nd Int'l NephroCAGE Symposium, Aug 16, 2022

Klemens Budde, Charité Berlin

Supported by: Federal Ministry for Economic Affairs and Climate Action on the basis of a decision by the German Bundestag

Serum creatinine after 1 year: a good predictor for graft-loss?





2nd Int'l NephroCAGE Symposium, Aug 16, 2022

$r^2 = 0.62$

$p<0.00001$

N=245

Klemens Budde, Charité Berlin

Supported by: Federal Ministry for Economic Affairs and Climate Action on the basis of a decision by the German Bundestag

Surrogate Endpoints for Late Kidney Transplantation Failure

Maarten Naesens[1*], Klemens Budde[2], Luuk Hilbrands[3], Rainer Oberbauer[4], Maria Irene Bellini[5], Denis Glotz[6], Josep Grinyó[7], Uwe Heeman[8], Ina Jochmans[1], Liset Pengel[9], Marlies Reinders[10], Stefan Schneeberger[11], Alexandre Loupy[12]

- Limitation of single parameters such as eGFR, Proteinuria, Histology, DSA and combined functional markers (e.g. eGFR+proteinuria) due to variability, interference with drugs, unclear cut-off values, controversial results….

- Potential of composite scores, but frequent lack of external validation and rigorous prospective studies….

TABLE 4 Value of composite scores as surrogacy for long-term graft survival [9,14,71–76]

Study	Kasiske *et al.*, 2010 [71]	Foucher *et al.*, 2010 [72]	Moore *et al.*, 2011 [73]	Schnitzler *et al.*, 2012 [74]	Shabir *et al.*, 2014 [74]; Gonzales *et al.*, 2016 [75]	Gonzales *et al.*, 2016 [75]	Prémaud *et al.*, 2017 [76]	Loupy *et al.*, 2019 [9]
Parameter	USRDS Risk Prediction Tool	KFRS	LOTESS Composite Risk Score	USRDS Predictive Model	Birmingham Risk Score	Birmingham-Mayo Histology-Based Model	AdGFS	iBox Risk Prediction Score
Development set	USRDS registry data (N=39,091)	Multicentre French registry (N=2169)	Multicentre national cohort study (N=2763)	USRDS registry data (N=87,575)	Single-center UK data (N=651)	Single-center US data (N=1465)	Single-center French data (N=664)	French multicentre cohort (N=4000)

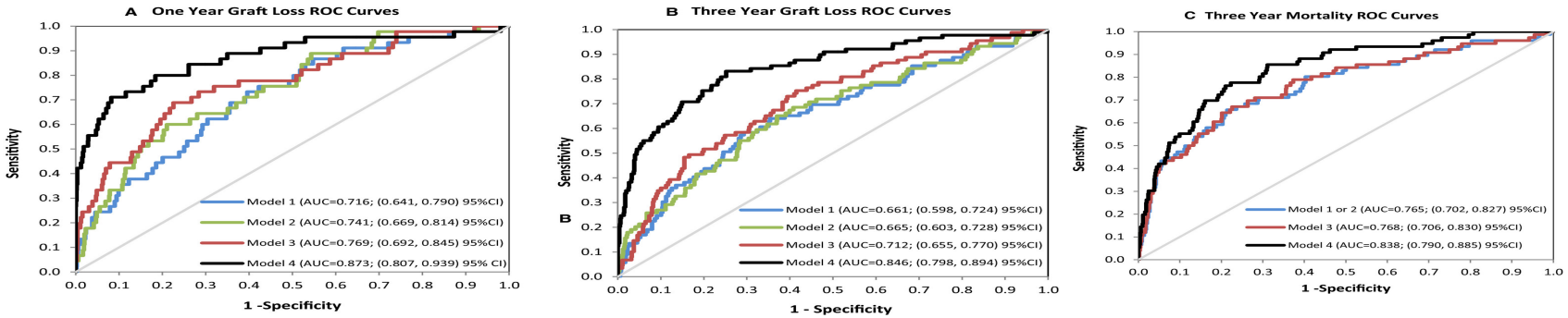
Naesens M et al Transplant International 2022

NEPHROCAGE
German-Canadian consortium on AI for improved kidney transplantation outcome
2nd International NephroCAGE Symposium, Aug 16, 2022

Supported by: Federal Ministry for Economic Affairs and Climate Action on the basis of a decision by the German Bundestag

Example for an individual prediction model

- Prediction of graft loss and mortality at 1 and 3 years
- 10% 3-year graft loss in 891 patients (2007–2015).
- Firth multivariable logistic regression
- eGFR trajectory: the absolute value of the slope from max. eGFR to eGFR at day 365
  - Data Model 1: UNOS only.
  - Data Model 2: UNOS + Transplant database.
  - Data Model 3: UNOS + Transplant database + EHR comorbidity.
  - Data Model 4: UNOS + Transplant database + EHR comorbidity + EHR posttransplant trajectory + natural language processing (Banff scores and vital signs)
- AUCs for model 4: 0,84-0,87



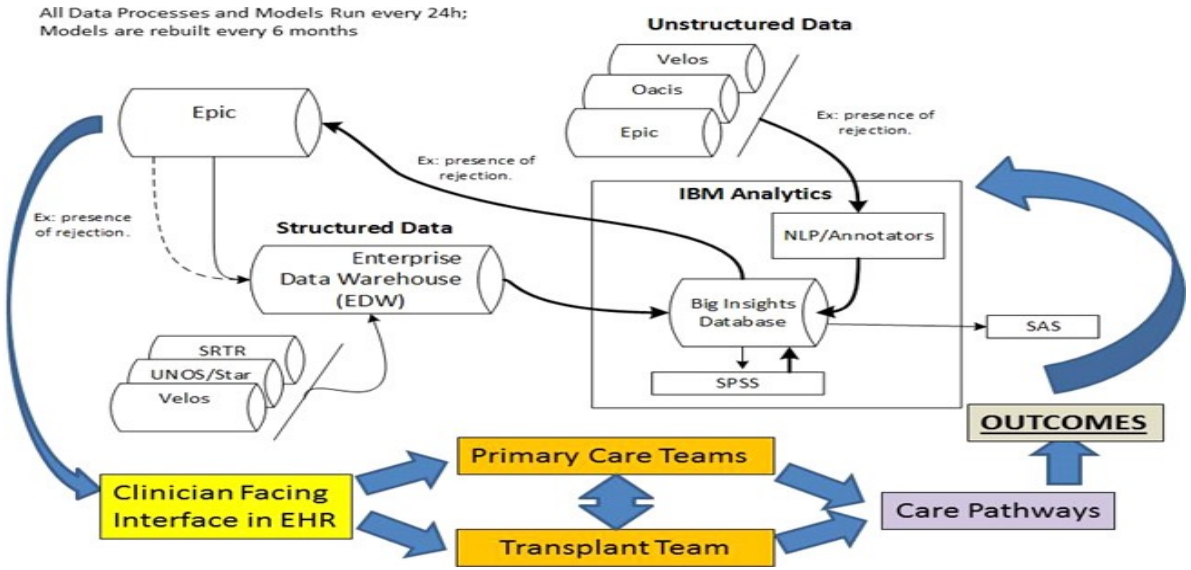
2nd Int'l NephroCAGE Symposium, Aug 16, 2022

Klemens Budde, Charité Berlin

Supported by: Federal Ministry for Economic Affairs and Climate Action
on the basis of a decision by the German Bundestag

Integration of prediction model into clinic

- **Inclusion of eGFR trajectories and dynamically evolving structured and unstructured longitudinal patient-level data using Big Data approaches improves the accuracy of prediction of graft loss and mortality**



Srinivas T et al. AJT 2017, 671-681

2nd Int'l NephroCAGE Symposium, Aug 16, 2022

Klemens Budde, Charité Berlin

Kidney Pilot

Data & Model

- Baseline: TBase data
  - includes patients of ~20 years
  - general: demographics, medications, diagnosis, ...
  - visits: vitals, clinical notes
  - lab data,hospitalizations
- >1400 patients, >100,000 data points
- Data Characteristics:
  - real life: noisy, unbalanced
- Method: Gradient Boosted Regression Trees
- about 1200 different features
- modelling of data complex, missing noisy data

Roland Roller et al.
under revision

2nd Int'l NephroCAGE Symposium, Aug 16, 2022

Klemens Budde,
Charité Berlin

Supported by: Federal Ministry for Economic Affairs and Climate Action
on the basis of a decision by the German Bundestag

Kidney Pilot

Data & Model

- Results on retrospective data for 90, 180, and 360 days

	days	Our Model	Esteban
Rejection	90	0.80 (0.04)	-
	180	0.82 (0.03)	0.778 (0.01)
	360	0.80 (0.04)	0.768 (0.01)
Tx-Loss	90	0.93 (0.02)	-
	180	0.94 (0.02)	0.821 (0.01)
	360	0.95 (0.01)	0.814 (0.01)
Infection	90	0.80 (0.01)	-
	180	0.78 (0.01)	-
	360	0.77 (0.01)	-

Roland Roller et al.
under revision

2nd Int'l NephroCAGE Symposium, Aug 16, 2022

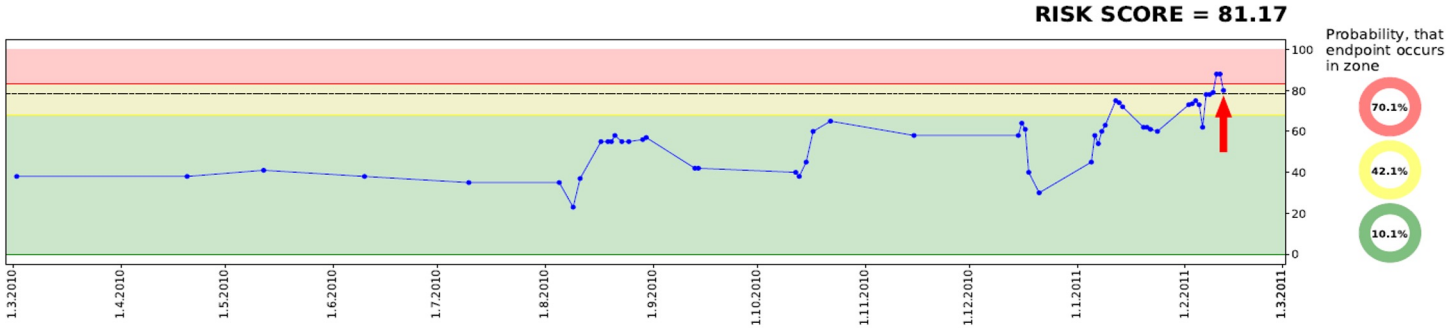
Klemens Budde,
Charité Berlin

Supported by: Federal Ministry for Economic Affairs and Climate Action
on the basis of a decision by the German Bundestag

Kidney Pilot

Data & Model visualization

- A dashboard with risk score, traffic light system and feature relevance
- Overview about the development of risk scores of one patient



Most influencial features for this decision:

Feature 1 with value (relevance: 22.49)
Feature 2 with value (relevance: 9.09)
Feature 3 with value (Relevanz: 7.88)
Feature 4 with value (relevance: 4.66)
Feature 5 with value (relevance: 4.63)

Most influencial features for overall model:

Feature A (relevance:35.12)
Feature B (relevance:23.76)
Feature C (relevance:8.99)
Feature D (relevance:3.04)
Feature E (relevance:2.26)

Roland Roller et al.

under revision

2nd Int'l NephroCAGE Symposium, Aug 16, 2022

Klemens Budde, Charité Berlin

Supported by: Federal Ministry for Economic Affairs and Climate Action
on the basis of a decision by the German Bundestag

iBox prediction model

- Development of a robust risk stratification model for better prediction and endpoint in trial design
- 7 European and 3 American centers with 7557 pts
- 7.12 years median follow-up
- 1067 (14,1%) graft failures
- Derivation cohort: 4000 consecutive French patients
- Validation cohort: 3557 patients from Europe and North America
- Transplanted between 2002 and 2014
- >100 prognostic factors for integrated data analysis
- **For 3, 5, and 7-year prediction at year 1 post Tx**
- Final hazard ratios from Cox model with 8 prognostic variables

A Loupy et al BMJ 2019: 366: 14923

2nd Int'l NephroCAGE Symposium, Aug 16, 2022

Klemens Budde, Charité Berlin

Supported by: Federal Ministry for Economic Affairs and Climate Action on the basis of a decision by the German Bundestag

iBox – in- and output

Data input		Output
Mandatory	Additional	
• Time from transplant • eGFR (mL/Min/1.73m2) • Proteinuria (g/g)	• Banff lesions grading: g,i,t,ptc,cg,IFTA • Histology diagnoses • Anti HLA DSA (MFI)	• Individual patient prediction of allograft survival 3, 5 and 7 years after evaluation time

Prediction system for risk of allograft loss in patients receiving kidney transplants: international derivation and validation study

- Loupy, Aubert, Orandi, et al., *BMJ* 2019;366:l4923
- http://dx.doi.org/10.1136/bmj.l4923

2nd Int'l NephroCAGE Symposium, Aug 16, 2022

Klemens Budde, Charité Berlin

Supported by: Federal Ministry for Economic Affairs and Climate Action on the basis of a decision by the German Bundestag

iBox prediction model

- Excellent c-statistics (AUC: 0,79-0,83)
- Similar values in validation cohorts
- Confirmed in 3 randomized controlled trials
- Dynamic prediction beyond the first year useful
- Score responsive to treatment changes (via eGFR/Proteinuria)
- Histology improved the model only marginally

Limitations

- Nomogramm and website (were) available, but now commercialized
- Prognostic, but how good for prediction in the individual??
- Not used in prospective trials
- HLA Mismatch, donor factors, DGF and rejection not in model
- Death excluded

2nd Int'l NephroCAGE Symposium, Aug 16, 2022

Klemens Budde, Charité Berlin

Supported by: Federal Ministry for Economic Affairs and Climate Action on the basis of a decision by the German Bundestag

External Validation of the Abbreviated Ibox Scoring System at One-Year Post Kidney Transplant as a Surrogate Endpoint for Death-Censored Graft Survival

A. Klein[1], L. Kosinski[2], R. Muse[2], A. Loupy[3], O. Aubert[3], G. Divard[3], M. Stegall[4], I. Helanterä[5], W. E. Fitzsimmons[1], I. O'Doherty[1]

[1]Transplant Therapeutics Consortium, Critical Path Institute, Tucson, AZ, [2]Critical Path Institute, Tucson, AZ, [3]Paris Transplant Group, Paris, France, [4]Mayo Clinic, Rochester, MN, [5]Helsinki University Hospital, Helsinki, Finland

Meeting: 2022 American Transplant Congress

Abstract number: 464

- Transplant Therapeutics Consortium
- new version **without biopsy @1 year post Tx**
- surrogate endpoint for **5-year death-censored** graft survival after kidney transplantation
- 4 data sets: 2 RCTs, 2 observational datasets
- CNI and CNI-free patients
- acceptable c-statistics (AUC: 0,70-0,84)

- CNI and CNI-free pts.
- acceptable c-statistics (AUC: 0,70-0,84)
- Potential as a surrogate endpoint for trials with group comparisons
- Public availability

Dataset	Subjects	C-statistic	Poisson Calibration				
			Observed graft loss	Predicted graft loss	Observed / Predicted	z score for Observed / Predicted	P-value
Observational	841	0.81	41	40.61	1.01	0.06	0.95
Mayo Clinic Rochester	344	0.84	21	16.19	1.30	1.19	0.23
Helsinki University Hospital	497	0.77	20	24.41	0.82	-0.89	0.37
RCTs	872	0.74	38	41.74	0.91	-0.58	0.56
BENEFIT RCT	515	0.70	15	18.77	0.80	-0.87	0.39
BENEFIT-EXT RCT	357	0.78	23	22.97	1.00	0.01	1.00

2nd Int'l NephroCAGE Symposium, Aug 16, 2022

Klemens Budde, Charité Berlin

15

Supported by: Federal Ministry for Economic Affairs and Climate Action
on the basis of a decision by the German Bundestag

Predicting Long Term Kidney-Allograft Failure: Machine Learning vs Traditional Statistical Models

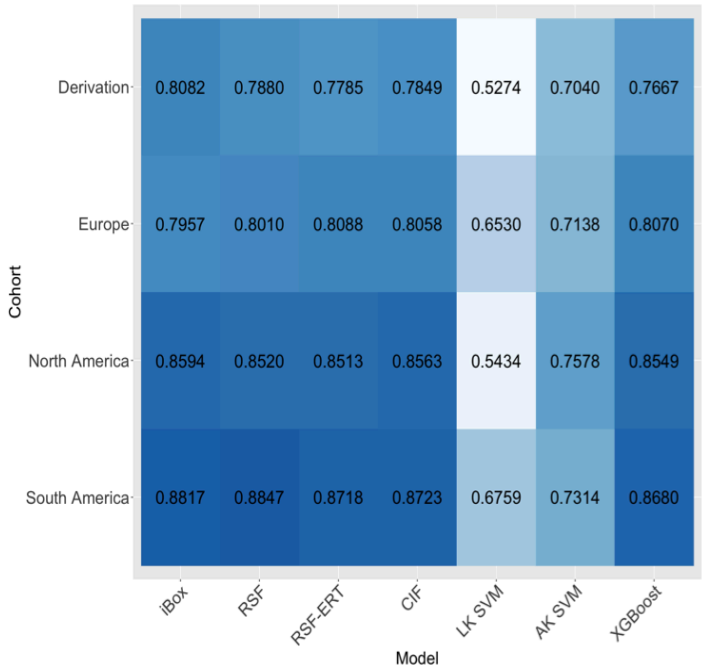
A. Truchot[1], M. Raynaud[1], c. lefaucheur[2], O. Aubert[1], A. Loupy[1]

[1]Paris Transplant Group, Paris, France, [2]Hôpital Saint-Louis, Paris, France

Meeting: 2022 American Transplant Congress

Abstract number: 538

- 4000 patients in derivation cohort and 5054 patients in validation cohort
- 1165 graft losses after 5,2 years
- 24 parameters for 6 ML models for 7 year graft survival vs. iBox
- risk assessment after 0,98 years
- iBox slightly better



2nd Int'l NephroCAGE Symposium, Aug 16, 2022

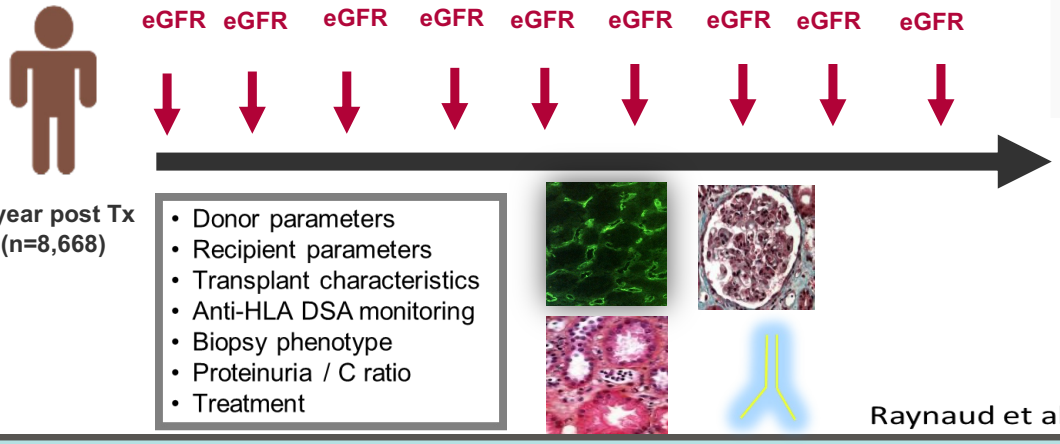
Klemens Budde, Charité Berlin

FAILING ALLOGRAFTS DATA DRIVEN PATTERN RECOGNITION

Cohorts: **8,668** kidney transplant recipients
12 centers in Europe and the USA between 2001 and 2016.

Inclusion criteria:

- Age > 18 years
- Solitary kidney transplants
- At least two eGFR measures after 1-year post-Tx



Raynaud et al, KI 2021

Cohort	Number of patients	Centers involved
Development cohort	4,140	Necker, Saint-Louis, Foch, Toulouse, Lyon and Nantes
European validation cohort	2,688	Leuven, Montpellier and Nancy
North-American validation cohort	1,840	Cedars-Sinai, Los Angeles Johns-Hopkins, Baltimore, Mayo clinic, Rochester

8,668 patients.
109,049 eGFR
Median follow-up : **6-years** post transplantation (IQR 4-9)

2nd Int'l NephroCAGE Symposium, Aug 16, 2022

Klemens Budde, Charité Berlin

Supported by: Federal Ministry for Economic Affairs and Climate Action
on the basis of a decision by the German Bundestag

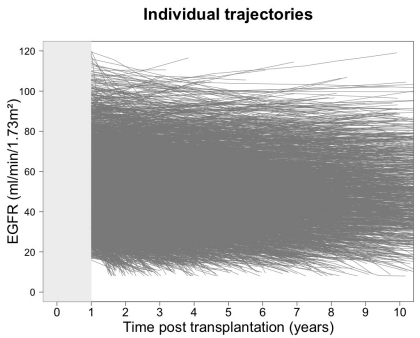
FAILING ALLOGRAFTS DATA DRIVEN PATTERN RECOGNITION

Latent Class Mixed Model (LCMM)

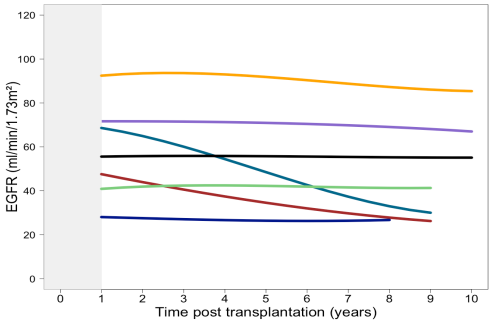
- Unsupervised method adapted to repeated measurements allowing the detection of underlying (latent) groups in longitudinal data.
- Statistical parameters are tested to obtain the best profiles of eGFR trajectories.
- LCMM associates logistic regression and a mixed model.

$$Y_{ij} = X_{Li}(t_{ij})^\top \beta + Z_i(t_{ij})^\top u_i + w_i(t_{ij}) + \epsilon_{ij}$$

Proust-Lima et al. Journal of statistical Software, 2017



10 hours of computation time per model



Patients are given likelihoods of belonging to latent classes

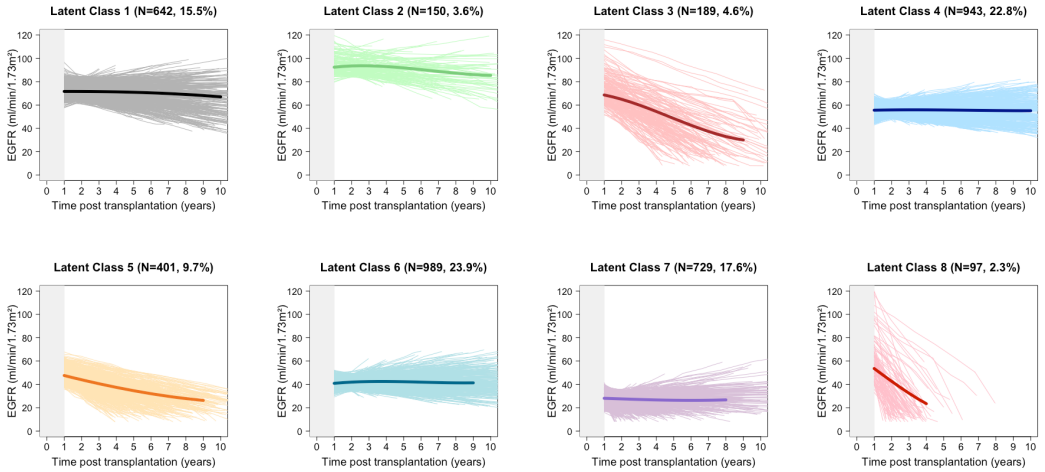
$$\pi_{ig} = P(c_i = g|X_{ci}) = \frac{e^{\xi_{0g} + X_{ci}^\top \xi_{1g}}}{\sum_{l=1}^{G} e^{\xi_{0l} + X_{ci}^\top \xi_{1l}}}$$

2nd Int'l NephroCAGE Symposium, Aug 16, 2022

Klemens Budde, Charité Berlin

Supported by: Federal Ministry for Economic Affairs and Climate Action, on the basis of a decision by the German Bundestag

FAILING ALLOGRAFTS DATA DRIVEN PATTERN RECOGNITION



Raynaud et al, KI 2021

Latent class	eGFR baseline	eGFR slope per year	Functional correspondence
#1	**71.6** (10.4)	**-0.75** (3.10)	**High baseline, stable**
#2	**91.6** (11.4)	**-1.04** (3.37)	**Very high baseline, slightly decreasing**
#3	**70.1** (15.9)	**-8.88** (3.44)	**High baseline, fast declining**
#4	**55.6** (8.11)	**-0.13** (2.51)	**Middle baseline, stable**
#5	**48.2** (10.3)	**-5.38** (2.46)	**Low baseline, decreasing**
#6	**41.0** (7.1)	**0.12** (2.40)	**Low baseline, stable**
#7	**28.4** (6.4)	**-2.97** (6.00)	**Very low baseline, slow decreasing**
#8	**58.0** (18.5)	**-23.9** (8.76)	**Middle baseline, fast declining**

2nd Int'l NephroCAGE Symposium, Aug 16, 2022

Klemens Budde, Charité Berlin

Supported by: Federal Ministry for Economic Affairs and Climate Action, on the basis of a decision by the German Bundestag

Multidimensional Prognostication Tool for Kidney Transplant Patient Survival: The Mortality Mbox

C. Debiais-Deschamps[1], O. Aubert[1], D. Yoo[1], G. Divard[1], C. Lefaucheur[2], C. Legendre[3], A. Loupy[1]

[1]Université de Paris, PARCC, INSERM, Paris, France, [2]Service de Transplantation Rénale, AP-HP, Hôpital Saint-Louis, Paris, France, [3]Service de Transplantation Rénale Adulte, AP-HP, Hopital Necker, Paris, France

Meeting: 2022 American Transplant Congress

Abstract number: 1791

- 1446 patients from France, transplanted 2004-2014
- 309 patients died after 7,6 years
- 160 parameters from Tx to 1 year post Tx
- Multivariable Cox model plus Lasso regression
- 19 predictors selected
- Mortality score AUC 0,81

2nd Int'l NephroCAGE Symposium, Aug 16, 2022

Klemens Budde, Charité Berlin

Supported by: Federal Ministry for Economic Affairs and Climate Action on the basis of a decision by the German Bundestag

Computer vs Human-Based Prediction and Stratification of the Risk of Long-Term Kidney Allograft Failure

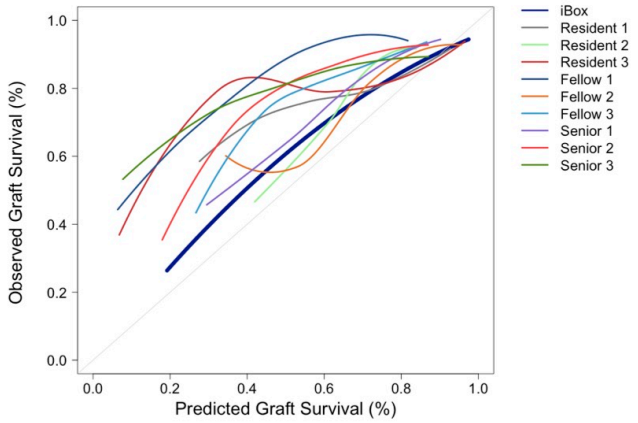
G. Divard[1], M. Raynaud[1], V. Tataputii[2], B. Abdalla[3], C. Legendre[1], C. Lefaucheur[1], O. Aubert[1], A. Loupy[1]

[1]Paris Transplant Group, Paris, France, [2]NYU Langone Health, New York, NY, [3]UCLA, Los Angeles, CA

Meeting: 2022 American Transplant Congress

Abstract number: 298

- 400/4000 random patients, 84 (21%) graft losses after 7,2 years
- 44 parameters during first year
- 9 physicians
- Predicting allograft survival at 1 year
- iBox AUC of 0,79 with calibration error of 5,8%
- Physicians overestimated graft failure with higher calibration error



2nd Int'l NephroCAGE Symposium, Aug 16, 2022

Klemens Budde, Charité Berlin

21

Supported by: Federal Ministry for Economic Affairs and Climate Action, on the basis of a decision by the German Bundestag

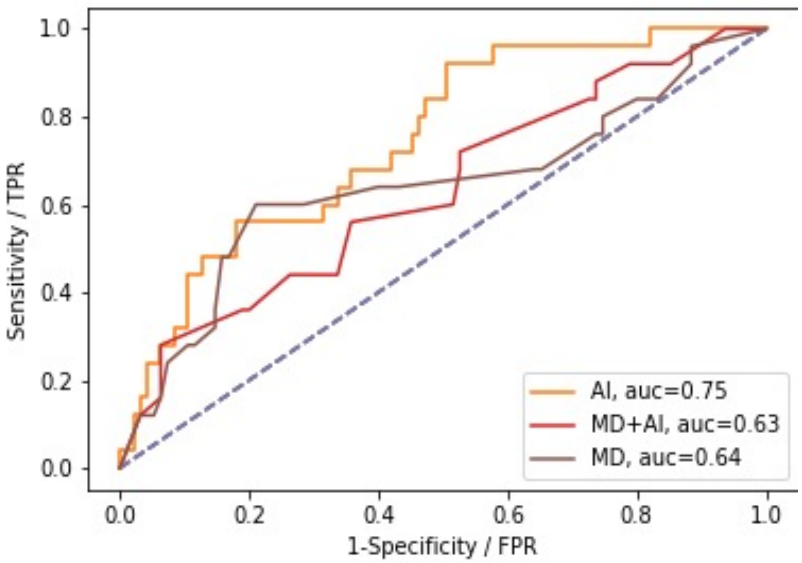
Kidney Pilot

Study results:

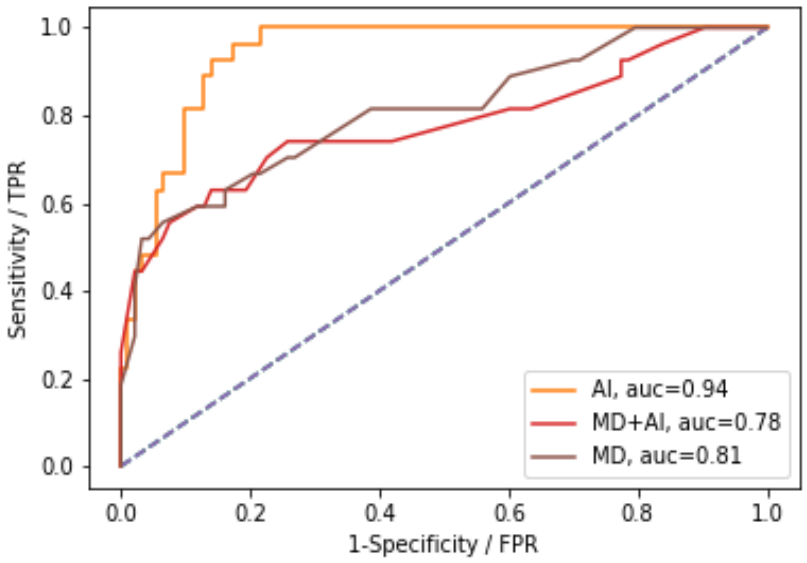
AI *versus* Medical Doctor (MD) *versus* AI+MD

AI outperforms MDs

Rejection: ROC Curve (all)



Graft Loss: ROC Curve (all)



2nd Int'l NephroCAGE Symposium, Aug 16, 2022

Klemens Budde, Charité Berlin

Supported by: Federal Ministry for Economic Affairs and Climate Action on the basis of a decision by the German Bundestag

Summary:

"Prediction is very difficult,
especially if it's about the future!"

Niels Bohr

2nd Int'l NephroCAGE Symposium, Aug 16, 2022

Klemens Budde,
Charité Berlin

Supported by: Federal Ministry for Economic Affairs and Climate Action
on the basis of a decision by the German Bundestag

Thank you for your attention!

2nd Int'l NephroCAGE Symposium, Aug 16, 2022

Klemens Budde,
Charité Berlin

Supported by:
Federal Ministry for Economic Affairs and Climate Action
on the basis of a decision by the German Bundestag