A specific urinary peptidomic profile predicts outcome in SARS-CoV-2 - infected patients

Short title: UPP predicts COVID-19 outcome

Ralph Wendt, Lutgarde Thijs, Sven Kalbitz, Harald Mischak, Justyna Siwy, Julia Raad, Jochen Metzger, Barbara Neuhaus, Heiko von der Leyen, Emmanuel Dudoignon Alexandre Mebazaa,Goce Spasovski, Mimoza Milenkova, Aleksandra Canevska-Tanevska, Beata Czerwieńska, Andrzej Wiecek, Björn Peters, Åsa Nilsson, Matthias Schwab, Katja Rothfuss, Christoph Lübbert, Jan A. Staessen, Joachim Beige, the CRIT-COV-U investigators*

Word counts: manuscript 6466, article text 3170, summary 343

Number: 4 tables, 1 figure; references 36

Correspondence
Joachim Beige,
Interdisciplinary Translational Research Unit (ITF)
Department of Infectious Diseases/Tropical Medicine,
Nephrology/KfH Renal Unit and Rheumatology,
Hospital St. Georg,
Martin-Luther-University Halle/Wittenberg,
Delitzscher Strasse 141,
DE-04129, Leipzig, Germany

Telephone: +49-341-909-4898 (office)
+49-151-11307581(mobile)
Facsimile: +49-341-909-1493 (e-office)
email: joachim.beige@sanktgeorg.de
joachim.beige@kfh.de

Affiliations

Department of Infectious Diseases/Tropical Medicine, Nephrology/KfH Renal Unit and Rheumatology, St. Georg Hospital Leipzig, Germany (R Wendt MD, S Kalbitz MD, Prof C Lübbert MD, DTM&H, Prof J Beige MD)**, Research Unit Hypertension and Cardiovascular Epidemiology, KU Leuven Department of Cardiovascular Diseases, University of Leuven, Belgium** (L Thijs)**, Institute of Cardiovascular and Medical Sciences, Glasgow, United Kingdom** (Prof H Mischak PhD)**, Mosaiques-Diagnostics GmbH, Hannover, Germany** (Prof H Mischak PhD, J Siwy PhD, J Raad, PhD, J Metzger PhD)**, Hannover Clinical Trial Center, Medizinische Hochschule, Hannover, Germany** (B Neuhaus PhD, Prof H von der Leyen MD), Université de Paris, **Department of Anaesthesiology and Intensive Care, Hôpital Saint Louis-Lariboisière, U942 Inserm MASCOT, Paris, France** (Prof A Mebazaa MD, E Dudoignon MD)**, Department Nephrology, Cyril and Methodius University, Skopje, Republic of North Macedonia** (Prof G Spasovski MD, M Milenkova MD, A Canevska-Tanevska MD), **Department of Nephrology, Transplantation and Internal Medicine, Medical University of Silesia, Katowice, Poland** (B Czerwieńska MD, Prof A Wiecek MD)**, Department of Nephrology, Skaraborg Hospital, Skövde, Sweden** (B Peters MD)**, Department of Molecular and Clinical Medicine, Institute of Medicine, the Sahlgrenska Academy at University of Gothenburg, Gothenburg, Sweden** (B Peters MD)**, Research and Development Centre (FoU), Skaraborg Hospital, Skövde, Sweden** (Å Nilsson, MSc)**, Margarete-Fischer-Bosch Institute for Clinical Pharmacology and Department for Gastroenterology, Hepatology and Endocrinology, Robert Bosch Hospital, Stuttgart, Germany** (Prof M Schwab MD, K Rothfuss MD)**, Departments of Clinical Pharmacology, and of Biochemistry and Pharmacy, University of Tuebingen, Germany** (Prof M Schwab MD), **Division of Infectious Diseases and Tropical Medicine, Leipzig University Medical Center, Leipzig, Germany** (Prof C Lübbert MD, DTM&H)**, Non-Profit Research Institute Alliance for the Promotion of Preventive Medicine, Mechelen** (Prof JA Staessen MD)**, Belgium, Biomedical Sciences Group, Faculty of Medicine, University of Leuven, Leuven, Belgium** (Prof JA Staessen MD), **Martin-Luther-University Halle-Wittenberg, Halle an der Saale, Germany** (Prof J Beige MD).

* CRIT-COV-U investigators and key investigators are listed in the appendix (pp 06-07).

Correspondence to Prof Joachim Beige, Martin-Luther-University Halle/Wittenberg, Delitzscher Strasse 141, DE-04129, Leipzig, Germany.

E-mail: joachim.beige@sanktgeorg.de or joachim.beige@kfh.de

Summary

Background

Prediction models, commonly based on clinical characteristics, routine biochemistry and imaging features, were developed for coronavirus disease 2019 (COVID-19), but to our knowledge, none includes proteomic markers reflecting the molecular pathophysiology of disease progression. (words 34)

Methods The Prospective Validation of a Proteomic Urine Test for Early and Accurate Prognosis of Critical Course Complications in Patients with SARS-CoV-2 Infection Study (Crit-COV-U) is recruiting consecutive patients (≥18 years) with PCR-confirmed SARS-CoV-2 infection at six European sites. A urinary proteomic biomarker (COV50) developed by CE-MS technology, which consisted of 50 sequenced peptides and identified the parental proteins was evaluated in 228 patients (derivation cohort) with replication in 99 patients (validation cohort). Death and progression along the WHO scale were assessed up to 21 days from the initial PCR test. Statistical methods included logistic regression, receiver operating curve (ROC) analysis with comparison of the area under curve (AUC) between nested models. (words 111/145)

Findings In the derivation cohort, 23 patients died and 48 developed worse WHO scores. Odds ratios (OR) for death per 1-SD increment in COV50 were 3·52 (95% CI, 2·02–6·13, p<0·0001) unadjusted and 2·73 (1·25–5·95, p=0·012) adjusted for sex, age, baseline WHO score, body mass index and comorbidities. For progression along the WHO scale, the corresponding OR were 2·63 (1·80–3·85, p<0·0001) and 3·38 (1·85–6·17, p<0·0001), respectively. The AUC for COV50 as continuously distributed variable was 0·80 (0·72–0·88) for mortality and 0·74 (0·66–0·81) for worsening WHO score. The optimised COV50

thresholds for mortality and worsening WHO score were 0·47 and 0·04, resulting in sensitivity/specificity of 87·0/74·6% and 77·1/63·9%, respectively. On top of sex, age, body mass index, comorbidities, COV50 analysed as continuously distributed variable and per threshold improved the AUC, albeit borderline for death, that is from 0·78 to 0·82 (p=0·11) and 0·84 (p=0·052) for mortality and from 0·68 to 0·78 (p=0·0097) and 0·75 (p=0·021) for worsening WHO score. Findings in the validation cohort were confirmatory. (words 169/314)

Interpretation This first CRIT-COV-U report proves the concept that urinary proteomic profiling generates biomarkers indicative of adverse COVID-19 outcomes, even at an early stage of disease, including WHO stage 1-3. (words 29/343)

Funding German Federal Ministry of Health acting upon a decree from the German Federal Parliament.

Key words COVID-19, disease severity, risk score, SARS-CoV-2, urinary proteomics, peptidomics, World Health Organization COVID-19 severity scale, polymeric immunoglobulin receptor, CD99, collagen homeostasis

Introduction (words 311/654)

Since the outbreak of the COVID-19 pandemic in Wuhan, People´s Republic of China, an exponentially growing literature described the clinical characteristics of infected patients at high risk of severe disease and death.[1-3] The burden on health care led to the development of models predicting progression to adverse outcomes with a principal objective to support clinical decision making. A systematic review of the literature published in April 2020 and updated thereafter summarised over 50 prognostic models commonly including sex, age, comorbidities, C-reactive protein, lymphocyte count, body temperature, serum creatinine and imaging features, but qualified all models as vulnerable to bias and not clinically applicable.[4] More recently published models, including the 4C score, were properly calibrated and gained in accuracy,[5-7] but none considered the molecular pathophysiology of the progression from silent infection to critical disease.

SARS-CoV-2 preferentially infects the cells of the respiratory tract, but penetrates the heart, liver, brain, kidneys and blood vessels.[8] The infection is therefore a systemic disease, leading to potential multiorgan failure.[9] Urine contains an array of over 20,000 endogenous peptides, which are partly generated along the nephron or from the circulation passing through the glomerular barrier. The urinary peptidome profile (UPP) therefore provides a systemwide molecular signature of progressing SARS-CoV-2 infection. Sequencing of urinary peptides allows identification of the parental proteins.[10,11] Multidimensional urinary peptide profiles already provide a specific molecular signature in the preclinical phase of heart failure,[12] chronic kidney disease (CKD)[13] or diabetic nephropathy.[14] A proof-of-concept study suggested the feasibility that UPP at the initial SARS-CoV-2 infection stage may predict outcome.[15] The Prospective Validation of a Proteomic Urine Test for Early and Accurate Prognosis of Critical Course Complications in Patients with SARS-CoV-2 Infection

Study (CRIT-COV-U) was therefore designed to develop and validate a UPP biomarker for prediction of outcome of SARS-CoV-2 – infected patients.[5-7]

Methods (words 810/1464)

The CRIT-COV-U project complies with the Helsinki declaration.[16] The Ethics Committee of the German-Saxonian Board of Physicians, Dresden, Germany (number, EK-BR-88/20.1) and the Institutional Review Boards of the recruiting sites provided ethical clearance. The protocol was registered at the German Register for Clinical Studies (www.drks.de), number DRKS00022495, which is interconnected with the WHO International Clinical Trial Registry Platform (www.who.int/clinical-trials-registry-platform).

CRIT-COV-U is a prospective multicentre cohort study with the objective to identify UPP biomarkers predictive of the clinical course in adults with PCR-confirmed SARS-CoV-2 infection.[17] To be eligible, patients had to be ≥18 years, not to be anuric, and to be able to give informed written consent. Six European study sites, located in Germany (n=2), France, the Republic of North Macedonia, Poland and Sweden, enrolled consecutive patients. Patients were diagnosed, while in ambulatory or first-day hospital care, and followed up for at least 21 days or until hospital discharge or death, whichever event occurred first. At each of three timepoints (day 0-2; 4-7; 10-21 after baseline; following the PCR-based diagnosis), patients were staged according to the WHO criteria in eight categories:[18] (1) ambulatory without limitation of activity; (2) ambulatory with limited activity; (3) hospitalised without oxygen therapy; (4) hospitalised on oxygen therapy by mask or nasal prongs; (5) hospitalised receiving non-invasive ventilation or high-flow oxygen therapy; (6) hospitalised with intubation and mechanical ventilation; (7) hospitalised with mechanical ventilation and additional organ support, such as vasopressors, renal replacement therapy or extracorporeal membrane oxygenation; and (8) death. The information collected via electronic case-report

forms (MARVIN EDC, XClinical GmbH, Munich, Germany) included clinical characteristics, such as ethnicity, sex, age, body mass index, blood pressure and routine biochemical measurements, such as glomerular filtration derived from serum creatinine.[19]

For UPP, 8 ml-urine samples were collected in borated test tubes (ExactoBac-U®, Sarstedt, Nümbrecht, Germany) at timepoints of clinical staging (day 0-2, 4-7, and 10-21). The samples were kept at -20 °C until assayed. The methods for the capillary electrophoresis coupled with mass spectrometry, peptide sequencing, and for the evaluation, calibration and quality control of the mass spectrometric data have been published[11,20,21] and are outlined in the appendix (pp 2-5). For identification of the urinary biomarker, 186 urine samples were randomly selected from those available in the derivation dataset at timepoints 2 and 3, excluding samples from patients at COVID-19 stages 4 and 5 according to the WHO classification, allowing contrasting the UPP profiles at stages 1-3 (n=116) and stages 6-8 (n=88). After transformation of the mass spectrometric spectra, the levels of peptides with known amino-acid sequence were compared between patients with mild and critical disease, using Wilcoxon rank sum test with adjustment of the significance for multiple comparisons by the Benjamini and Hochberg method.[22] The disease-specific classifier was developed using support vector machine modelling and cross-validated by a take-one-out procedure with significance adjusted for the false-discovery rate set at 0.05.

Sample size calculations informed by the proof-of-concept study,[15] proposed a derivation phase sample size of 212 patients with critical COVID-19 (WHO stage ≥6) to be contrasted with 271 patients with mild symptoms to identify an UPP, yielding a sensitivity and specificity of 75% and 80%, respectively. Following a request from the German regulators faced with the extraordinary burden placed on health care, the CRIT-COV-U database was frozen on 17 December 2020 for an interim analysis with 228 and 99 patients enrolled in the derivation and validation cohorts, respectively.

For database management and statistical analysis, SPSS (version 22·0) and SAS (version 9·4) software were used. Significance was a two-tailed significance of 0·05 or less. Means and proportions were compared using the large-sample z test or ANOVA and Fisher's exact test, respectively. The predefined endpoints were mortality and progression across the WHO scale of COVID19 severity. In the derivation dataset, the incidence of endpoints was related to the proteomic classifier using single and multiple logistic regression analysis taking into account previously reported risk factors, such as sex, age, the WHO scale at timepoint 1 and comorbidities including hypertension, diabetes, heart failure and cancer. Performance of COV50 in risk stratification was assessed by the area (AUC) under the receiver operating curve (ROC) and the Delong approach to compare the AUCs between nested models. Internal validation was performed through the calculation of the leave-one-out cross-validated AUC. Prior to computing sensitivity and specificity, the COV50 thresholds were optimised, using Youden index. The AUC, sensitivity and specificity in the validation dataset were calculated from the logistic model and the thresholds derived in the derivation cohort. For further external validation, the distribution of the COV50 classifier was evaluated in 981 controls, randomly selected from the human CE-MS proteome database available at Mosaiques-Diagnostics, Hannover, Germany. Controls were sampled before end of 2019, therefore free of COVID-19 and matched for sex, age (±5 years) and body mass index (±1 kg/m^2) in a 3:1 proportion with the 327 patients enrolled in the current analysis. Finally, we compared the performance of COV50 with the Coronavirus Clinical Characterisation Consortium score (4C)[6] to predict mortality.

Role of the funding source

The funder of the study had no role in the study design, data collection, data analysis, data interpretation, or the writing of the report. All authors had full access to all of the data in the study and had the final responsibility for the decision to submit the manuscript.

Results (words 715/2178)

Comparing the UPP at stages 1-3 with stages 6-8 of COVID disease at timepoints 2 and 3 identified 1132 significantly deregulated peptides. To generate the COV50 classifier, 100 peptides in the top tail of the significance distribution were combined by support vector modelling and reduced to 50 by applying take-one-out cross-validation. The 50 sequenced peptides making up the UPP biomarker and the parental proteins from which the peptide fragments were derived are listed in the appendix (pp 8-10). Using the urine samples collected at timepoint 1, the association of the severity of infection during follow-up was prospectively studied in relation to the 50-peptide urinary biomarker (COV50) and potential confounders, first in the 228 patients included in the derivation dataset and next in 99 patients enrolled in the validation dataset. The 228 participants enrolled in the derivation cohort (table 1) were on average 63·1 years old, included 94 (41·2%) women, 152 (66·7%) patients with comorbidities (appendix figure 1, p 12), including hypertension (n= 137), heart failure (n=30), diabetes mellitus (n=65), or cancer (n=13), and 119 (52·2%) patients on treatment with inhibitors of the renin-angiotensin system, either angiotensin-converting enzyme inhibitors (n=67) or angiotensin-receptor blockers (n=58). The WHO score at enrollment was 1-3 in 90 (39·5%) patients, 4-5 in 107 (46·9%), and 6 in 31 (13·6%) participants.

Across increasing fourths of the COV50 distribution (table 2), the proportion of women decreased, and age and the prevalence of hypertension, heart failure, and diabetes

increased. During follow-up, 23 patients died and 48 progressed across WHO scores. The baseline COV50 distribution shifted upward (p<0·0001), when plotted against the worst WHO score attained during follow-up (appendix figure 2, p 13). For death (table 3), the relative risk expressed per 1-SD increment in COV50 was 3·52 (95% CI, 2·02–6·13, p<0·0001) unadjusted and 2·73 (1·25–5·95, p=0·012) when fully adjusted for sex, age, the baseline WHO score, body mass index and the presence of comorbidities. For progression in WHO score (table 3), the corresponding odds ratios (OR) were 2·63 (1·80–3·85, p<0·0001) and 3·38 (1·85–6·17, p<0·0001), respectively. The cross-validated AUC for COV50 analysed as a continuously distributed variable was 0·80 (95% confidence interval, 0·72–0·88) for total mortality and 0·74 (0·66–0·81) for progressing WHO score (table 4). The optimised COV50 thresholds for total mortality and progressing WHO score were 0·47 and 0·04 and resulted in estimates of sensitivity/specificity of 87·0/74·6% and 77·1/63·9%, respectively (table 4).

The cross-validated AUCs of baseline risk factors were 0·57 (0·46-0·68) for age, 0·65 (0·54-0·75) for the WHO score, and 0·80 (0·72-0·82) for COV50 in relation to mortality (appendix table 2, p 11), and 0·59 (0·51-0·68) for age, 0·52 (0·43-0·61) for the WHO score, and 0·74 (0·66-0·81) for COV50, respectively, in relation to worsening WHO score (appendix p 11). In the derivation cohort, on top of sex, age, body mass index, comorbidities, and the baseline WHO score, COV50 analysed as continuously distributed variable and per threshold (figure 1) slightly enlarged or significantly improved the AUC. For mortality in relation to the continuously distributed COV50 marker and the COV50 threshold, the AUC increased from 0·78 to 0·82 (p=0·11) and 0·84 (p=0·052), respectively; for worsening WHO score, the AUC increased from 0·68 to 0·78 (p=0·0097) and 0·75 (p=0·021).

Compared to the derivation cohort (table 1), the baseline characteristics of the validation cohort, including the distribution of COV50 and comorbidities (appendix figure 1 and 3, pp 12 and 14) were broadly similar. Using the predicted probabilities and the optimised thresholds

derived in the derivation cohort, the results in the validation cohort confirmed the discriminatory performance of the COV50 biomarker, irrespective of whether it was analysed as a continuously distributed variable or as categorised risk factor (figure 1). Compared with the 327 patients included in the current analyses, the 981 matched COVID-19 - free controls had comparable characteristics (appendix p 15; 0.084≤p≤0.87). When applying 0·47 and 0·04 as COV50 thresholds, only two and seven controls scored positive, yielding specificities of 99·8% and 99·3%, respectively. Finally, the 4C mortality score consisting of eight variables to grade was applicable only in 257 hospitalized CRIT-CoV-U patients without missing data, of whom 31 died. In these 257 patients, a 4C score of ≥15, indicating critical disease, and COV50 as a stand-alone biomarker had a similar AUC in relation to mortality (0·77 versus 0·76; p=0·79; appendix figure 5, p 16).

Discussion (words 994/3170)

COV50 is a novel multidimensional urinary biomarker (appendix table 1, pp 8-10), consisting of 50 deregulated urinary peptides mainly derived from collagen alpha 1(1), but also from other proteins previously recognised to be involved in the pathogenesis of COVID-19. On its own and adjusted for clinical risk factors, COV50 predicted the incidence of death and progression across WHO stages. This association was robust and withstood internal validation in the derivation cohort by the leave-one-out AUC approach and by correction for overfitting. External validation in the validation cohort produced confirmatory results. Moreover, on top of established clinical risk factors commonly used in predictive models, COV50 analysed as continuously distributed variable and per threshold (figure 1) improved the AUC, albeit the data were stronger for worsening WHO score than for mortality, given the number of study endpoints.

COV50 is registered in Germany and therefore immediately applicable for clinical and research purposes. The UPP does not undergo significant changes when urine is stored for 5 days at room temperature in borated test tubes,[23,24] thereby providing a wide time window for handing urine samples, for instance as collected at the homes of patients with PCR-confirmed SARS-CoV-2 infection. Furthermore, urine can be stored for years at -20 °C without UPP alteration opening opportunities for research.[25] From a clinical perspective, COV50 might contribute to the personalised management of COVID-19 patients, which can range from observation and follow-up at home, to non-invasive and invasive hospital care, such as treatment with Remdesivir, corticosteroids, monoclonal antibodies, convalescent plasma, intensified oxygen delivery, or mechanical ventilation combined or not with other life supporting interventions. Patient with a COV50 level of less than -1 can be managed at home; those with a level ranging from -1 to 0.40 might require in-hospital management with intermediate care, such as intensified oxygenation; and those with a level of ≥0·40 are likely to require intensive care and invasive life-support measures. The discriminatory performance of COV50 as stand-alone test is comparable with the 4C score, but has the advantage not to include any clinical or biochemical variable, which is already indicative of evolving respiratory insufficiency. From this perspective, UPP followed by the identification of the parental proteins by sequencing the urinary peptides is a powerful instrument in generating multidimensional biomarkers, which reflect the molecular processes underlying various illnesses. Disease-specific peptidomic signatures have become evident in the subclinical run-on to critical illness, as demonstrated for diastolic left ventricular dysfunction (HF1)[12] and CKD or diabetic nephropathy (CKD273).[13,14] The number of peptide fragments making up HF1 is 85 and 273 for CKD273. These UPP are mutually exclusive, highlighting their specificity for the target disease. COV50 shares 13 urinary peptides with CKD273 and only one with HF1. Only two fragments are common to COV50, HF1 and CKD273 (appendix

figure 4, p 15). Along similar lines, COV50 levels 0·47 and 0·04 scored seven or fewer of 981 matched controls as at risk for critical COVID-19, thereby confirming the >99·0% specificity of the marker.

The most prominent characteristic of the COV50 signature (appendix table 1, pp 8-10) is the shift in collagen fragments, in particular collagen alpha 1(1). Deregulation of collagen homeostasis is a hallmark of SARS-CoV-2 infection[26] and was also observed in CKD.[13,27] Several studies reported that CKD and biomarkers indicative of renal impairment predict critical COVID-19, while survivors remain at high CKD risk.[28] The COV50 urinary signature showed upregulation of α1-antitrypsin degradation products in line with reports that α1-antitrypsin deficiency is a major risk factor for life-threatening COVID-19.[29] No information in the context of COVID-19 is currently available on CD99, which is involved in cell recruitment, leukocyte trans-endothelial migration and in maintaining the integrity of the endothelial barrier.[30,31] Reduction of CD99 might interfere with appropriate immune responses and indicate endothelial damage. The polymeric immunoglobulin receptor (pIgR), highly expressed in trachea and the lung and responsible for the transcytosis especially of IgA, has not yet been investigated in COVID-19. It is downregulated in chronic obstructive pulmonary disease and associated with disease severity.[32] In our study, the reduction in urinary pIgR fragments is associated with COVID-19 severity. In line with the reduction of urinary gelsolin fragments, patients with an unfavourable COVID-19 outcome have lower plasma levels of gelsolin.[33] The sodium/potassium-transporting ATPase subunit gamma (FYXD2) is highly expressed in the kidney. Reduced abundance of a peptide from FYXD2 in our current study was associated with severe COVID-19 in keeping with the same observation in IgA nephropathy.[34]

The urgency associated with the COVID-19 pandemic in Germany, Europe and beyond justified the generation of this interim report, as other investigators did working in this field.[35]

However, we did comply with all quality criteria as outlined in a recent commentary on the COVID-19 literature.[36] In accordance with the scientific rigor required in this research field, the CRIT-COV-U consortium is preparing a protocol amendment describing the statistical analysis plan and significance levels required for a second look at the CRIT-CoV data in the final analysis. Furthermore, as can be expected for an interim report, the current study has potential limitations. First, the sample size of the validation cohort was small compared with the derivation cohort. Second, models were not adjusted for glomerular filtration rate, because intravenous fluid administration confounds this renal function measurement. Presumably, this limitation is also applicable to other scoring algorithms. Finally, at the current stage of data collection, calibrating the predictive models was not yet possible, limiting the generalisability of the COV50 biomarker.

In conclusion, this first CRIT-COV-U report proves the concept that UPP generates biomarkers indicative of adverse COVID-19 outcomes, even at WHO stages 1-3. The current findings obviously need consolidation in the full dataset of 1000 patients, but open perspectives for patient management, health policy planning and for providing an intermediate UPP endpoint in randomised clinical trials of novel COVID-19 treatment modalities. COV50 is licensed in Germany and immediately available for clinical use. Based on the current results, attempts are being initiated in Germany to apply pre-emptive treatment in COVID-19 patients predicted to experience unfavourable outcomes.

Contributors

RW, CL, HM, JM, HvdL and JB conceptualised the study. HM, JS, JR and JM did the proteomic urine analyses. LT, JAS and JB did the statistical analysis and wrote the first draft of the manuscript, JAS was the principal writer of the final draft. RW, SK, AM, ED, GS, MM, ACT, BC, AW, BP, ÅN, MS, KR, CL, and JB were study investigators or participated in the conduct of the study, including recruitment

and follow-up of patients. BN and HvdL conducted and supervised eCRF database construction, data retrieval, management and quality control. All authors interpreted the results, commented on successive drafts of the manuscript and approved the final version.

Data sharing

The study protocol is available at the German Register for Clinical Studies (www.drks.de), number DRKS00022495. Anonymised participants data will be made available when the study is complete, upon request directed to the corresponding author. Proposals will be reviewed and approved by the funder, investigators and collaborators based on scientific merit. After approval of a proposal, data can be shared through a secure online platform after signing a data access and confidentiality agreement. All data will be made available for a minimum of 5 years from the end of the study.

Acknowledgments

Mrs Vera De Leebeeck and Mrs Renilde Wolfs provided expert clerical assistance.

Conflict of interest

HM, JM, JS and JR are past (JM) or current employees of Mosaiques-Diagnostics GmbH, Hannover, Germany. Dr. Mebazaa reports personal fees from Orion, Servier, Otsuka, Philips, Sanofi, Adrenomed, Epygon and Fire 1 and grants and personal fees from 4TEEN4, Abbott and Sphingotec

Panel: **Research in context**

Evidence before this study

A PubMed search without limitations of publication date or language using the terms "*COVID-19*" AND "*risk prediction*" produced 1734 results. Risk factors commonly predictive of death or adverse outcomes commonly included male sex, age, obesity, hypertension, diabetes, and other comorbidities. Adding the term "proteomics" to the search identified 11 articles, including one duplicate, published from 2020 until 2021, of which the full text was read. One study focused on the molecular basis of the COVID-19 manifestations by a proximity network analysis and incorporating SARS-CoV-2 virus-host protein-protein interactions, transcriptomics, and proteomics into the human interactome. A second study applied multiplatform metabolomics on the blood of 17 SARS-CoV-2 infected patients and 27 sex-matched controls and identified a discriminatory model with 100% sensitivity. One preliminary report ahead of peer-review described a proteomic profile in the serum of 49 COVID-19 patients predicting critical illness and death. Other studies addressed inflammatory, immunological and T cell responses to infection or were literature reviews.

Added value of this study

This study is the first to include a specific urinary proteomic biomarker (COV50) into a model predicting death and worsening WHO score up to 21 days from the PCR-confirmed SARS-CoV-2 infection. In the derivation cohort (n=228), the standardised odds ratios adjusted for sex, age, baseline WHO score and comorbidities were 2·73 (p=0·012) for mortality (n=28) and 3·38 (p<0·0001) for worsening WHO score (n=48). The COV50-associated AUC was 0·80 (95% CI, 0·72–0·88) for mortality and 0·74 (0·66-0·81) for worsening WHO score, yielding a sensitivity/specificity for optimised COV50 thresholds of 87·0%/74·6% and 77·1%/63·9%, respectively. On top of a base model including sex, age, body mass index,

comorbidities and the baseline WHO score, COV50 analysed as continuously distributed variable and per threshold improved the AUC for worsening WHO score from 0·68 to 0·78 (p=0·0097) and 0·75 (p=0·021), respectively. Findings in the validation cohort were confirmatory.

Implications of all the available evidence

SARS-CoV-2 infection affects multiple organs. The urinary proteome contains over 20,000 peptides and provides a specific molecular signature of systemwide pathophysiological processes. This first report of the Prospective Validation of a Proteomic Urine Test for Early and Accurate Prognosis of Critical Course Complications in Patients with SARS-CoV-2 Infection Study (CRIT-COV-U) proves the concept that UPP generates biomarkers indicative of disease outcome even at the preclinical stage of SARS-CoV-2 infection. COV50 is licensed in Germany and immediately available for clinical use. Based on the current results, attempts are being initiated in Germany to apply pre-emptive treatment in COVID-19 patients predicted to experience unfavourable outcomes.

References

1 Huang C, Wang Y, Li X, et al. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *Lancet* 2020; **395:** 497-506.

2 Guan W, Ni Z, Hu Y, et al. Clinical characteristics of coronavirus disease 2019 in China. *N Engl J Med* 2020; **382:** 1708-20.

3 Williamson EJ, Walker AJ, Bhaskaran K, et al. Factors associated with COVID-19-related death using OpenSAFELY. *Nature* 2020; **584:** 430-36.

4 Wynants L, Van Calster B, Collins GS, et al. Prediction models for diagnosis and prognosis of covid-19: systematic review and critical appraisal. *Br Med J* 2020; **369:** m1328

5 Montero-Odasso M, Hogan DB, Lam R, et al. Age alone is not adequate to determine health-care resource allocation during the COVID-19 pandemic. *Can Geriatr J* 2020; **23:** 152-54.

6 Knight SR, Ho A, Pius R, et al. Risk stratification of patients admitted to hospital with covid-19 using the ISARIC WHO Clinical Characterisation Protocol : development and validation of the 4C Mortality Score. *Br Med J* 2020; **370:** m3339

7 Myrstad M, Ihle-Hansen H, Tveita AA, et al. National Early Warning Score 2 (NEWS2) on admission predicts severe disease and in-hospital mortality from Covid-19 - a prospective cohort study. *Scan J Trauma Resusc Emerg Med* 2021; **28:** 66

8 Puelles VG, Lütgehetmann M, Lindenmeyer MT, et al. Multiorgan and renal tropism of SARS-CoV-2. *N Engl J Med* 2020; **383:** 590-92.

9 Kimhofer T, Lodge S, Whiley L, et al. Integrative modeling of quantitative plasma lipoprotein, metabolic, and amino acid data reveals a multiorgan pathological signature of SARS-CoV=2 infection. *J Proteome Res* 2020; **19:** 4442-54.

10 Zürbig P, Renfrow MB, Schiffer E, et al. Biomarker discovery by CE-MS enables sequence analysis via MS/MS with platform-independent separation. *Electrophoresis* 2006; **27:** 2111-25.

11 Klein J, Papadopoulos T, Mischak H, Mullen W. Comparison of CE-MS/MS and LC-MS/MS sequencing demonstrates significant complementarity in natural peptide identification in human urine. *Electrophoresis* 2014; **35:** 1060-64.

12 Zhang ZY, Nkuipou-Kenfack E, Staessen JA. Urinary peptidomic biomarker for personalized prevention and treatment of diastolic left ventricular dysfunction. *Proteomics Clin Appl* 2019; **13:** e1800174

13 Pontillo C, Jacobs L, Staessen JA, et al. A urinary proteome-based classifier for the early detection of decline in glomerular filtration. *Nephrol Dial Transplant* 2017; **32:** 1510-16.

14 Tofte N, Lindhardt M, Adamova K, et al. Early detection of diabetic kidney disease by urinary proteomics and subsequent intervention with spironolactone to delay progression (PRIORITY) : a prospective observational study and embedded randomised placebo-controlled trial. *Lancet Diabetes Endocrinol* 2020; **8:** 301-12.

15 Wendt R, Kalbitz S, Lübbert C, et al. Urinary proteomics associate with COVID-19 severity : Pilot proof-of-principle data and design of multicentric diagnostic study. *Proteomics* 2020; **20:** e2000202

16 World Medical Association. World Medical Association Declaration of Helsinki. Ethical principles for medical research involving human subjects. *J Am Med Assoc* 2013; **310:** 2191-94.

17 Corman VM, Landt O, Kaiser M, et al. Detection of 2019 novel coronavirus (2019-nCoV) by real-time RT-PCR. *Eurosurveillance* 2020; **25:** 2000045

18 World Health Organization. Ordinal scale for clinical improvement. WHO R&D Blueprint Novel Coronavirus. 2020;

19 Levey AS, Bosch JP, Lewis JB, et al. A more accurate method to estimate glomerular filtration rate from serum creatinine : a new prediction equation. *Ann Intern Med* 1999; **130:** 461-70.

20 Mischak H, Vlahou A, Ioannidis JPA. Technical aspects and inter-laboratory variability in native peptide profiling: the CE-MS experience. *Clin Biochem* 2013; **46:** 432-43.

21 Latosinska A, Siwy J, Mischak H, Frantzi M. Peptidomics and proteomics based on CE-MS as a robust tool in clinical application: the past, the present, and the future. *Electrophoresis* 2019; **40:** 2294-308.

22 Benjamini Y, Hochberg Y. Controlling the false discovery rate : a practical and powerful approach to multiple testing. *J Royal Stat Soc B* 1995; **57:** 289-300.

23 Schaub S, Wilkins J, Weiler T, Sangster K, Rush D, Nickerson P. Urine protein profiling with surface-enhanced laser-desorption/ionization time-of-flight mass spectrometry. *Kidney Intern* 2004; **2004:** 323-32.

24 Theodorescu D, Wittke S, Ross MM, et al. Discovery and validation of new protein biomarkers for urothelial cancer : a prospective analysis. *Lancet Oncol* 2006; **7:** 230-40.

25 Fliser D, Novak J, Thongboonkerd V, et al. Advances in urinary proteome analysis and biomarker discovery. *J Am Soc Nephrol* 2007; **18:** 1057-71.

26 Nunnari G, Sanfilippo C, Castrogiovanni P, et al. Network perturbation analysis in human bronchial epithelial cells following SARS-CoV2 infection. *Exp Cell Res* 2020; **395:** 122204

27 Schanstra JP, Zürbig P, Alkhalaf A, et al. Diagnosis and prediction of CKD progression by assessment of urinary peptides. *J Am Soc Nephrol* 2015; **26:** 1999-2010.

28 ERA-EDTA Council and the ERACODA Working Group. Chronic kidney disease is a risk factor for severe COVID-19: a call to action by ERA-EDTA. *Nephrol Dial Transplant* 2021; **36:** gfaa314

29 Yang C, Chapman KR, Wong A, Liu M. a1-antitrypsin deficiency and the risk of COVID-19: an urgent call to action. *Lancet Respir Med* 2021; **9:** doi.org/10.1016/ S2213-2600(21)00018-7

30 Winger RC, Koblinski JE, Kanda T, Ransohoff RM, Muller W. Rapid remodeling of tight junctions during oaracellular diapedesis in a human model of the bloodpbrain barrier. *J Immunol* 2014; **193:** 2427-37.

31 Muller WA. Transendothelial migration: unifying principles from the endothelial perspective. *Immunol Rev* 2016; **273:** 61-75.

32 Gohy ST, Detry BR, Lecocq M, et al. Polymeric immunoglobulin receptor down-regulation in chronic obstructive pulmonary disease. Persistence in the cultured epithelium and role of transforming growth factor-β. *Am J Respir Crit Care Med* 2014; **190:** 509-21.

33 Abers MS, Delmonte OM, Ricotta EE, et al. An immune-based biomarker signature is associated with mortality in COVID-19 patients. *JCI Insight* 2021; **6:** e144455

34 Rudnicki M, Siwy J, Wendt R, et al. Urine proteomics for prediction of disease progression in patients with IgA nephropathy. *Nephrol Dial Transplant* 2020; **35:** 9faa307

35 Chen P, Nirula A, Heller B, et al. SARS-CoV-2 neutralizing antibody LY-CoV555 in outpatients with Covid-19. *N Engl J Med* 2020; **384:** 229-37.

36 Ferreira JP, Epstein M, Zannad F. The decline of the experimental paradigm during the COVID-19 pandemic: a template for the future. *Am J Med* 2021; **134:** 166-75.

Legends to figure

Performance of COV50 on top of other baseline risk factors in the derivation cohort to discriminate death from survival (panels A-C) and progression from non-progression in the baseline WHO score during follow-up (panels D-F) in the derivation cohort

The base model included sex, age, body mass index and the presence of comorbodities: hypertension, heart failure, diabetes or cancer. In subsequent steps, the baseline WHO score was added and next COV50 as a continuously distributed variable (panels B and E) or as a categorised variable based on an optimised threshold of -0.47 for mortality (panel C) or -0.04 for a worsening WHO score (panel F).

Table 1: **Baseline characteristics**

Characteristic	Derivation cohort	Validation cohort	p value
Number in cohort	228	99	
Main study variables			
WHO score			
1-3	90 (39·5)	9 (37·4)	
4-5	107 (46·9)	76 (60·6)	<0·0001
6	31 (13·6)	14 (2·0)	
COV50 level	-0·19 (1·52)	-0·17 (1·23)	0·92
Number with characteristic (%)			
White ethnicity	205 (89·9)	91 (91·9)	0·68
Women	94 (41·2)	43 (43·4)	0·72
Non-smoker	109 (47·8)	58 (58·6)	0·031
Hypertension	137 (60·1)	66 (66·7)	0·27
Heart failure	30 (13·6)	27 (27·8)	0·0039
Body mass index ≥30 kg/m^2	68 (29·8)	26 (26·3)	0·60
Diabetes mellitus	65 (28·5)	41 (41·4)	0·028
Cancer	13 (5·7)	7 (7·1)	0·62
Use of RAS blockers,	119 (52·2)	55 (55·6)	0·27
Mean (SD)of characteristic			
Age	63·1 (17·1)	66·8 (16·1)	0·68
Systolic blood pressure, mm Hg	130·0 (23·4)	127·5 (20·2)	0·35
Diastolic blood pressure, mm Hg	79·9 (55·0)	75·6 (12·2)	0·45
Heart rate, beats per minute	83·4 (15·0)	82·8 (17·9)	0·75
Body mass index, kg/m^2	28·1 (6·0)	27·4 (4·6)	0·24
Glomerular filtration rate, ml/min/1.73 m^2	76·7 (30·9)	78·7 (30·4)	0·63

RAS blockers indicate blocker of the renin-angiotensin system, including angiotensin-converting enzyme inhibitors and angiotensin-receptor blockers. The glomerular filtration rate was derived from serum creatinine, using the Chronic Kidney Disease Epidemiology Collaboration. The p value indicates the difference between the baseline characteristics of the derivation and validation cohort.

Table 2: **Baseline characteristics by fourths of the baseline COV50 distribution in the derivation cohort**

Characteristic	Low	Medium-low	Medium-high	High	p value for trend
COV50 limits	-1·23	[-1·23, -0·20[	[-0·20, 0·90[	≥0·90	
Number in group	57	57	57	57	
Main study variables					
WHO score					
1-3	50 (8·7)	20 (35·1)	19 (33·3)	1 (1·8)	
4-5	7 (12·3)	35 (61·4)	37 (64·9)	28 (49·1)	<0·0001
6-8	0	2 (3·5)	1 (1·8)	28 (4·9)	
COV50 level	-2·13 (0·50)	-0·77 (0·30)	0·29 (0·28)	1·85 (0·59)	<0·0001
Number with characteristic (%)					
Women	28 (49·1)	27 (47·4)	25 (43·9)	14 (24·6)	0·0087
Non-smoker	32 (56·1)	22 (38·6)	22 (38·6)	33 (57·9)	0·36
Hypertension	23 (40·4)	35 (61·4)	42 (73·7)	37 (64·9)	0·0034
Heart failure	1 (1·8)	7 (12·5)	14 (25·5)	8 (14·5)	0·016
Body mass index ≥30 kg/m^2	14 (24·6)	18 (31·6)	16 (28·1)	20 (35·1)	0·30
Diabetes mellitus	6 (10·5)	9 (15·8)	20 (35·1)	30 (52·6)	<0·0001
Cancer	2 (3·5)	6 (10·5)	2 (3·5)	3 (5·3)	0·62
Use of RAS blockers,	16 (28·1)	30 (52·6)	39 (68·4)	34 (59·4)	0·0024
Mean (SD)of characteristic					
Age	49·5 (16·8)	63·9 (17·2)	71·0 (13·8)	67·8 (12·1)	<0·0001
Systolic blood pressure, mm Hg	128·9 (23·9)	130·1 (23·1)	134·8 (21·6)	126·3 (24·5)	0·83
Diastolic blood pressure, mm Hg	79·8 (13·0)	77·9 (13·0)	77·1 (11·8)	70·6 (20·0)	0·0014
Heart rate, beats per minute	81·6 (12·1)	81·3 (13·5)	81·6 (14·1)	89·0 (18·5)	0·011
Body mass index, kg/m^2	27·2 (5·2)	28·2 (6·3)	28·2 (5·3)	29·0 (7·0)	0·12
Glomerular filtration rate, ml/min/1.73 m^2	86·3 (23·1)	81·8 (26·7)	74·7 (31·9)	69·6 (35·3)	0·0083

RAS blockers indicate blocker of the renin-angiotensin system, including angiotensin-converting enzyme inhibitors and angiotensin-receptor blockers. The glomerular filtration rate was derived from serum creatinine, using the Chronic Kidney Disease Epidemiology Collaboration.

Table 3: **Odds ratios relating outcome to COV50**

Outcome	Number E/R	Odds ratio (95% confidence interval)	p value
Mortality			
Unadjusted	23/228	3·52 (2·02–6·13)	<0.0001
Adjusted			
Sex and age		3·23 (1·81–5·74)	<0.0001
+ baseline WHO score		2·63 (1·21–5·69)	0·014
+ body mass index and comorbidities		2·73 (1·25–5·95)	0·012
Progressing WHO score			
Unadjusted	48/228	2·63 (1·80–3·85)	<0·0001
Adjusted			
Sex and age		2·37 (1·58–3·54)	<0·0001
+ baseline WHO score		3·34 (1·83–6·07)	<0·0001
+ body mass index and comorbidities		3·38 (1·85–6·17)	<0·0001

Number E/R indicates the number of events/number at risk. Odds ratios express the risk for 1-SD increment in COV50. Comorbidities include hypertension, heart failure, diabetes and cancer.

***Table 4*: Discriminative performance of COV50**

Outcome	Derivation cohort	Validation cohort
Mortality		
Number events/at risk	23/228	10/99
Continuously distributed COV50		
AUC (95% confidence interval)	0·82 (0·74–0·89)	0·83 (0·71–0·94)
Cross-validated AUC (95% confidence interval)	0·80 (0·72–0·88)	NA
Categorised COV50		
Youden cut-off threshold	0·47	0·47
Sensitivity	87·0 (73·2–1·00)	80·0 (55·0–1·00)
Specificity	74·6 (68·7–80·6)	70·8 (61·3–80·2)
Progressing WHO score		
Number events/at risk	48/228	23/99
Continuously distributed COV50		
AUC (95% confidence interval)	0·75 (0·67–0·82)	0·70 (0·58–0·88)
Cross-validated AUC (95% confidence interval)	0·74 (0·66–0·81)	NA
Categorised COV50		
Youden cut-off threshold	0·04	0·04
Sensitivity (95% confidence interval)	77·1 (65·2–89·0)	73·9 (56·0–91·9)
Specificity (95% confidence interval)	63·9 (56·9–70·9)	63·2 (52·3–74·0)

AUC indicates area under the curve. The AUC in the validation cohort was derived from the probabilities as predicted by the logistic model in the derivation cohort. Sensitivity and specificity in the validation cohort were based on the thresholds obtained in the derivation cohort. NA indicates not applicable.

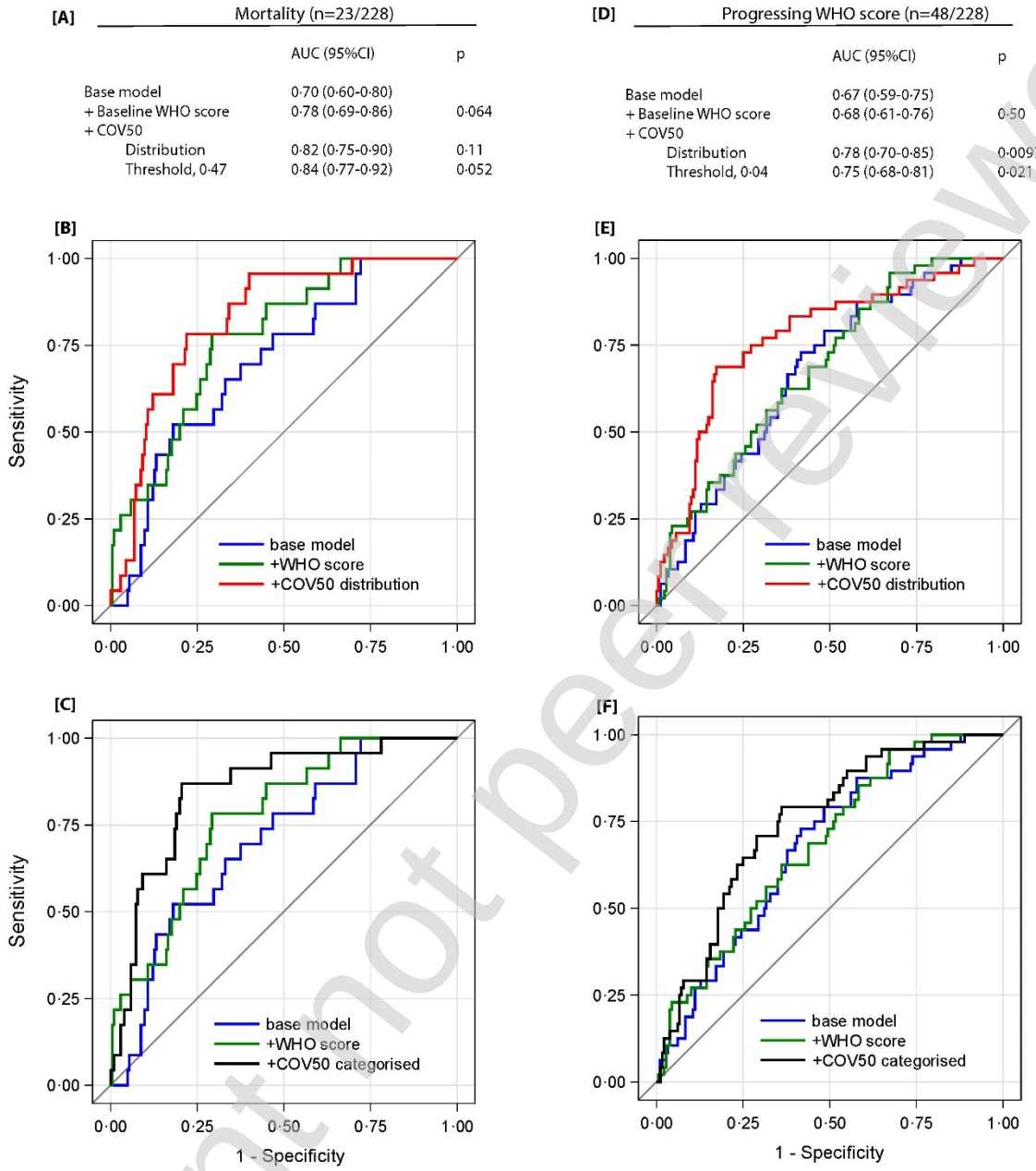

Figure 1:

Performance of COV50 on top of other baseline risk factors in the derivation cohort to discriminate death from survival (panels A-C) and progression from non-progression in the baseline WHO score during follow-up (panels D-F) in the derivation cohort

The base model included sex, age, body mass index and the presence of comorbidities: hypertension, heart failure, diabetes or cancer. In subsequent steps, the baseline WHO score was added and next COV50 as a continuously distributed variable (panels B and E) or as a categorised variable based on an optimised threshold of 0.47 for mortality (panel C) or 0.04 for a worsening WHO score (panel F).