

Systems biology derived biomarkers to predict progression of renal function decline in type 2 diabetes mellitus

Gert Mayer, MD^{1*}, Hidde J.L. Heerspink, PhD^{2*}, Constantin Aschauer, MD³, Andreas Heinzl, MS⁴, Georg Heinze, PhD⁵, Alexander Kainz, PhD³, Judith Sunzenaue, MSr³, Paul Perco, PhD⁴, Dick de Zeeuw, MD², Peter Rossing, MD⁶, Michelle Pena, PhD² and Rainer Oberbauer, MD³ for the SYSKID consortium

¹ Department of Internal Medicine IV (Nephrology and Hypertension), Medical University of Innsbruck, Innsbruck, Austria

² Department of Clinical Pharmacy and Pharmacology, University Medical Center Groningen, University of Groningen, Groningen, Netherlands

³ Department of Nephrology, Medical University of Vienna, Vienna, Austria

⁴ Emergentec biodevelopment GmbH, Vienna, Austria

⁵ Center for Medical Statistics, Informatics and Intelligent Systems (CeMSIIS), Section for Clinical Biometrics, Medical University of Vienna, Vienna, Austria

⁶ Steno Diabetes Center, Gentofte, Denmark, University of Copenhagen, Copenhagen Denmark

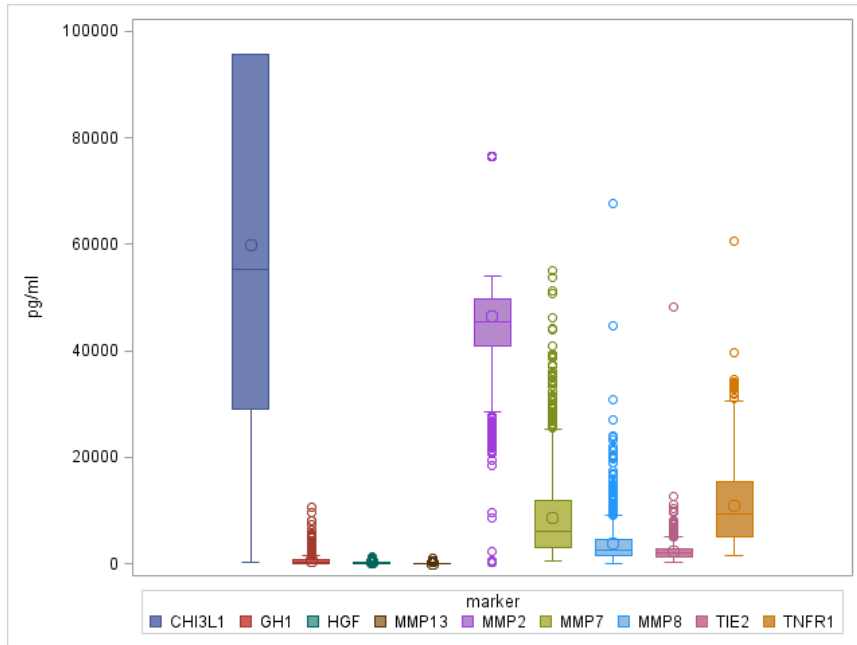
*equal contribution

SUPPLEMENTARY DATA

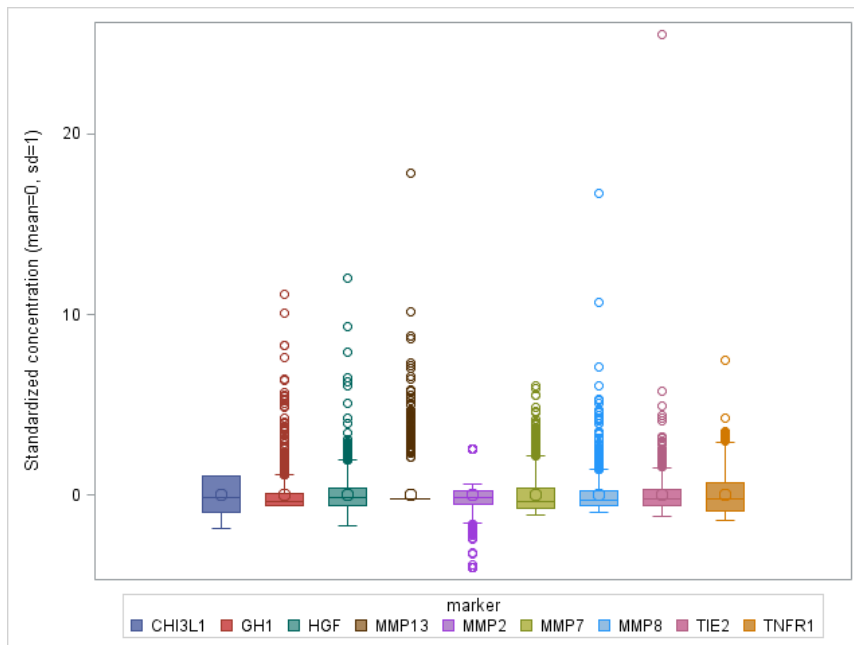
Supplementary Figure 1

Concentration of markers in baseline serum obtained at study initiation. (A) measured concentration; (B) concentrations standardized to mean=0 and standard deviation=1

A



B

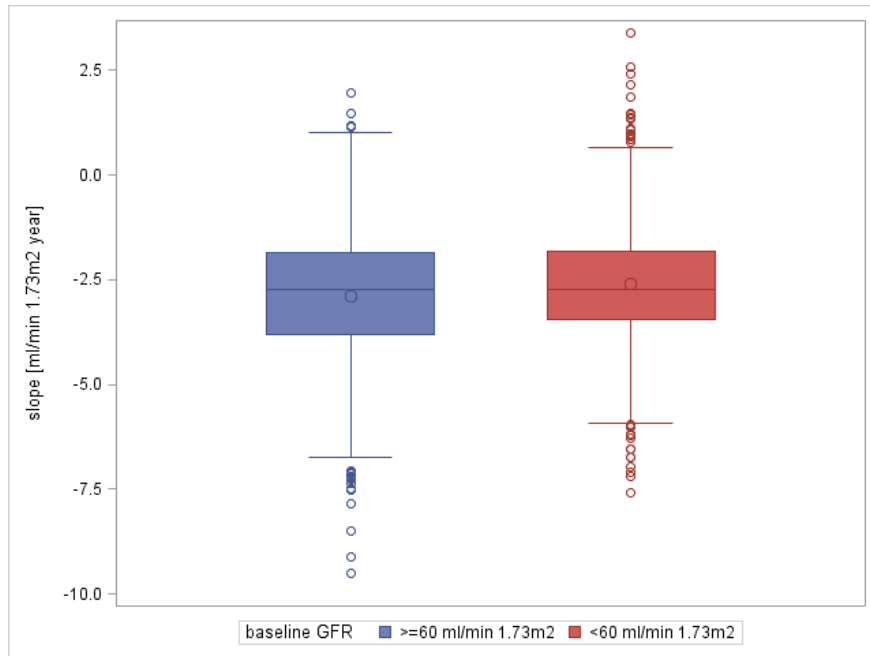


SUPPLEMENTARY DATA

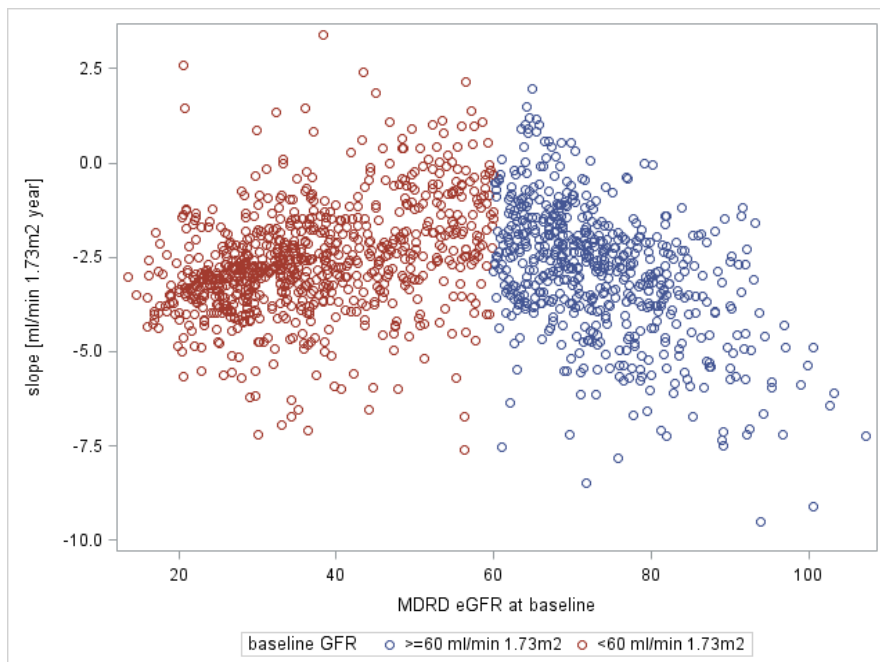
Supplementary Figure 2

(A) Boxplot and (B) scatterplot of calculated slope of eGFR for each patient by study group

A)



B)



SUPPLEMENTARY DATA

Supplementary Table 1

Systems biology derived biomarker panel

protein	synonyms	name	UniProt ID
CHI3L1	GP-39, CGP-39, ASRT7, YKL-40, YYL-40, hCGP-39, HC-gp39, HCGP-3P, GP39, YKL40	Chitinase 3 – like 1	P36222
GH1	IGHD1B, GH-N, GH, GHB5, hGH-N, GHN	growth hormone 1	P01241
HGF	F-TCF, HPTA, SF, DFNB39, HGFB	hepatocyte growth factor (hepapoietin A; scatter factor)	P14210
MMP13	CLG3, MANDP1, MMP-13	matrix metalloproteinase 13	P45452
MMP2	TBE-1, MMP-II, MONA, MMP-2, CLG4A, CLG4	matrix metalloproteinase 2	P08253
MMP7	MPSL1, MMP-7, PUMP-1, PUMP1	matrix metalloproteinase 7	P09237
MMP8	HNC, CLG1, PMNL-CL, MMP-8	matrix metalloproteinase 8	P22894
TIE2	CD202b, CD202B, TIE-2, TEK, VMCM1, VMCM	TEK tyrosine kinase, endothelial	Q02763
TNFR1	TNFRSF1A, TNFR60, TNF-R55, TNFAR, TNFR1-d2, TBP1, TNFR55, TNF-R-I, MS5, p60, CD120a, FPF, p55-R, TNF-R, p55	tumor necrosis factor receptor superfamily member 1A	P19438
CXCL1	GROa, GROA, MGSA-a, GRO, FSP, MGSA, SCYB1, GRO1, NAP-3	chemokine (C-X-C motif) ligand 1 (melanoma growth stimulating activity, alpha)	P09341
CXCL10	crg-2, INP10, gIP-10, mob-1, C7, IP-10, SCYB10, IFI10	chemokine (C-X-C motif) ligand 10	P02778
LEP	LEPD, OBS, OB	leptin	P41159
CCL2	SMC-CF, HSMCR30, MCAF, GDCE-2, HC11, SCYA2, MCP1, MGC9434, MCP-1	chemokine (C-C motif) ligand 2	P13500
VEGFA	VEGF, MVCD1, VPF, VEGF-A	vascular endothelial growth factor A	P15692
EGF	HOMG4, URG	epidermal growth factor	P01133
IL1B	IL-1, IL1F2, IL1-BETA, IL-1B	interleukin 1 beta	P01584

SUPPLEMENTARY DATA

Supplementary Table 2

Multivariable model for loss of eGFR resulting from LASSO selection and 1000 bootstrap resamples based on data of the discovery cohort (study of Pena et al). The nine most often selected markers in the bootstrap, highlighted in bold, were selected for validation in the present study. Parameters are ranked based on their selection probability in the bootstrapped LASSO selection in descending order. In the original sample and in each bootstrap resample the tuning parameter of the LASSO was optimized by minimizing AIC.

beta, LASSO-estimated regression coefficients; mean, mean coefficients over 1000 bootstrap resamples; 2.5%, 97.5%, 2.5th and 97.5th percentiles of distribution of bootstrapped coefficients; p-value computed by 'twice smaller tail' rule (twice the minimum of proportion of bootstrapped coefficients lower than or equal to 0 and proportion of bootstrapped coefficients greater than or equal to 0); selection probability, proportion of bootstrap resamples where the variable had a non-zero regression coefficient.

	beta	mean	2.5%	97.5%	p-value	selection probability
(Intercept)	-90.9014	-104.8530	-218.3065	7.2328	0.064	1.000
log baseline UACR	-0.5863	-0.6123	-0.9604	-0.2866	0.000	1.000
SBP	0.0456	0.0484	0.0049	0.0921	0.036	0.982
MMP2	7.0473	8.0619	0.6148	15.4056	0.048	0.977
Tie2	-0.7675	-0.6785	-1.3786	0.0000	0.056	0.973
MMP7	-0.4476	-0.4227	-0.9421	0.0000	0.092	0.959
DM drugs*	-1.0285	-0.9866	-2.2774	0.0000	0.218	0.899
current or former smoker	-1.9337	-1.5535	-4.0655	0.0000	0.230	0.896
baseline eGFR	-0.0212	-0.0288	-0.0842	0.0000	0.392	0.819
TNFR1	-0.8351	-0.9398	-2.7729	0.0000	0.410	0.802
MMP8	0.1964	0.3366	0.0000	0.9297	0.528	0.748
MMP13	-0.1372	-0.4404	-2.1688	1.0669	0.814	0.732
HGF	-0.1342	-0.3474	-1.1344	0.0000	0.584	0.726
GH1	-0.0543	-0.1088	-0.4969	0.2074	0.816	0.702
female gender	0.6830	0.5008	-1.3835	2.1213	0.764	0.686
CHI3L1	-0.0328	-0.1318	-0.8145	0.5700	0.948	0.647
CXCL10	0.0000	0.1786	-0.5812	1.0047	1.000	0.640
CXCL1	0.0000	-0.1521	-2.0604	1.5487	1.000	0.611
CXCL2	0.0000	0.1390	-1.4370	1.9030	1.000	0.586
DBP	0.0000	-0.0059	-0.0600	0.0497	1.000	0.580
Lep	0.0000	0.1512	-0.2928	0.8702	1.000	0.557

* DM drugs ...number of glucose lowering drugs (see OA_Table 5, in the DIRECT study, 17% received insulin, 62% oral hypoglycemic agents only and 19% both, insulin and oral hypoglycaemic agents. In SMACRO the specific antidiabetic agents were not recorded.)

SUPPLEMENTARY DATA

Supplementary Table 3

Linear mixed model for patients with eGFR ≥ 60 ml/min/1.73m² at baseline relating demographic and clinical variables and biomarkers to eGFR values at baseline and eGFR slope per year

Variable	baseline	95%CI		p-value	slope	95%CI		p-value
constant	189.34				1.5722	-3.2473	6.3917	0.5219
female gender	-6.1889	-7.8095	-4.5683	<.0001	-0.3543	-0.7563	0.0476	0.0840
age (per year)	-0.2758	-0.3738	-0.1778	<.0001	-0.2758	-0.3738	-0.1778	0.3621
current or former smoker	-0.3893	-2.0667	1.2881	0.6490	0.3130	-0.1045	0.7306	0.1416
total cholesterol (mmol/l)	-0.02393	-0.6197	0.5719	0.9372	0.1286	-0.02665	0.2838	0.1044
body mass index (kg/m ²)	0.1791	0.009826	0.3484	0.0381	-0.03714	-0.07862	0.004335	0.0792
log2 UACR (mg/g)	-0.7522	-1.3794	-0.1249	0.0188	-0.1846	-0.3638	-0.00536	0.0435
mean arterial pressure (mmHg)	-0.03035	-0.1155	0.05476	0.4844	0.000487	-0.02104	0.02202	0.9646
HbA1c (%9)	0.1571	-0.3072	0.6214	0.5071	-0.1162	-0.2355	0.002958	0.0560
log2 MMP2 (pg/ml)	-2.1428	-3.1891	-1.0966	<.0001	0.2242	-0.01261	0.4610	0.0635
log2 Tie2 (pg/ml)	1.1232	0.3513	1.8951	0.0635	n.s.			
log2 CHI3L1 (pg/ml)	0.5414	-0.08352	1.1663	0.0895	n.s.			
log2 MMP8 (pg/ml)	0.8674	0.1797	1.5551	0.0135	-0.1725	-0.3163	-0.02883	0.0186
log2 MMP7 (pg/ml)	0.3636	-0.6632	1.3905	0.4874	-0.2245	-0.4737	0.02471	0.0774
log2 HGF (pg/ml)	1.5020	0.4134	2.5906	0.0069	n.s.			
log2 TNFRI (pg/ml)	-8.5091	-10.3902	-6.6280	<.0001	n.s.			
log2 GH1 1 (pg/ml)	-0.07095	-0.2070	0.5588	0.3677	-0.07095	-0.1673	0.02544	0.1490

n. s.: not selected (see main text), MMP13 is not listed because of n.s. in both outcomes, baseline and slope

SUPPLEMENTARY DATA

Supplementary Table 4

Linear mixed model for patients with eGFR ≤ 60 ml/min/1.73m² at baseline relating demographic and clinical variables and biomarkers to eGFR values at baseline and eGFR slope per year

Variable	baseline	95%CI		p-value	slope	95%CI		p-value
constant	125.11				16.371 8	4.37	28.37	0.0076
female gender	-2.49	-3.8972	-1.0847	<0.0001	-0.1631	-0.8252	0.4989	0.6289
age (per year)	-0.02485	-0.09956	0.04987	0.5143	0.01165	- 0.02756	0.05086	0.5601
current or former smoker	0.1924	-1.0516	1.4364	0.7617	1.0000	0.3887	1.6114	0.0014
total cholesterol (mmol/l)	0.2155	-0.4187	0.01799	0.3371	-0.2004	-0.4187	0.01799	0.0721
body mass index (kg/m ²)	-0.06550	-0.1629	0.03192	0.1874	0.00242 9	- 0.05085	0.05571	0.9288
log2 UACR (mg/g)	-1.1205	-1.4084	-0.8325	<.0001	-0.4268	-0.5689	-0.2847	<.0001
mean arterial pressure (mmHg)	0.07138	0.00888 5	0.1339	0.0252	- 0.02697	- 0.05768	0.00374 4	0.0852
HbA1c (%)	-0.1886	-0.5680	0.1909	0.3299	0.06177	-0.1234	0.2470	0.5131
log2 MMP2 (pg/ml)	2.7356	1.4639	4.0074	<.0001	-0.8673	-1.4858	-0.2488	0.0060
log2 Tie2 (pg/ml)	1.2376	0.6417	1.8335	<.0001	-0.3853	-0.6866	- 0.08405	0.0122
log2 MMP7 (pg/ml)	-1.2410	-2.0699	-0.4122	0.0034	-0.3595	-0.8228	0.1038	0.1282
log2 HGF (pg/ml)	2.0009	1.1332	2.8687	<.0001	n.s.			
log2 MMP13 (pg/ml)	0.4765	-1.0701	2.0232	0.5457	-0.6628	-1.4872	0.1617	0.1151
log2 TNFRI (pg/ml)	-9.9139	-11.2758	-8.5520	<.0001	0.9630	0.2966	1.6294	0.0046

n. s.: not selected (see main text), MMP8, GH1 and CHI3L1 are not listed because of n.s. in both outcomes, baseline and slope

SUPPLEMENTARY DATA

Supplementary Table 5

Baseline characteristics of participants in the two studies analyzed (DIRECT and SMACRO) Values are mean (standard deviation) or N (%).

	DIRECT		SUN-MACRO	
	Placebo arm	Active treatment arm	Placebo arm	Active treatment arm
age (years)	57.8 (7.6)	57.4 (7.4)	63.8 (9.3)	62.8 (9.3)
gender (female)	214 (52.7)	200 (55.0)	108 (25.6)	88 (21.6)
smoking status at baseline (previous or current)	102 (25.1)	83 (22.8)	193 (52.9)	176 (50.9)
BMI (kg/m ²)	29.5 (4.7)	29.5 (4.5)	32.7 (13.7)	33.5 (19.6)
systolic blood pressure baseline (mmHg)	134 (13)	134 (14)	139 (14)	139 (14)
diastolic blood pressure baseline (mmHg)	78 (7)	78 (7)	73 (10)	74 (10)
heart rate	--	--	69 (11)	70 (12)
HbA1c (%)	8.2 (1.5)	8.2 (1.6)	7.9 (1.5)	8.0 (1.6)
serum glucose (mmol/l)	--	--	8.6 (3.5)	8.8 (4.0)
serum cholesterol (mmol/l)	5.5 (1.2)	5.4 (1.1)	4.7 (1.4)	4.5 (1.0)
serum creatinine (μmol/l)	62 (15)	63 (15)	190 (44)	194 (47)
baseline MDRD eGFR (ml/min/1.73m ²)	70.0 (12.0)	68.8 (10.9)	33.6 (9.5)	33.3 (9.6)
baseline UACR (mg/g)	--	--	1809 (1682)	1824 (1538)