

Validation of calibration transfer in ATR-FTIR spectroscopy combined with machine

UNIVERSITAT DE VALÈNCIA Facultat de Química



Hospital Universitario Puerta de Hierro Majadahonda

2 University Hospital Puerta de Hierro de Majadahonda, Majadahonda, Spain. RISCORS2040 RD24/0004/0028

3 Department of Medicine. Universidad Autónoma de Madrid, Spain

4 University Hospital Doctor Peset Aleixandre, Valencia, Spain

5 University Hospital La Fe, Valencia, Spain

6 Neuroimmunology Unit, Instituto de Investigación Sanitaria Puerta de Hierro-Segovia de Arana, Madrid, Spain.

7 Biobank, Instituto de Investigación Sanitaria Puerta de Hierro-Segovia de Arana, Madrid, Spain.

8 Centre for Biospectroscopy, Monash University, Clayton, Victoria, Australia

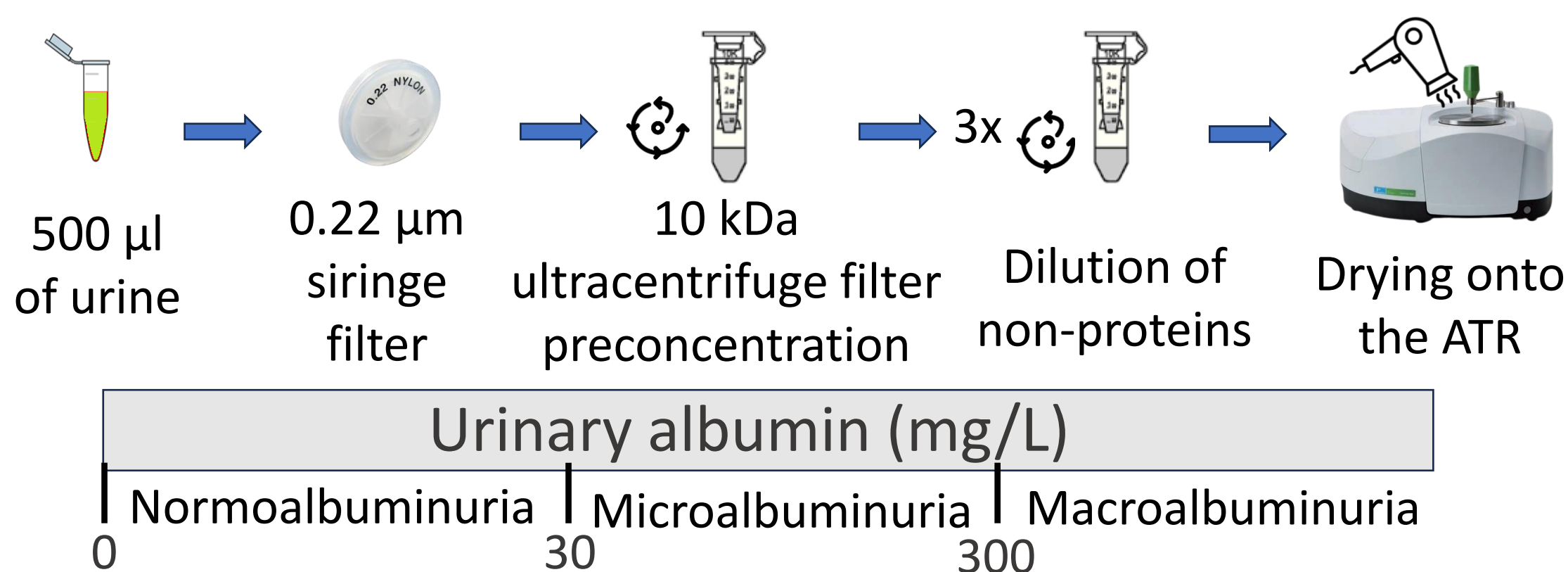


*angel.illana@uv.es

INTRODUCTION & OBJECTIVES

Infrared (IR) spectroscopy of urine extracts coupled with Machine Learning methods [has been proposed](#) as a promising technique for **Diabetic Kidney Disease (DKD)** screening. However, for **clinical translation**, predictive models based on machine learning techniques necessitate substantial sets of samples for calibration and testing under various experimental parameters and populations, posing a critical barrier to the development of **globally deployable tools**. Here, we aim to assess the methodology's ability to establish **DKD diagnostic models** applicable across **diverse populations, instruments and experimental conditions worldwide** based on urinary albumin.

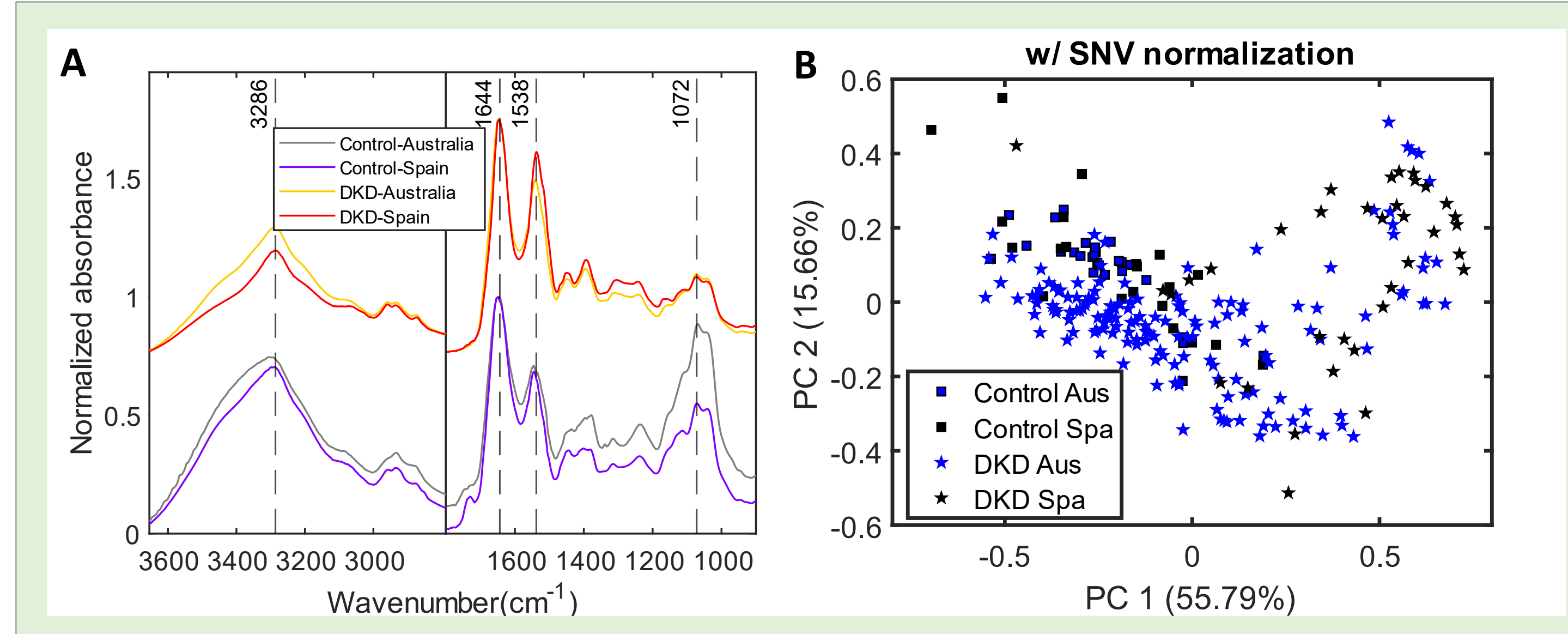
EXPERIMENTAL METHODOLOGY



SAMPLES OVERVIEW

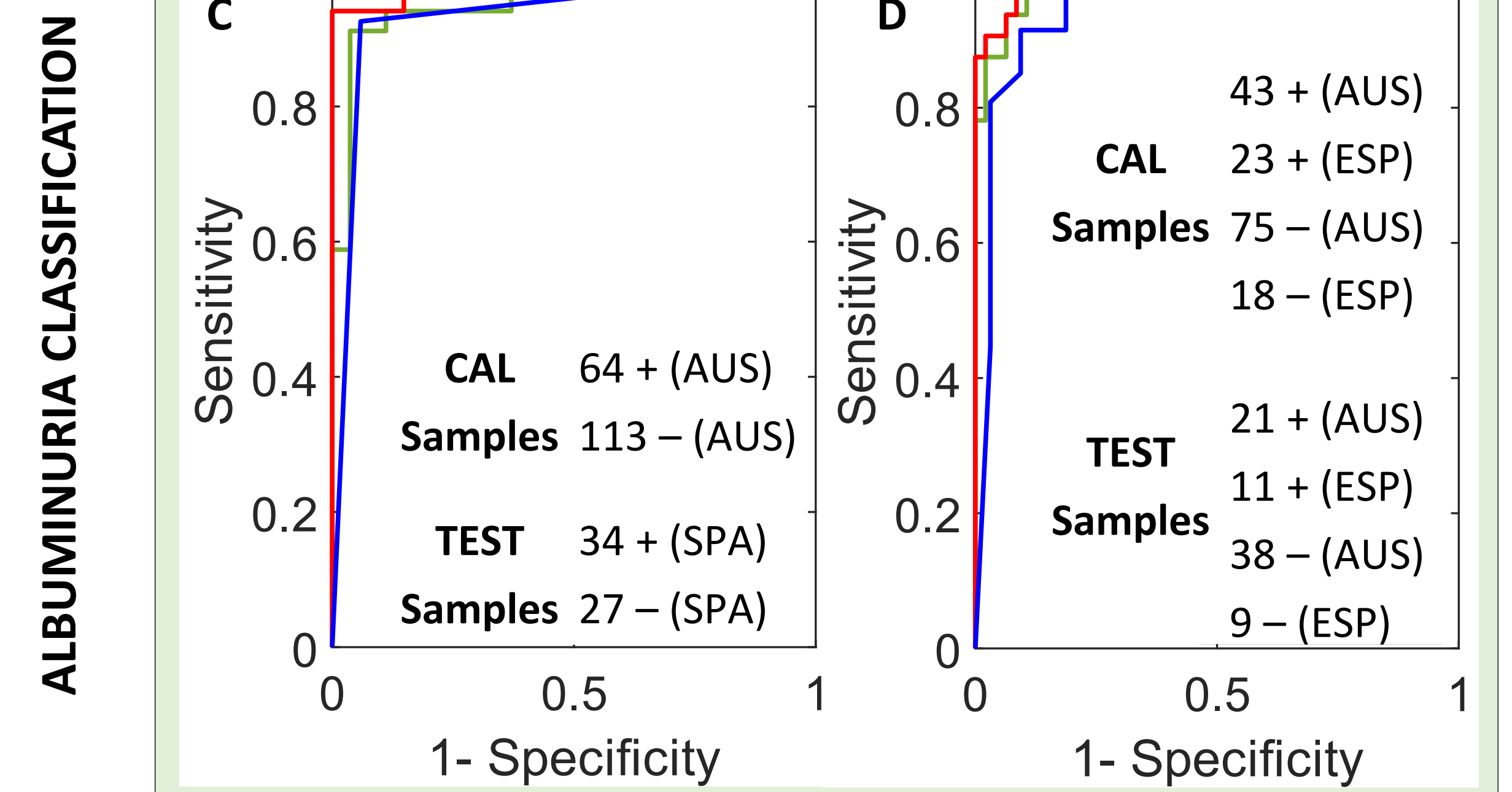
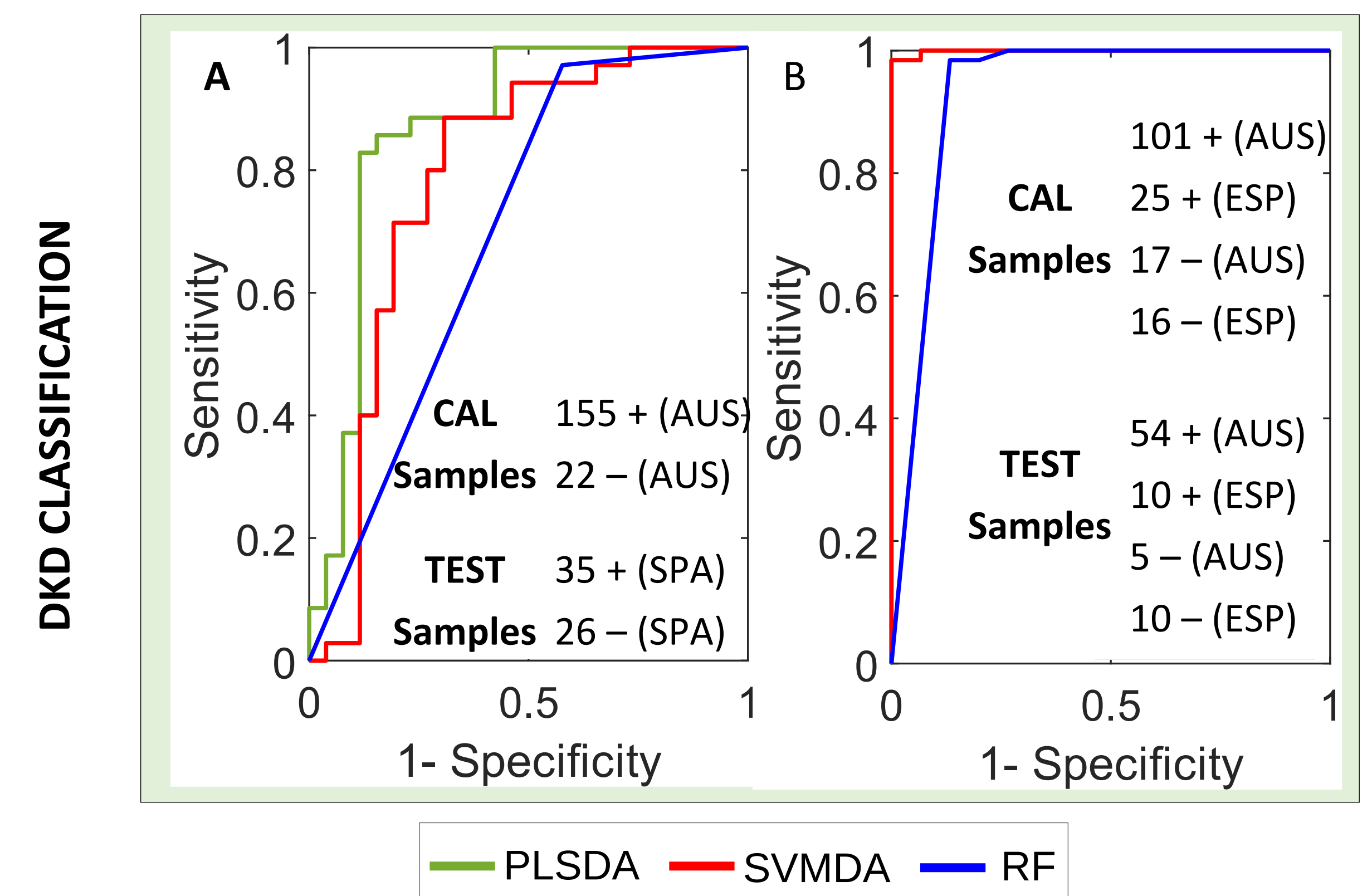
Characteristics of the samples of this study including the different devices used for processing and measurement

Samples origin	DKD [N (%Males)]	Control [N (%Males)]	Collection	0.22 µM filter	10kDa filter	System	Volume (µL)
Melbourne (AUS) 	155 (100)	22 (100)	24 h	Cellulose acetate membrane	Amicon® Regenerated Cellulose	Bruker Alpha	0.8
Valencia (ESP) 	20 (70)	26 (62)	Spot urine	Nylon membrane	Vivaspin® polyethersulfone (PES)	Perkin-Elmer Spectrum Two	1.0
Madrid (ESP) 	15 (80)	0 (0)					



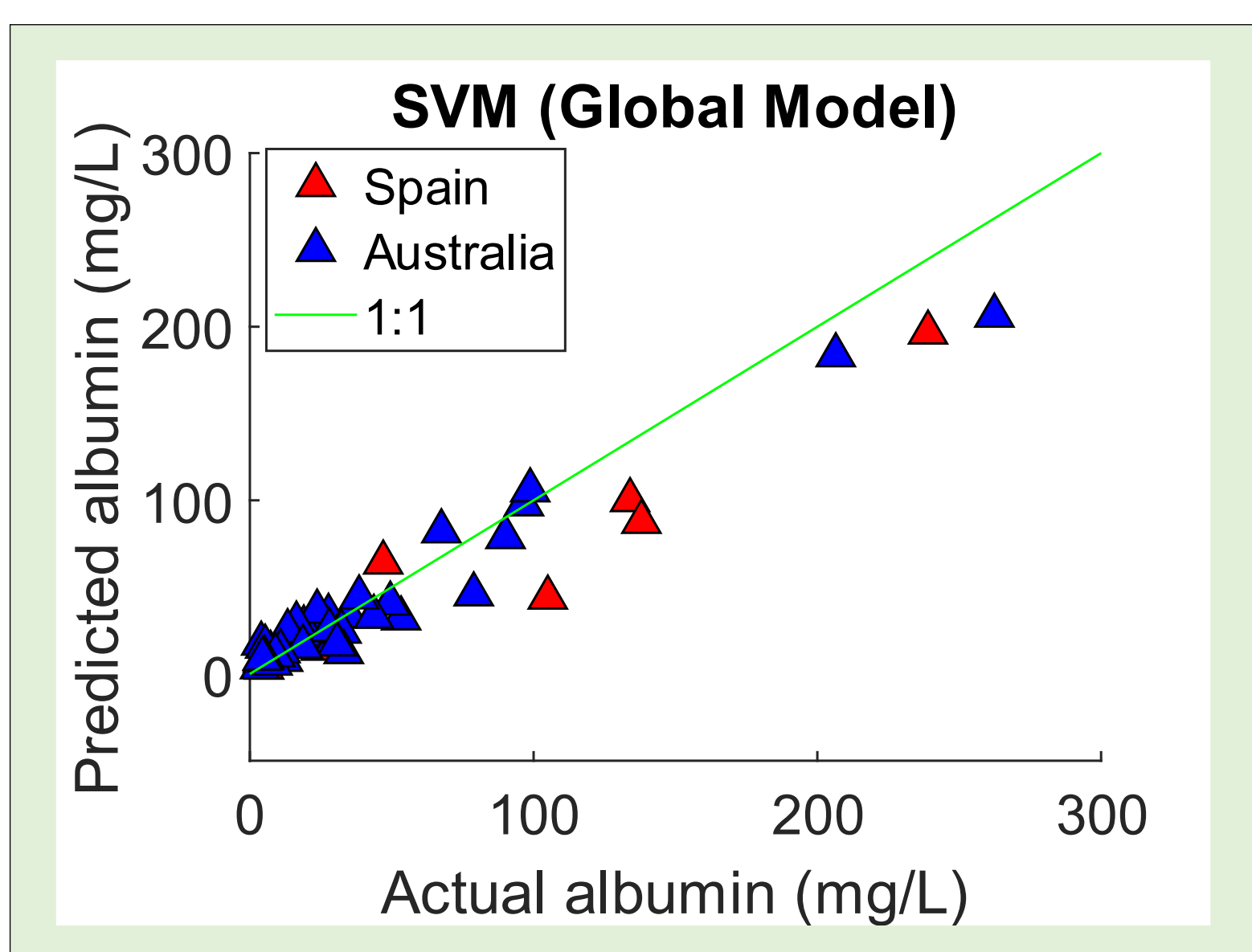
Normalized spectra (A) indicate differences between DKD and Control. PCA (B) depicts that most **data variability is associated with the disease** rather than the sample origin or instrument.

CLASSIFICATION



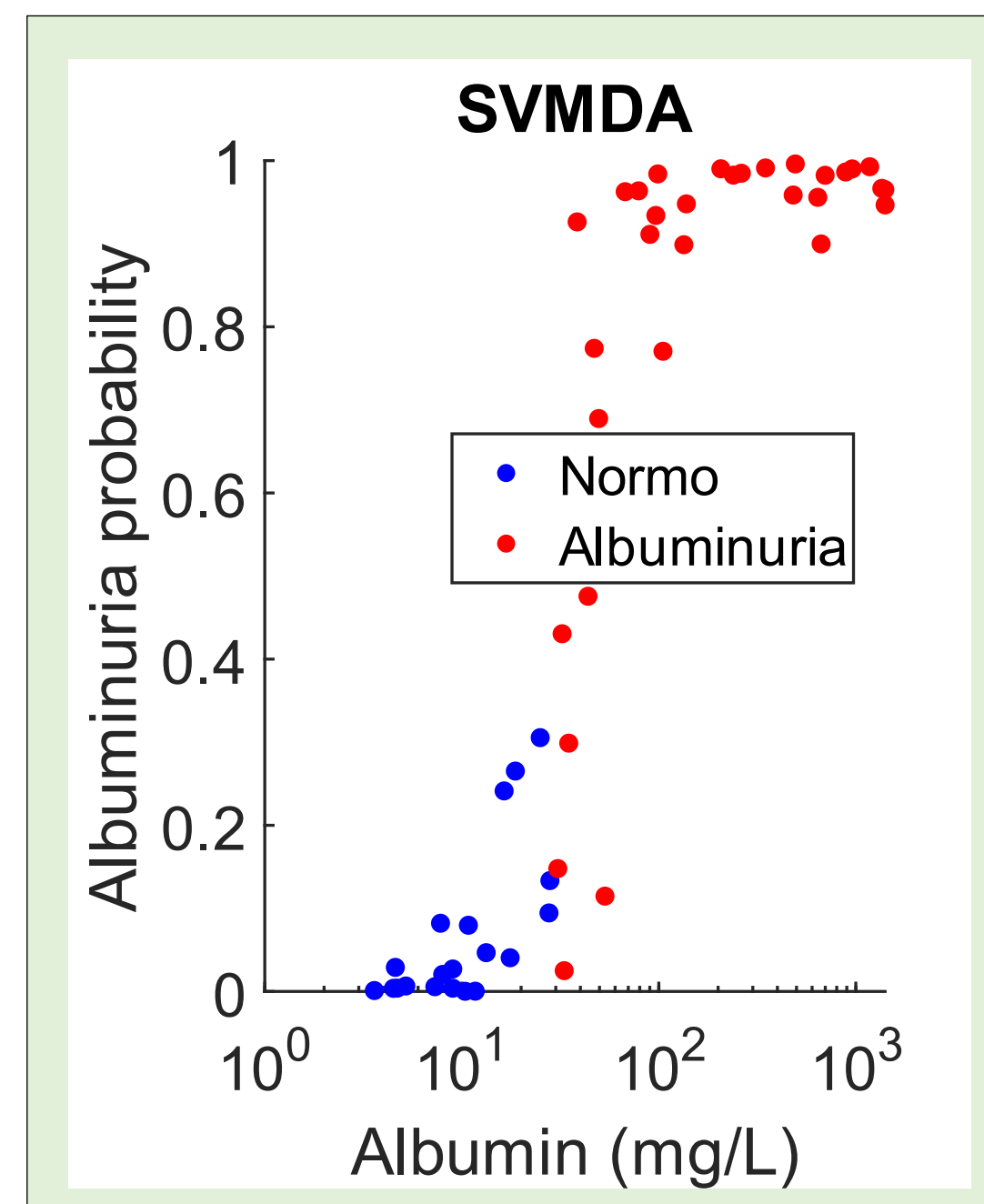
Models **developed using Australian spectra (A, C)** successfully **predicted Spanish samples**, achieving AUROC values of 0.87 and 0.98 for DKD and microalbuminuria identification, respectively. Both values improved to 0.99 when a global model (B, D) was calibrated and independently tested with a combined set integrating samples from both countries.

QUANTIFICATION



	R ² (Pred)	RMSEP (mg/L) (NormoAlb)	RMSEP (mg/L) Normalized (NormoAlb)	RMSEP (mg/L) (MicroAlb)	RMSEP (mg/L) Normalized (MicroAlb)
PLSR	0.89	18	1.26	28	0.29
SVR	0.93	7	0.52	28	0.30
RF	0.89	12	0.87	28	0.30

Models exhibited **strong predictive performance for albuminuria concentrations across the entire microalbuminuria range**. Urinary albumin concentrations from Australia and Spain were predicted similarly, and do not show major clustering or bias in their predictions.



There is a classification grey zone around 30 mg/L of albumin (the cut-off for albuminuria diagnostic)

CONCLUSIONS

The robustness of spectral markers in IR spectra, minimizes variability from sample type and instrumentation. Due to the absence of a certified reference material (CRM) for albuminuria and urinary albumin-to-creatinine ratio (UACR), model validation was performed through comparative analysis with reference methods.

Recalibrating the model with the new samples significantly improves the prediction results without the need of measuring a subset of samples in both instruments.

Want the whole picture? Read our preprint here: [10.26434/chemrxiv-2025-pxsq8](https://doi.org/10.26434/chemrxiv-2025-pxsq8)



ACKNOWLEDGEMENTS

This study has been funded by Instituto de Salud Carlos III (ISCIII) through the projects PI23/00135, PI23/01518 and research network RICORS RD24/0004/0028 and co-funded by the European Union. Authors also acknowledge the financial support from the projects RYC2019-026556-I and RPID2020-119326RA-I00 funded by MCIN/AEI/10.13039/501100011033. V.N.E. acknowledges the financial support by the grant PRE2021-098833 from MCIN/AEI/10.13039/501100011033 and the FSE+. A.S.I. acknowledges the support of grant JDC2022-049354-I funded by MCIN/AEI/10.13039/501100011033 and by the "European Union NextGenerationEU/PRTR". The authors wish to thank the donors, and the Hospital Universitario Puerta de Hierro Majadahonda (HUPHM)/Instituto de Investigación Sanitaria Puerta de Hierro-Segovia de Arana (IDIPHSA) Biobank (Carlos III Health Institute Biomodels and Biobanks Platform – PT23/00015) for the human specimens used in this study.

