eGFR mini project

# Aim

This mini project stems from my master's thesis conducted at Stockholm University[[1]](#footnote-1),[[2]](#footnote-2). The focus is on investigating the eGFR calculation in the MIMIC-IV database while identifying potential issues within the dataset and proposing possible solutions to mitigate them. The primary aim is to develop an understanding of the eGFR calculation process and contribute to enhancing the accuracy and reliability of the results. The insights gained from this project can be valuable for future researchers seeking to extract information on kidney health from the dataset. Top of Form

Bottom of Form

# eGFR and the CKD-EPI(2021) equation

The acronym eGFR refers to estimated Glomerular Filtration Rate, which serves as a cost-effective and accurate method for evaluating kidney function. Depending on the eGFR value and other factors beyond the scope of this project, healthcare professionals may recommend lifestyle modifications and/or prescribe specialized medications tailored to each individual patient.The National Kidney Foundation offers an informative fact sheet[[3]](#footnote-3) and a CKD heat map[[4]](#footnote-4) documents that provide a comprehensive introduction to eGFR and Chronic Kidney Disease (CKD).

For the most up-to-date calculation of eGFR, researchers have proposed the CKD-EPI(2021) equation [1]. The National Kidney Foundation also offers a concise information page dedicated to the CKD-EPI equation [2]. The eGFR is expressed in milliliters per minute per 1.73 square meters (), and the equation is as follows:

Where:

* **Scr**is the serum creatinine in mg/dL
* **K** is a constant, 0.7 for females or 0.9 for males
* **a** is a constant, -0.241 for females or -0.302 for males
* **age** is the patient’s age in years

## The standardized Scr

Standardized Scr refers to laboratory measurements of serum creatinine (Scr) that have been calibrated based on the National Institute of Standards and Technology (NIST) standard reference material (SRM) 967. References [3], [4] outline the methodology of the standardization and its importance in eGFR results, respectively. With the introduction of the SRM 967 in 2007, manufacturers have been required to follow this standard ever since [4].

# The MIMIC-IV dataset

*“MIMIC (Medical Information Mart for Intensive Care) is a large, freely-available database comprising deidentified health-related data from patients who were admitted to the critical care units of the Beth Israel Deaconess Medical Center”* [5]. While the database is freely-available, individuals are required to complete the necessary training in [CITI Data or Specimens Only Research](https://physionet.org/about/citi-course/), and subsequently apply for data access through the [physionet.org](https://physionet.org/) website. You can find my training certificate confirming completion of the required training [here](https://www.citiprogram.org/verify/?wf6b514f4-d64e-4448-96bb-977173e0a3b1-47385825). Access to MIMIC is offered through two methods: cloud-based [6] or local [7]. For my master thesis I chose to access the data in the BigQuery db. For my master's thesis, I opted to access the data in the BigQuery database. In this project, I have chosen to set up the local database, specifically utilizing a Postgres installation as outlined in the mimic-IV [GitHub page](https://github.com/MIT-LCP/mimic-code/tree/main/mimic-iv/buildmimic/postgres), as I wanted to explore the capabilities of the PostgresSQL language.

# References

[1] L. A. Inker *et al.*, ‘New Creatinine- and Cystatin C–Based Equations to Estimate GFR without Race’, *N. Engl. J. Med.*, vol. 385, no. 19, pp. 1737–1749, Nov. 2021, doi: 10.1056/NEJMoa2102953.

[2] ‘CKD-EPI Creatinine Equation (2021)’, *National Kidney Foundation*, Oct. 01, 2021. https://www.kidney.org/professionals/kdoqi/gfr\_calculator/formula (accessed May 19, 2023).

[3] N. G. Dodder, S. S.-C. Tai, L. T. Sniegoski, N. F. Zhang, and M. J. Welch, ‘Certification of Creatinine in a Human Serum Reference Material by GC-MS and LC-MS’, *Clin. Chem.*, vol. 53, no. 9, pp. 1694–1699, Sep. 2007, doi: 10.1373/clinchem.2007.090027.

[4] H. Pottel *et al.*, ‘Standardization of serum creatinine is essential for accurate use of unbiased estimated GFR equations: evidence from three cohorts matched on renal function’, *Clin. Kidney J.*, vol. 15, no. 12, pp. 2258–2265, Dec. 2022, doi: 10.1093/ckj/sfac182.

[5] ‘About MIMIC’, *MIMIC*. https://mimic.mit.edu/docs/about/ (accessed May 24, 2023).

[6] ‘Cloud’, *MIMIC*. https://mimic.mit.edu/docs/gettingstarted/cloud/ (accessed May 24, 2023).

[7] ‘Local database setup’, *MIMIC*. https://mimic.mit.edu/docs/gettingstarted/local/ (accessed May 24, 2023).

1. https://drive.google.com/file/d/1R\_RTf9gSW9CxZMnw5xjPyk8NLDtpDk8B/view?usp=sharing [↑](#footnote-ref-1)
2. https://github.com/fdiamant/Thesis [↑](#footnote-ref-2)
3. https://www.kidney.org/sites/default/files/01-10-8374\_2212\_patflyer\_egfr.pdf [↑](#footnote-ref-3)
4. https://www.kidney.org/sites/default/files/kidney-numbers\_ckd-heatmap.pdf [↑](#footnote-ref-4)