Master's thesis

NTNU
Norwegian University of Science and Technology
Faculty of Medicine and Health Sciences
Department of Circulation and Medical Imaging

Daria Obratov

NECESSARY STEPS FOR THE APPLICATION OF AN INTEGRATIVE "OMICS" SOLUTION TO THE DETECTION OF RECOMBINANT HUMAN ERYTHROPOIETIN

Master's thesis in Exercise Physiology
Supervisor: Professor Ulrik Wisløff
Co-supervisor: Professor Yannis Pitsiladis
April 2024

NTNU
Norwegian University of
Science and Technology

Daria Obratov

NECESSARY STEPS FOR THE APPLICATION OF AN INTEGRATIVE "OMICS" SOLUTION TO THE DETECTION OF RECOMBINANT HUMAN ERYTHROPOIETIN

Master's thesis in Exercise Physiology
Supervisor: Professor Ulrik Wisløff
Co-supervisor: Professor Yannis Pitsiladis
April 2024

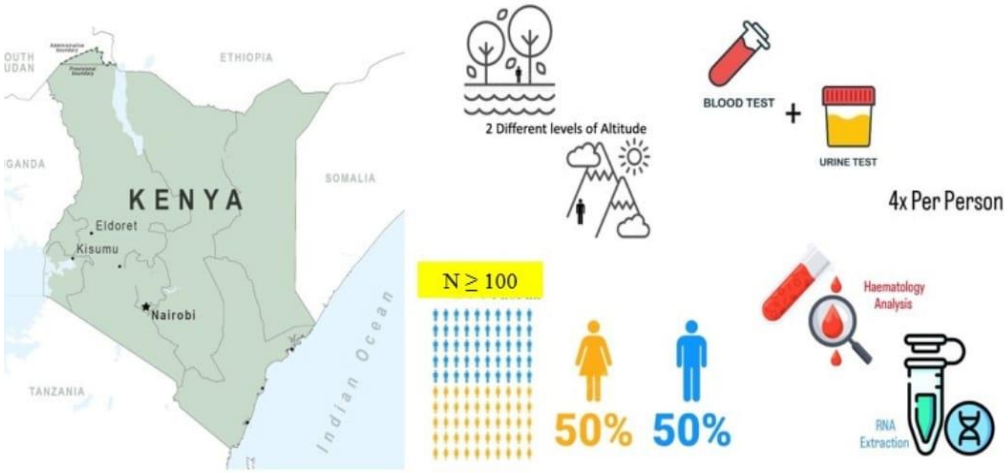
Norwegian University of Science and Technology
Faculty of Medicine and Health Sciences
Department of Circulation and Medical Imaging

NTNU
Norwegian University of
Science and Technology

NECESSARY STEPS FOR THE APPLICATION OF AN INTEGRATIVE "OMICS" SOLUTION TO THE DETECTION OF RECOMBINANT HUMAN ERYTHROPOIETIN

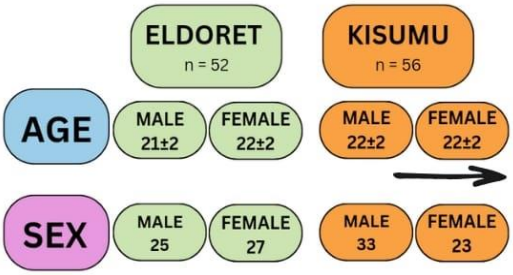
Aim of the study was to establish a biobank from blood samples of non-doping participants. Future analyses of these samples will focus on their omics profiles, which will serve as a basis for developing transcriptomic references and tests. These tests will be designed to enhance the detection of Recombinant Human Erythropoietin (rHuEpo), thereby improving anti-doping tests.

METHODS



CONCLUSION

We are non-doping cohort group whose blood samples have been instrumental in creating a biobank. Our blood samples will be analysed for "omics profile" and used for transcriptomic reference. This will help in developing transcriptomic tests designed to improve detection of RHuEPO.



Abstract

Background: The World Anti-Doping Agency (WADA) bans the use of recombinant human erythropoietin (rHuEpo) in sports, challenging to detect with the Athlete Biological Passport (ABP) that monitors haematological data longitudinally. Since 2009, the ABP has identified potential doping trends, but the individual variability in transcriptomic signatures related to rHuEpo, high altitude, and exercise among non-doping individuals is still uncertain. This study seeks to create a biobank of non-doping samples to improve ABP's detection capabilities and establish transcriptomic reference ranges to reduce false doping results.

Methods: Four blood and urine samples were collected from 108 university students based in Eldoret, Kenya (~2100 above sea level) and Kisumu, Kenya (~1000 m above sea level) with 4 to 6 weeks between each collection. The students included Eldoret males (21±2 years), Eldoret females (22±2 years), Kisumu males (22±2) and Kisumu females (22±2 years). Blood was collected into a K2EDTA and a Tempus™ Blood RNA Tube for haematological and transcriptomic analysis, respectively. Haematological variables used as blood doping markers in the ABP include Red Blood Cells (RBC), Haematocrit (HCT), Haemoglobin (HGB), Mean Corpuscular Haemoglobin (MCH), Mean Corpuscular Haemoglobin Concentration (MCHC), Mean Corpuscular Volume (MCV), Number of Reticulocytes (RET#) and Reticulocytes percentage (RET%). The Off-score was calculated for each sample using the formula: Hgb x 10 - 60(√RET%). The "clean" status of athletes was assessed using an ABP style model, created in MATLAB (version 6.1.0 with Statistics Toolbox version 3.0). Cut-off was applied with an adaptive Bayesian model to calculate individualized upper and lower limits for these variables, incorporating factors such as mean subject variance, between-subject variance, sex, and baseline data. This method aimed to distinguish between drug-free samples, which stayed within these personalized limits and suspicious samples which deviated significantly. Statistical analysis of haematological variables such as HGB, RET% and OFF-score, crucial for doping detection, were performed using R (R Studio, Version 1.2.5042, ABPS package, Vienna, Austria).

Results: Males from both Eldoret and Kisumu consistently exhibited higher ($p<0.05$) haematological variables than their female counterparts. However, female participants from both Kisumu and Eldoret showed a significantly higher ($p<0.05$) RET% compared to males. None of the participants from Eldoret exceeded the Bayesian cut-off for any haematological variable. Participants from Kisumu exceeded the cut-offs at only three time points for both sexes, OFF-score values, for females HGB values. Sixty participants surpassed the ABPS cut-off. Transcriptomic analysis has not yet been conducted, but results are anticipated by July 2024.

Conclusions: The blood samples collected in this study offer invaluable insights into the haematological reference values for healthy, non-doping Kenyan student-athletes and serve as the critical establishment of a control group. This foundational step is crucial for the next phase of this research, which involves developing transcriptomic tests designed to improve the detection of rHuEpo doping.

Acknowledgments

I extend my deepest gratitude to my NTNU supervisor Professor Ulrik Wisløff, whose approval and support made it possible for me to pursue my master's thesis beyond the NTNU University. This unique opportunity has been instrumental in my academic and personal growth, and for that, I am profoundly thankful.

My heartfelt appreciation goes to Professor Yannis Pitsiladis (Hong Kong Baptist University) who has been my mentor and guide throughout this journey. His unwavering support and guidance have been pivotal in every stage of my research, from conceptualization to completion. Professor Pitsiladis has been more than a supervisor; he has been a supporter in the truest sense, assisting me with challenges both big and small, and offering encouragement throughout the writing process. His dedication to my success has been an inspiration, for which I am grateful.

I extend my heartfelt appreciation to Prof. Yannis Pitsiladis team, along with Dr. Shaun Sutehall and Dr. Fernanda Malinsky, for their substantial guidance and insightful feedback during the writing process.

I would like to acknowledge the study participants, without whom this research would not have been possible.

This research was funded by the World Anti-Doping Agency (WADA).

Contents

Abbreviations

ABP	Athlete Biological Passport
CAS	Court of Arbitration for Sport
EAAS	Externally administered Anabolic Androgenic Steroids
ESAS	Erythropoiesis Stimulating Agents
Hbmass	Haemoglobin mass
HCT	Haematocrit
HGB	Haemoglobin
MCH	Mean Corpuscular Haemoglobin
MCHC	Mean Corpuscular Haemoglobin Content
MCV	Mean Corpuscular Volume
OFF– hr Score	Haemoglobin level and the percentage of Reticulocytes
RBC	Red Blood Count
RET #	Number of Reticulocytes
RET %	Reticulocytes Percentage
rHuEpo	Recombinant Human Erythropoietin
RNA	Ribonucleic Acid
SARMS	Androgen Receptor Modulators
SD	Standard Deviation
VO2 MAX	Maximal Oxygen Consumption
WADA	World Antidoping Agency
*	Significant difference between sex within location
§	Significant difference between location for the same sex

Introduction

Doping is commonly recognized as the use of prohibited substances and methods by athletes to enhance their performance in sports (Lippi et al., 1999). The origin of the term 'doping' is believed to derive from 'dope', a term once used for a basic alcoholic concoction utilized as a stimulant in ceremonial dances in South Africa (Lippi et al., 1999). Doping in sports dates back to the Ancient Olympics, where athletes used figs for performance enhancement (Holt et al., 2009). In the 19th century, athletes experimented with drug mixtures to boost strength and endurance with no initial legal restrictions, but extensive records exist of such practices (Holt et.al, 2009). However, following several fatalities, regulations against performance-enhancing drugs were gradually introduced (Holt et al., 2009). Doping enables the creation of a "super athlete," but it does so by disregarding the principles of fair play in sports and it poses significant health risks to athletes (Brzeziańska et al., 2014). Survey conducted by Bamberg et al., (1997) among athletes, revealed that 98% of respondents would be willing to use doping if they were guaranteed an Olympic medal without the risk of detection. When asked if they would take doping although it meant risking their lives, but with a guarantee of winning every competition for the next 5 years without getting caught, 50% answered affirmatively (Bamber et al., 1997).

According to the World Anti-Doping Agency (WADA), blood doping is characterized as the manipulation of blood or its components through physical or chemical methods (WADA, 2011). With a history of abuse in sports, spanning over fifty years (Lundby et al., 2012), the prevalence of blood doping can largely be attributed to the relative simplicity of the methods used, the significant enhancement to athletic performance (Lundby et al., 2012) and the relative ease with which one can avoid detection by anti-doping efforts (Ashenden et al., 2011).

By the early 1990s, Erythropoietin (EPO) has become the preferred performance enhancing drug among endurance athletes, due to its logistical ease compared to blood doping and its low detectability, being a naturally occurring hormone (Sawka et al., 1996).

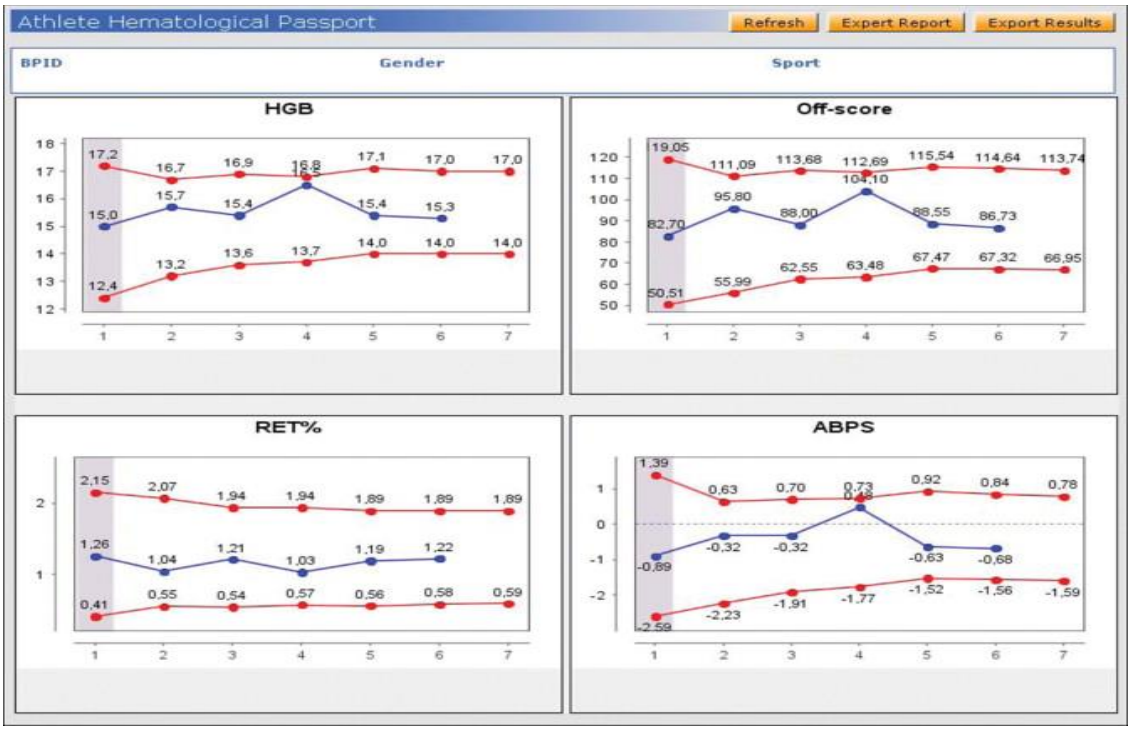
Regrettably, EPO remains one of the most abused substances in sports to enhance endurance performance (Debeljak et al., 2012). EPO use led to numerous scandals, notably in the Tour de France. The 2012 United States Anti−Doping Agency (USADA) report on Lance Armstrong, who admitted use of doping after initially denying it. It culminated in a lifetime ban and stripped titles, highlighting the controversial issue of EPO in sports (Atkinson et al., 2020). EPO serves as the primary hormonal controller of red blood cell production (Bieber et al., 2001). It is a glycoprotein hormone, primarily produced in the kidney (Bieber et al., 2001). It is released by renal cortical interstitial cells in reaction to tissue oxygen deficiency and promotes the production of haemoglobin (Bieber et al., 2001). Synthetic EPO known as recombinant EPO, has emerged as the predominant medication for managing anemia deriving from various sources (Jelkman et al., 2013).

Detection of the recombinant human erythropoietin (rHuEpo) is troublesome since it is structurally alike endogenous erythropoietin and rapidly disappears from circulation (Durussel et al., 2016). Despite improvements in detection methods over the years,

significant challenges persist, particularly in identifying microdoses of blood doping substances (Ashenden et al., 2011).

The race between doping athletes and scientists has been ongoing throughout the history of competitive sports (Azzazy et al., 2007). In 2009, WADA adopted the Athlete Biological Passport (ABP) as a new method to detect blood doping (WADA, 2023). The ABP longitudinally monitors specific haematological variables, Red blood cells (RBC), Haematocrit (HCT), Haemoglobin concentration (HGB), Mean Corpuscular Haemoglobin (MCH), Mean Corpuscular Haemoglobin Content (MCHC), Mean Corpuscular Volume (MCV), Number of Reticulocytes (RET#), Reticulocytes percentage (RET%) and Off-score (Hgb x 10 – 60(√RET%)) which serves as an indicator of blood doping (WADA, 2023). The ABP programme consists of two modules haematological and steroid (WADA, 2023). The haematological module (Figure 1), aims to detect enhancements in oxygen transport, encompassing the usage of erythropoiesis-stimulating agents (ESAs) such as rHuEpo and other forms of blood manipulation or transfusion (WADA, 2023).

Figure 1. *Illustration of the haematological module of the ABP (Schumacher et al., 2014).*



The ABP uses a Bayesian method for the evaluation of the doping based on several variables and/or factors which can be used as evidence of doping (Sottas et al., 2008, 2010). The ABP estimates the probability of blood doping based on previous individual test history and heterogeneous factors known to influence blood parameters such as sex, ethnicity, altitude exposure, age, and sporting discipline (Sottas et al., 2008, 2010). Although it is noteworthy in a sporting context, strenuous exercise can make changes in blood parameters (Ashenden et al., 2004).

The effectiveness of the ABP in detecting rHuEpo has been a topic of immense debate within the anti-doping community. Research conducted by Mørkeberg et al. and Pottgiesser et al., in 2011 highlighted that the ABP when analysing haemoglobin mass

(Hbmass) and RET%, showed superior sensitivity in detecting the highest dosages of transfused blood. However, the OFF-score or "stimulation index" which is calculated with haemoglobin level and reticulocytes percentage level using the formula: (Hgb x 10 – 60(√RET%) (Zorzoli et al., 2011), demonstrated equal or superior sensitivities at detecting lower dosages of doping (Mørkeberg et al. and Pottgiesser et al., 2011). Further investigations by Bornø et al., (2010) and Ashenden et al., (2011) revealed that athletes could enhance their performance significantly through microdoses of rHuEpo without triggering the ABP, highlighting the system's limitations. These findings support an urgent need for the development of more advanced and effective methods to detect blood doping. The scientific validity and interpretation of ABP data has been the subject of disagreement, particularly among athletes. The Court of Arbitration for Sport (CAS) in 2016 noted that athletes sometimes challenge the findings of their ABP by arguing that natural physiological variability, such as the effects of altitude, diet, hydration, and illness, could mimic the haematological markers of doping. In the case of Kristina Ugarova, significant deviations in HGB, RET%, and OFF-score suggested blood doping. However, Ugarova disagreed with these findings, attributing the anomalies to natural physiological responses to high-altitude training, thereby questioning the ABP's accuracy. While Ugarova's case was ultimately unsuccessful, it demonstrates that athletes may use physiological adaptations to altitude to justify haematological changes.

A previous study from (Haile et al., 2019) compared the performance benefits of rHuEpo in Kenyan runners training at moderate altitude to Caucasians training at sea level. The findings indicated similar relative improvements in performance for both groups, despite different baseline blood parameters. This finding suggests that the benefits from increased haemoglobin were not fully realized in performance gains, highlighting the complexity of detecting doping in athletes training at moderate altitudes. Notably, Bejder et al., (2021) demonstrated that the ABP's sensitivity to rHuEpo detection was higher at altitude than at sea level for most tested haematological parameters, emphasizing their importance to detect rHuEpo abuse.

In the same year the ABP was implemented, Varlet-Marie et al., (2009) utilized SAGE (i.e., serial analysis of gene expression method) to identify differentially expressed genes associated with rHuEpo administration. This research demonstrated that gene expression patterns could serve as a novel means to detect the presence of rHuEpo, offering a promising direction for future anti-doping efforts. They also point toward the potential of advanced methods such as gene expression analysis, to enhance the efficacy of doping detection in sports.

The impact of rHuEpo on blood parameters and athletic performance was assessed in a study by Durussel et al., (2013). Study observed significant improvements in performance, aerobic capacity ($\dot{V}O_2$ max), and Hgb mass over a 4-week rHuEpo injection regimen among 39 endurance-trained males. Subsequent transcriptomic analysis revealed significant changes in gene expression during and up to four weeks post rHuEpo administration, highlighting the potential of gene biomarkers to enhance anti-doping strategies (Durussel et al., 2016).

As markers for rHuEpo doping are developed, it is imperative that potential co-founders of a transcriptomic test are investigated. For example, results from studies by Buttner et al., (2007) and Connolly et al., (2004) showed that exercise significantly influences gene

expression profile. Other factors have been theorised that could significantly alter gene expression profiles, such as exercise induce haemolysis (Lippi et al., 2019), injury immobility (Schumacher et al., 2008) and residence at altitude (Durussel et al., 2014). Furthermore, Wang et al., (2017) recruited 14 healthy, endurance-trained individuals and administered weekly microdoses of rHuEpo for seven weeks. In addition to pilot studies investigating the effects of altitude exposure and exercise on gene expression. As hypothesised, gene expression was significantly altered following rHuEpo use, exercise and altitude, indicating the potential to develop transcriptomic signatures for these co-founders (Wang et al., 2017). This work was furthered by Sutehall et al., (2022) who identified a transcriptomic signature of acute altitude exposure in whole blood and peripheral blood mononuclear cells.

While significant work has been developed to address co-founders of a transcriptomic test for rHuEpo abuse, additional research is needed to develop normative reference ranges, to account for individual variations in gene expression.

Aims

The primary aim of the present study was to establish and exploit a biobank of samples from non-doping individuals. Through the collection and analysis of blood and urine samples, the limitations of the current ABP haematological method will be adressed.

Secondary aim of this research was to create transcriptomic reference ranges of doping-free individuals to be used in the future as a comparison with athlete samples.

Hypothesis

Hypothesis 1: The establishment of a non-doping cohort through the collection and analysis of blood samples will reveal specific hematological and genetic markers that are consistent among non-doped individuals. It will distinguish them from those who may have engaged in doping practices.

Hypothesis 2: Creation of the cohort group for reference ranges, will help in future to make transcriptomic profiles, which will account for individual variations related to population, sex, age, and physiological conditions, thereby reducing false positives and negatives in doping tests.

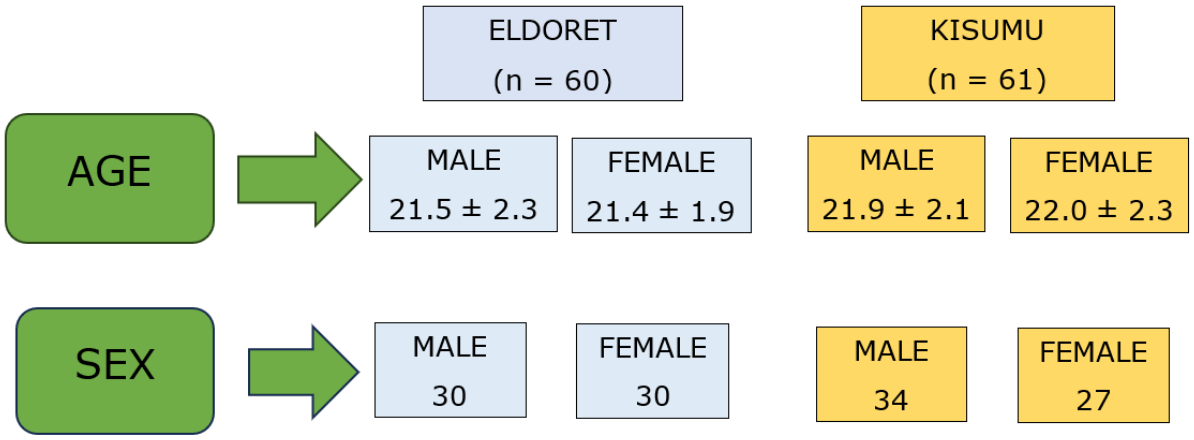
Methods

Participants

One hundred and twenty-one university students were recruited for this study, 61 from Eldoret, at moderate altitude (~2100 - 2700 m above sea level) and 61 from Kisumu (~1000 m above sea level). One hundred and eight participants fulfilled study requirements. Age and sex distribution for the recruited group is available in Table 1. None of the participants were elite athletes, but engaged in a range of sporting activities, including athletics, soccer, rugby, tennis, basketball, and badminton. The inclusion criteria required participants to be healthy students or athletes (self-reported good health) of both sexes, with an equal distribution of 50% male and 50% female. Participants found to have any health conditions (Anisocytosis, Microcytosis, Iron Deficiency, Anemia, PLT clump, fragments, NRBC, Atypical Lymph, Eosinophilia, Lymphocytosis, Abnormal Lymph) identified through haematological screening, if not influencing the ABP haematological variables cut off and ABPS results will not be excluded from the study. Participants were excluded from the study if they missed any of the four scheduled sample collection visits, or if they were found to be using drugs prohibited by the WADA code, such as erythropoietin or steroids. Participants in this study were non-professional athletes from a wide range of sports modalities, therefore weight, height, Body Mass Index as well as VO_2max were not relevant for the study and were not measured. Information about the menstrual cycle was not collected from female participants. As the primary objective of the study was to establish a biobank of the clean samples for future genetic analysis, information about menstrual cycle is not relevant for the study neither for the future analysis.

Potential participants were approached by members of the research staff, who visited the local sporting facilities and hostels in the colleges and universities in Eldoret and Kisumu. Those who met the study's eligibility criteria received full written (Appendix 1) and verbal overview of the study's objectives, procedures and intended outcome. Their level of comprehension was assessed through a structured interview process. It ensured participants fully understood the study's requirements and procedures. After confirming their comprehension, written informed consent was collected from each participant.

Table 1. Age and sex distribution of the participants recruited for the study.



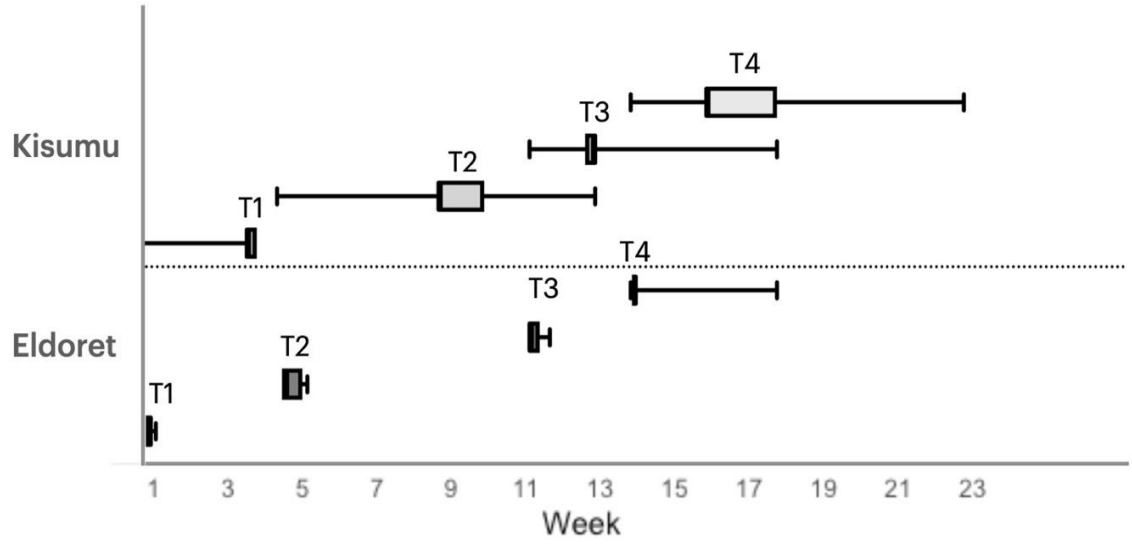
Age data is presented as mean ± SD.

Study design

Participants provided blood and urine samples four times over a 6-month period. For participants from Eldoret, the second (T2), third (T3), and fourth (T4) samples were collected at median (IQR) intervals of 29 (4), 77 (2), and 98 (0.25) days, respectively, after the first sample (T1) was collected. For participants from Kisumu, the second (T2), third (T3) and fourth samples (T4) were collected at median (IQR) of 59 (9), 90 (2) and 112 (14) days after the initial sample (T1) was collected (Figure 2). Approximately 10mL of urine was collected in a sterile urine bottle and stored at -20°C freezer for further analysis. A 3 mL whole blood sample from an antecubital vein was collected into a Tempus™ Blood RNA Tube (Thermo Fisher Scientific, Massachusetts, USA). It was stored at -80°C freezer within 12 hours from the time of collection to preserve RNA for subsequent extraction, while a 3 mL whole blood sample was also collected into K_2EDTA tube for haematological analysis.

Sample analysis

Figure 2. *Distribution of samples collected over the duration of the study.*



Box indicates the interquartile range, with the horizontal line indicating the range.

Blood samples collected for haematological analysis were transported to the laboratory, in a cold box, and processed within 12 hours from the time of collection as WADA protocol for the blood transportation and storage demands (WADA, 2021). Blood samples were analysed for haematological parameters assessed by the ABP, using fluorescence flow cytometry (Mindray Auto Haematology Analyzer BC-6800, Shenzhen, China) at the Haematology Laboratory within Moi Teaching & Referral Hospital (MTRH), Eldoret, Kenya. The Mindray Auto Haematology Analyzer is designed for various blood analysis tasks in clinical settings (Wang et al., 2019). This includes counting blood cells, classifying types of white blood cells, and measuring haemoglobin levels (Wang et al., 2019). It operates on the Coulter principle, which assesses the number and size distribution of white blood cells, basophils, red blood cells, and platelets. Haemoglobin levels are determined through colorimetric analysis (Wang et al., 2019). Additionally, it employs semiconductor laser flow cytometry for a comprehensive four-category statistical analysis of white blood cells (Wang et al., 2019). Based on these measurements, the device further computes other relevant parameters (Wang et al., 2019).

Specifically, RBC, HCT, HGB, MCH, MCHC, MCV, RET#, and RET% were determined. The Off-score was calculated for each sample using the formula: Hgb x 10 – 60(√RET%) (WADA,2014). The blood samples collected for RNA analysis were stored in the biobank on −80 Celsius in the preparation for gene expression analysis, results are expected in July 2024.

Adaptive Bayesian model

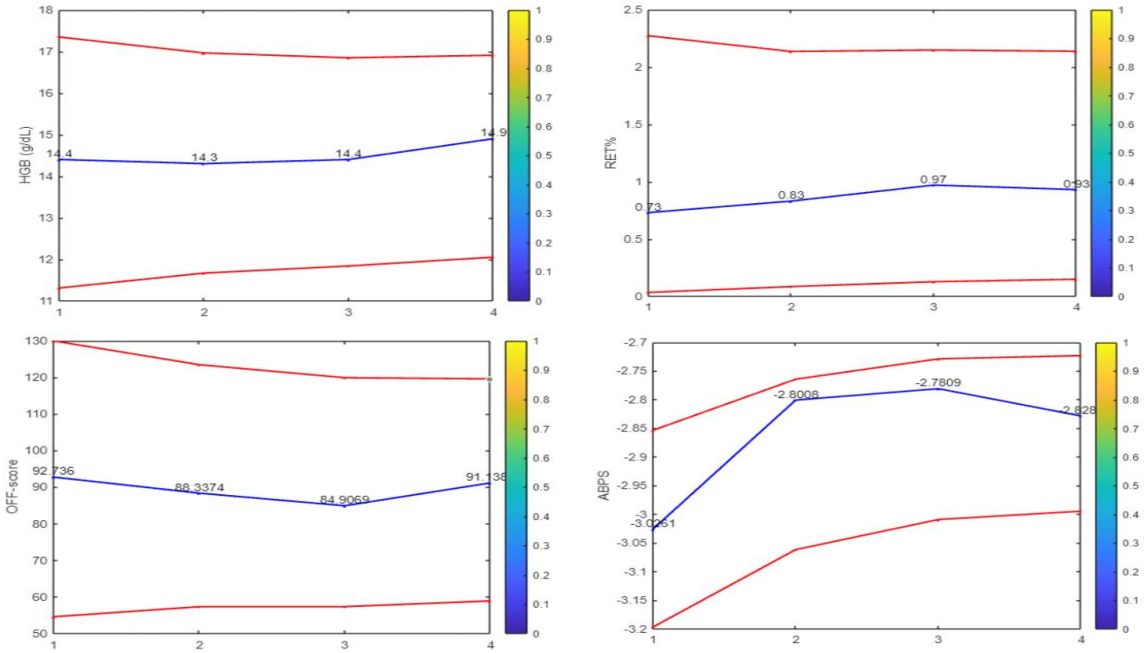
In the present study, drug-free status of the participants was determined using Bayesian networks (integrating prior information about doping prevalence with the results of current tests), applied to derive population reference ranges and individual test result history (Sottas et al., 2010). By inputting variables such as mean subject variance, between-subject variance, gender, gender tabulation, sequence, and baseline data into MATLAB (inhouse software developed), I attempted to replicate some of the key functionality of the official WADA ABP software. Specifically, MATLAB (version 6.1.0. with Statistics Toolbox version 3.0) was used to calculate the upper and lower individual limits by applying the adaptive Bayesian model to HGB, RET%, OFF-score, and the ABP score. ABP score (ABPS) was determined in R (R Studio, Version 1.2.5042, ABPS package, Vienna, Austria). This test can be repeated by other package users (Schütz et al., 2018). Haematological variables that were outside the reference dataset within the ABPS package were modified to the maximum or minimum acceptable value (Schütz et al., 2018). The initial limits are based on African population epidemiology and subsequent limits are adapted to the collected data including mean subject variance (me), between-subject variance (bsvar), sex (s), sex tabulation (st), sequence (seq), and baseline data (bd) (Sottas et al.,2006). Sequences within limits were considered to be drug-free (Schütz et al., 2018). On the other hand, sequences with significant deviations and out of limits were considered suspicious (Table 2). An example of a 'normal' biological profile, reflecting no irregularities or deviations from established physiological norms is presented in Figure 3.

Table 2. *Number of participants who fulfilled study requirements, who exceeded the upper or lower limits of the ABP-style Bayesian model, for each haematological variable.*

	Eldoret		**Kisumu**	
	Male (n=27)	Female (n=25)	Male (n=33)	Female (n=23)
HGB	0	0	0	1
RET%	0	0	0	0
OFF-score	0	0	1	1
ABPS	15	13	18	14

Data are presented as number of participants who exceeded the upper or lower limit of the ABP-style Bayesian model, for each haematological variable (HGB, RET%, OFF-score) and ABPS values. Model was created in MATLAB.

Figure 3. *Normal biological profile in the ABP-style Bayesian model.*



ABP-style Bayesian model, created in MATLAB. Blue lines indicate the participant's haematological variables and red lines indicate the upper and lower limits as determined by the Bayesian statistics. Y axis shows haematological variables in each figure from left to right (HGB, RET%, OFF-score) and ABPS values, X axis shows results of each of four collection points.

Statistical analysis

To assess the normality of the distribution, the Shapiro-Wilk test was applied. The Kruskal-Wallis test was used to evaluate differences in age across locations and sexes, since this test is designed to examine whether there are differences among multiple groups (Walters et al., 2021). This research has four groups (Kisumu females, Kisumu males, Eldoret female, and Eldoret males), their age are independent and non-parametric continuous data. Furthermore, the Chi-squared test was applied to examine the variations in the distribution of male and female participants and the prevalence of health conditions across the two locations. As for distribution of sex and health conditions, data are categorical and non-parametric and reasonably large (Walters et al., 2021), that is why Chi-squared test was used. To compare haematological data between various locations and sexes, the Mann-Whitney U test was used. Statistical significance was established at a p-value ≤ 0.05. The choice of a one-sided test was driven by its suitability where the research goal is to confirm a specific condition (Walters et al., 2021) in this case, the non-use of doping. This approach efficiently allocates statistical power towards demonstrating the lack of doping, thus directly supporting the study's objective to confirm the groups are doping-free. All analyses were conducted using the R programming language (R Studio, Version 1.2.5042, R Foundation for Statistical Computing, Vienna, Austria). Haematological data are presented as medians and interquartile ranges. Power calculations for this study were not done. In the context of this research, the primary objective is to establish reference values for a forthcoming diagnostic method. Given the exploratory nature of this study, the determination of an effect size, is not directly applicable. This limitation arises from the necessity to first

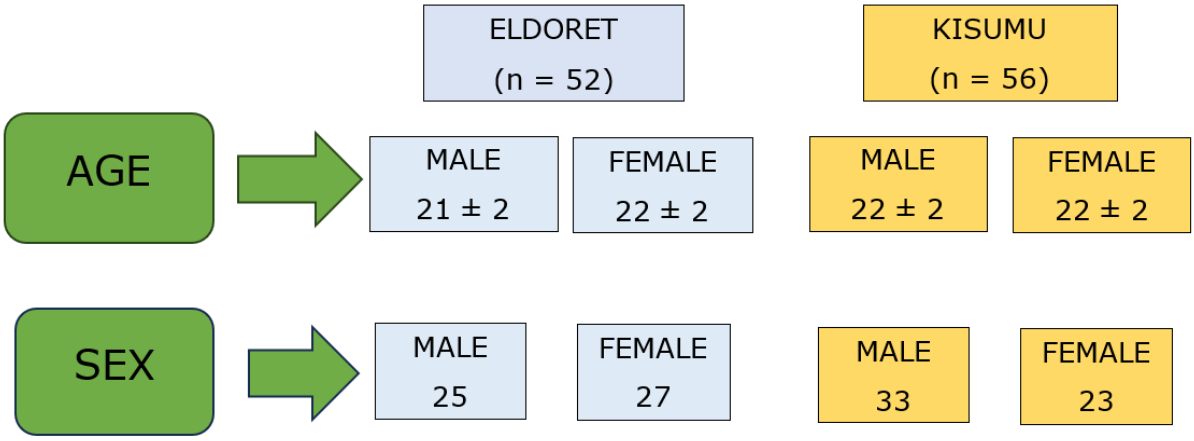
identify and understand the "normal" reference data before specific genes can be selected for the test. Consequently, power calculations rely on effect sizes to estimate the necessary sample size to detect a statistically significant effect, are not feasible at this stage. Notably, existing literature has documented variations from baseline levels for certain genes, providing effect sizes that could potentially inform power analyses. However, the application of these findings to the present study is constrained by the fact that the genes contributing to the diagnostic test have yet to be determined.

Results

Recruitment

Of the 121 recruited university students, 12 participants missed at least one of the four scheduled visits (due to personal engagements) and one participant's samples were compromised, and therefore were excluded from the final analysis. Leaving a total of 108 participants who fulfilled study requirements (Eldoret: 25 females, 27 males), (Kisumu: 23 females and 33 males). Participants found to have any health conditions (Anisocytosis, Microcytosis, Iron Deficiency, Anemia, PLT Clump, fragments, NRBC, Atypical Lymph, Eosinophilia, Lymphocytosis, Abnormal Lymph) identified through haematological screening did not significantly influence the ABP haematological variables cut off and ABPS results and were not excluded from the study. There were no statistically significant difference in age and sex between location (Table 3).

Table 3. *Demographics of participants who completed all four sample collections.*



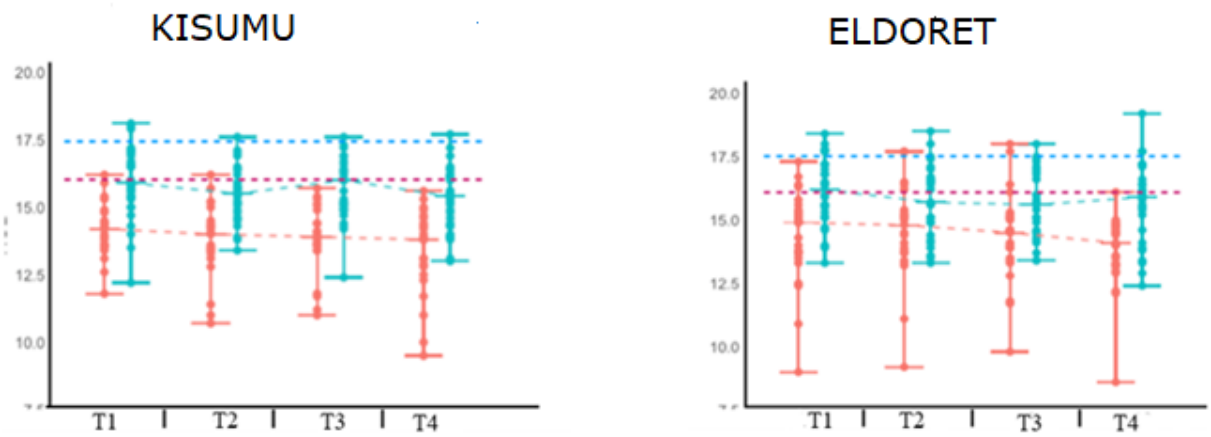
Age data is presented as mean ± SD. There were no significant differences in age and sex.

Haematological Analysis

Haemoglobin concentration

Men had a higher median HGB than women at first collection period (T1) (16.4 g/dL vs 14.9 g/dL in Eldoret and 15.9 g/dL vs 14.2 g/dL in Kisumu, $p<0.001$, respectively). Second collection period (T2) (15.7 g/dL vs 14.8 g/dL in Eldoret and 15.4 g/dL vs 14.0 g/dL in Kisumu, $p<0.05$). Third collection period (T3) (15.7 g/dL vs 14.5 g/dL in Eldoret and 15.6 g/dL vs 13.9 g/dL in Kisumu, $p<0.05$. Fourth collection period (T4) (15.9 g/dL vs 14.1 g/dL in Eldoret and 15.3 g/dL vs 13.8 g/dL in Kisumu, $p<0.001$) (Figure 4, Table 4).

Figure 4. *Haemoglobin grams per decilitre (HGB g/dl) over the four samples period.*



Data are presented as median and range. Dashed horizontal lines indicate male (in blue) and female (in red) cut-offs. X axis stands for collection period (T1, T2, T3, T4). Y axis stands for HGB (g/dL) values.

Table 4. *HGB (g/dL) across the timepoints collected.*

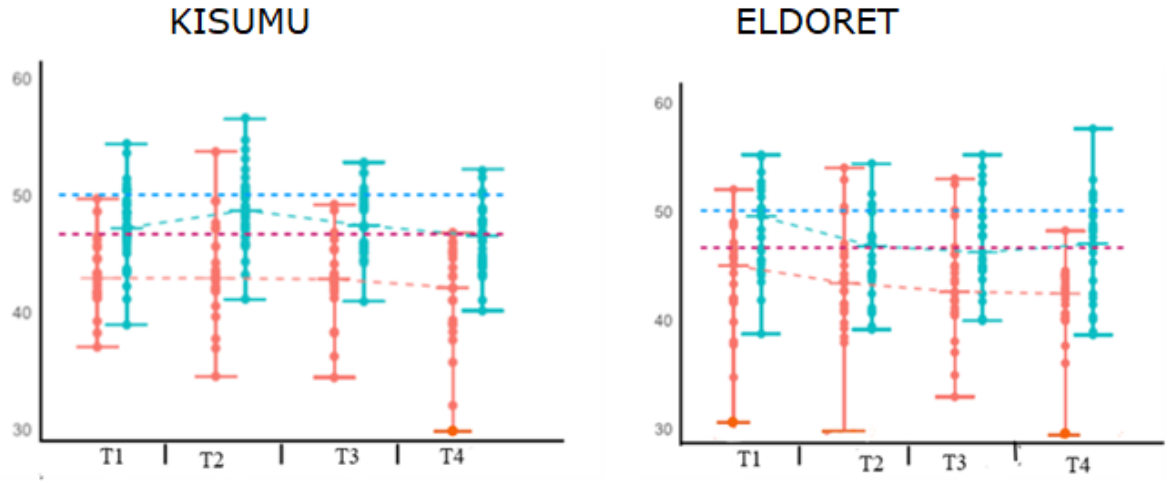
		T1	T2	T3	T4
Eldoret	Male	16.4 (15.2-17.1)*	15.7 (14.6-16.6)*	15.7 (14.5-16.8)*	15.9 (14.3-16.5)*
	Female	14.9 (13.7-15.4)	14.8 (13.7-15.3)	14.5 (13.5-15.1)	14.1 (13.2-14.8)
Kisumu	Male	15.9 (15.4-16.7)*	15.4 (14.9-15.9)*	15.6 (15.0-16.4)*	15.3 (14.3-16.1)*
	Female	14.2 (13.6-14.6)	14.0 (13.3-14.45)	13.9 (13.5-14.6)	13.8 (12.6-14.5)

*Data are presented as median (range). * Indicates significant difference between sex within location($p<0.05$).*

Haematocrit

Men had higher median HCT than women in Kisumu and Eldoret at T1 (49.8% vs 45.0% in Eldoret and 47.2% vs 42.9% in Kisumu, $p<0.001$, respectively). T2 (46.8% vs 43.4% in Eldoret and 48.4% vs 42.9% in Kisumu, $p<0.05$). T3 (47.3% vs 42.6 % in Eldoret and 47.3% vs 42.8% in Kisumu, $p<0.05$). T4 (47.0% vs 42.4 % in Eldoret and 46.2% vs 42.1% in Kisumu, $p<0.001$ (Figure 5, Table 5). Similarly, the median of HCT in male participants from Kisumu were higher than those from Eldoret at T2 (48.4% vs 46.8%, $p<0.01$).

Figure 5. *Haematocrit percentage (HCT %) over the four samples period.*



Data are presented as median and range. Dashed horizontal lines indicate male (in blue) and female (in red) cut-offs. X axis stands for collection period (T1, T2, T3, T4). Y axis stands for HCT (%) values.

Table 5. *HCT (%) across the timepoints collected.*

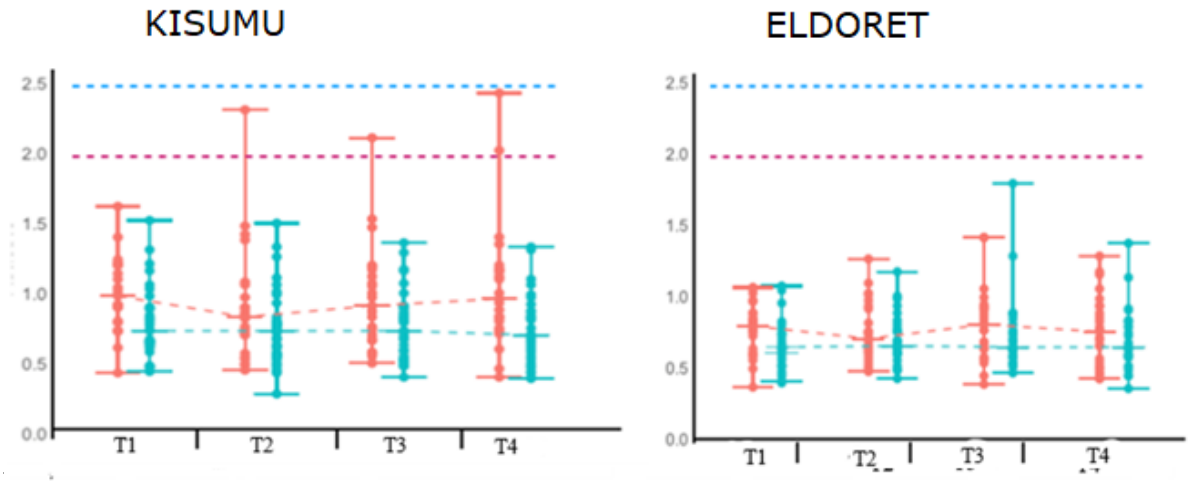
		T1	T2	T3	T4
Eldoret	Male	49.8 (45.3-50.8)*	46.8 (43.8-48.8)*§	47.3 (44.6-49.5)*	47.0 (43.0-49.6)*
	Female	45.0 (41.6-46.6)	43.4 (41.0-46.1)	42.6 (40.7-46.0)	42.4 (40.2-43.7)
Kisumu	Male	47.2 (45.8-49.3)*	48.4 (47-50.3)*	47.3 (45.7-49.8)*	46.2 (44.2-47.9)*
	Female	42.9 (41.6-45.0)	42.9 (41.8-46.3)	42.8 (41.8- 45.3)	42.1 (38.9-45.0)

Data are presented as median (range). * Indicates significant difference between sex within location and § indicates significant difference between location for the same sex ($p<0.05$).

Reticulocyte

Female participants in Kisumu had significantly higher median RET % values compared to females in Eldoret at T1 (0.98% vs 0.79%, $p<0.05$) and T4 (0.96% vs 0.75%, $p<0.05$) (Figure 5, Table 6). Additionally, the analysis also identified that females had higher reticulocytes than males at T1 in Eldoret (0.79% vs 0.60%, $p<0.05$) and in Kisumu (0.98% vs 0.67, $p<0.05$) (Figure 6, Table 6).

Figure 6. *Reticulocytes percentage (RET%) over the four samples period.*



Reticulocytes percentage (RET%) over the four samples collected are presented as median and range. Dashed horizontal lines indicate male (in blue) and female (in red) cut-offs. X axis stands for collection period (T1, T2, T3, T4). Y axis for RET% values.

Table 6. *RET% (%) across the timepoints collected.*

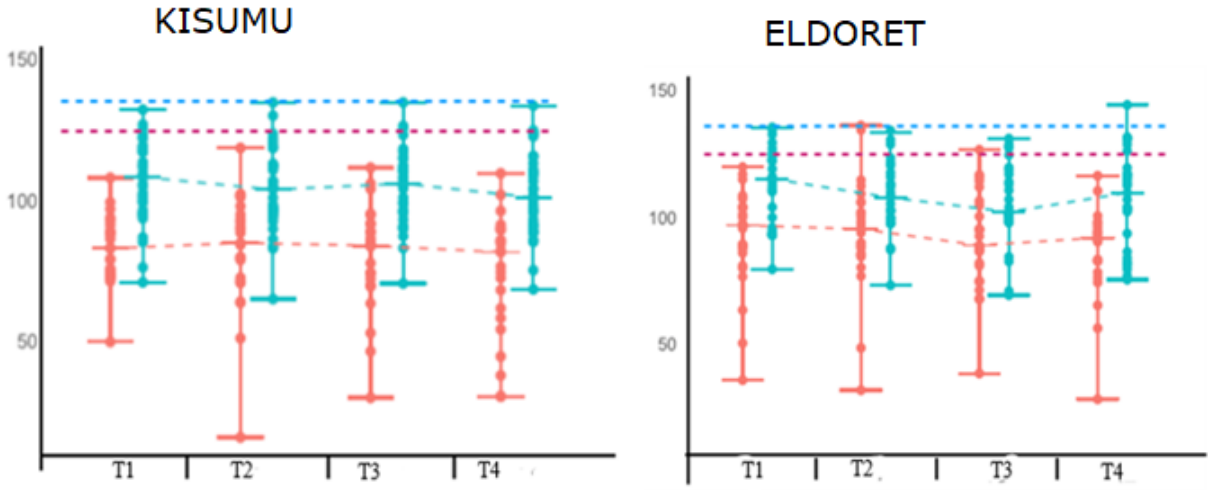
		T1	T2	T3	T4
Eldoret	Male	0.60 (0.53-0.73)*	0.68 (0.60-0.78)	0.63 (0.57-0.78)	0.63 (0.57-0.80)
	Female	0.79 (0.63-0.88)§	0.70 (0.61-0.91)	0.80 (0.63-0.92)	0.75 (0.53-0.93)§
Kisumu	Male	0.67 (0.63-0.89)*	0.75 (0.57-0.83)	0.79 (0.55-0.89)	0.69 (0.55-0.87)
	Female	0.98 (0.79-1.20)	0.83 (0.63-1.01)	0.91 (0.69-1.14)	0.96 (0.77-1.16)

*Data are presented as median (range). * Indicates significant difference between sex within location and § indicates significant difference between location for the same sex ($p<0.05$).*

OFF-score

Men had higher OFF-score than women at T1 (115.7 vs 96.7 in Eldoret and 108.0 vs 82.7 in Kisumu, $p<0.001$). T2 (107.5 vs 95.1 in Eldoret and 100.7 vs 84.7 in Kisumu, $p<0.001$) (Figure 6, Table 7). T1 (96.7 vs 82.7, $p<0.01$), T2 (95.1 vs 84.7, $p<0.05$) and T4 (91.7 vs 81.4, $p<0.05$) women at Eldoret demonstrated higher OFF-score of median than women in Kisumu (Figure 7, Table 7).

Figure 7. *OFF-score over the four samples period.*



Data are presented as median and range. Dashed horizontal lines indicate male participants (in blue) and female participants (in red) cut-offs. X axis stands for collection period (T1, T2, T3, T4). Y axis stands for OFF-score values.

Table 7. *OFF-score across the time points collected.*

		T1	T2	T3	T4
Eldoret	Male	115.7 (97.7-123.6)*	107.5 (98.3-116.5)*	103.7 (98.7-117.5)*	112.2 (89.8-118.0)*
	Female	96.7 (85.7-103.7)	95.1 (85.4-105.5)	88.7 (80.7-104.4)	91.7 (82.5-95.7)
Kisumu	Male	108.0 (96.3-117.9)*	100.7 (95.2-112.1)*	105.7 (95.7-112.0)*	108.3 (91.0-110.8)*
	Female	82.7 (73.7-90.0)§	84.7 (71.8-93.9) §	83.7 (70.9-91.5)	81.4 (59.8-89.4)§

*Data are presented as median(range). * Indicates significant difference between sex within location and § indicates significant difference between location for the same sex ($p<0.05$).*

Sample Storage and Future Analysis

As part of our comprehensive research protocol, we have ensured that all collected samples are securely stored for future analysis. RNA will be extracted from Tempus™ Blood RNA Tube according to manufacturer's protocol. Quantity and quality will be assessed using Nanodrop 2000 (Thermo Fisher Scientific, Massachusetts, USA) and 2100 Bioanalyzer (Agilent Technologies, California, USA). Samples that pass the quality controls parameters (260/280 ratio of ~2.0 and RIN value ~10) will be analysed using an RNA-seq platform (DNBSEQ-G400RS, MGI Tech, Shenzhen, China) to allow comparison with previously published research (Wang et al., 2021). This analysis will allow to determine the individual variation in transcriptomic expression of key genes previously identified as indicators of rHuEpo doping or co-founders (e.g., Durussel et al., 2016, Wang et al., 2017 and Sutehall et al., 2022).

Analysis of Urine Samples

The collection of urine samples in this study was primarily conducted to complement our haematology findings. These urine samples have been securely stored in the biobank to assure the integrity of the samples (stable temperature on -80 Celsius) and ready for analysis at a later stage to confirm the absence of doping. As such, urine samples are being stored for future analysis of urinary EPO using the standard WADA method (WADA, 2010).

Discussion

The purpose of this study was to establish a cohort of non-doping participants whose blood sample data can be used to determine the individual variation of genes used to create a transcriptomic test. Both cohorts recruited to this study (Eldoret and Kisumu) reside at low (~1000 m above sea level) or moderate altitude (~2100 – 2700 m above sea level) and therefore, no sea level participants were recruited. Since determining the individual variation of specific transcriptomic markers is the purpose of this study, both cohorts can be combined, regardless of altitude of residence.

In order to determine the clean status of the participants of this study, we created an ABP-style model in MATLAB. Using this model, which replicates the Bayesian statistics used in the ABP, we are able to verify the clean status of each participant and use the data generated in the transcriptomic analysis with confidence. As highlighted in Table 2, only three of 108 participants exceeded the haematological limits, over the duration of this study, suggesting the majority of the participants recruited in this study are currently not blood doping. Those participants who had samples that exceeded the Bayesian limits, will have their profiles reviewed, and if it confirmed that the profile is suspicious, they will be removed from the transcriptomic analysis. Our findings are in agreement with the study conducted by Pitsiladis et al., (2006), who observed Kenyan athletes residing at high altitudes exhibit HGB and HCT concentrations in the normal to high range compared to lowlanders. This alignment further validates our results, reinforcing the conclusion that the majority of this study group consists of participants who do not engage in doping practices.

The efficacy of the ABP is subject to various influencing factors, including ethnicity, sex, and the precision of measurement equipment, as highlighted by WADA in 2019. Notably, the impact of altitude exposure remains unstandardized within the ABP framework, posing significant challenges to its integrity (Sottas et al., 2008). Exercise induces significant changes in blood volume, with endurance exercise leading to an expansion of both plasma and erythrocyte volumes, suggesting a universal response to endurance training that is consistent across age and gender (Sawka et al., 2000). However, this physiological adaptation introduces analytical variation that poses significant challenge for the haematological module of the ABP.

The use of substances like rHuEpo by athletes can further complicate the detection of doping by modifying haematological variables (Sanchiz et al., 2009). Furthermore, the use of microdoses of rHuEpo, which often does not trigger the ABP detection thresholds, adds another layer of complexity to the doping detection challenge (Ashenden et al., 2011). In study by Ashenden et al., (2011) involving ten subjects administered microdoses of rHuEpo over up to 12 weeks reveals significant findings. Despite a notable 10% increase in Hgb mass, similar to approximately two bags of reinfused blood the ABP software did not flag any of the subjects as suspicious of doping while receiving rHuEpo treatment. This outcome illustrates the potential for athletes to use rHuEpo without triggering abnormal changes in the blood variables that the ABP currently monitors, further complicating the efforts to detect doping practices effectively (Ashenden et al., 2011). A prominent illustration of these challenges is the controversial case of Italian cyclist Franco Pellizotti, who was accused of doping based on his ABP data, resulting in his initial suspension (Sanchis et al., 2011). However, the Italian anti-doping tribunal later considered the evidence insufficient, a verdict subsequently reversed by a two-year suspension from the Court of Arbitration for Sport, following an appeal by the Union

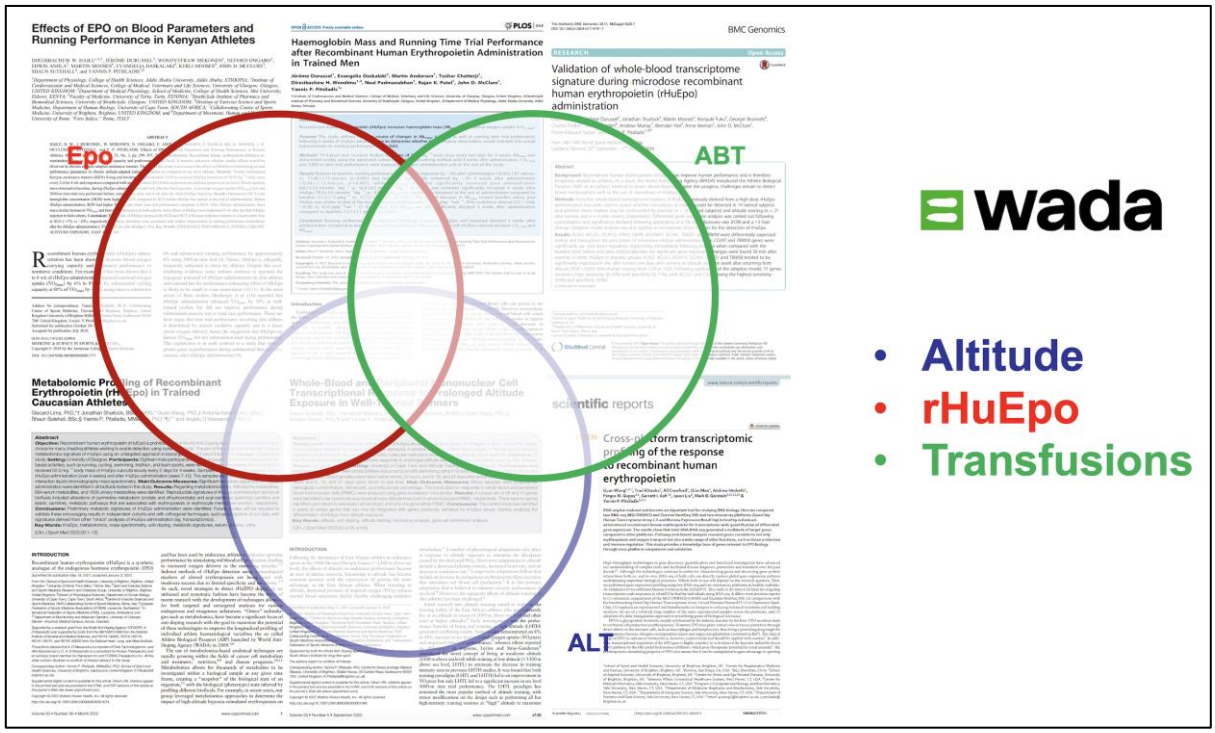
Cyclists Internationale (UCI). This case underscores the critical need for standardization within the ABP, particularly in how altitude training's effects on blood profiling are interpreted, as Pellizotti's case vividly demonstrates (Sanchis et al., 2011).

Our findings support existing literature (Sharp et al., 2002; Gore et al., 2003) indicating significant gender differences in HGB (Figure 4, Table 4), HCT (Figure 5, Table 5), OFF score (Figure 7, Table 7) with male participants exhibiting higher values than female participants. This aligns with the physiological understanding that males typically have higher HGB and HCT levels due to hormonal influences. The higher production of androgens in men, compared with females, enhances erythropoiesis through direct stimulation in the bone marrow and by increasing erythropoietin production in the kidneys (Murphy et al., 2014). Conversely, the higher of oestrogen in women inhibits these processes (Murphy et al., 2014). It is therefore expected that females haematological variables will be significantly lower than men. However, females from Kisumu in our study, had significantly higher RET% compared to males in Kisumu and both genders in Eldoret (Figure 6, Table 6). This gender difference in RET% distribution has also been reported elsewhere (Lombardi et al., 2013), where female athletes to exhibit higher RET% values across multiple seasons. This discrepancy highlights how a range of factors influence RET% in athletes, including training season, sports discipline, and individual biological differences. Given the significant differences in RET% observed between female participants in Kisumu and Eldoret, altitude may serve as a contributing factor (Lombardi et al., 2013). However, without controlling for variables such as diet, altitude history prior to sampling, and menstrual cycle phases, pinpointing the exact cause of these differences remains challenging. This limitation underscores the complexity of haematological analysis in athletes and the potential impact of various uncontrolled environmental and physiological factors.

Future Transcriptomic analysis

In the current phase of our study, we have laid the groundwork for an extensive analysis of 'omics' profiles, with a particular emphasis on transcriptomic variations. Specifically, plan is to sequence all samples (n>1200) from the previously funded/associated WADA studies with particular focus on altitude, rHuEpo and blood transfusion (Figure 8).

Figure 8. *The planned sequencing analysis.*



The planned sequencing analysis of all samples (n>1200), from previously funded/associated WADA studies and comparison between altitude, rHuEpo and blood transfusion illustrated with a Venn diagram.

While this analysis is scheduled to be conducted March to July 2024, this section outlines anticipated approach. Hypotheses suggest that differential gene expression responses will be evident, reflecting the participants' adaptations to their respective environments. Particular attention will be paid to how living at different altitudes might affect gene expression. Additionally, it is planned to explore potential gender-specific genetic expression responses.

In this upcoming analysis, it is planned to identify specific genes and pathways involved, thereby enriching our understanding of transcriptomic adaptations to various external and internal factors. The detailed analysis will also explore potential biomarkers for diverse applications, ranging from medical research to athletic performance optimization in varying environmental conditions.

Conclusions

The findings from this study will significantly enhance our understanding of the molecular changes induced by rHuEpo and its detection. The identification of a unique molecular signature associated with rHuEpo use marks as a crucial advancement in anti-doping research. This signature will provide a more nuanced and sophisticated method of detecting rHuEpo, which will surpass the current limitations faced by standard detection methods of the ABP.

The variability in HCT and RET%, particularly among female participants from Kisumu, demonstrates the complexity of interpreting haematological markers in the context of doping. This variability highlights the necessity for more nuanced approaches like omics profiling, which can differentiate between the effects of environmental factors, physiological conditions, and doping use. The implications of this research extend well beyond the confines of sports science. It signifies a critical advancement towards more ethical and effective anti-doping practices, aligning closely with the goals of preserving athlete health and ensuring fair competition on a global scale. The aim is to refine these innovative methods and adapt them for practical application in the dynamic arena of competitive sports.

References:

Ashenden, M. (2004). Contemporary issues in the fight against blood doping in sport. Haematologica, 89(8), 901-903. PMID: 15339669.

Ashenden, M., Gough, C. E., Garnham, A., Gore, C. J., & Sharpe, K. (2011). Current markers of the Athlete Blood Passport do not flag microdose EPO doping. European Journal of Applied Physiology, 111(9), 2307-2314

Atkinson, T. S., & Kahn, M. J. (2020). Blood doping: Then and now. A narrative review of the history, science and efficacy of blood doping in elite sport. Blood Reviews, 39, 100632.

Azzazy, H., & Mansour, M. (2007). Rogue athletes and recombinant DNA technology: Challenges for doping control. The Analyst, 132, 951-957.

Bamberger, M., & Yaeger, D. (1997). Over the edge: Special report. Sports Illustrated, 86, 64.

Bejder, J., Breenfeldt Andersen, A., Bonne, T. C., Linkis, J., Olsen, N. V., Huertas, J. R., & Nordsborg, N. B. (2021). Hematological adaptations and detection of recombinant human erythropoietin combined with chronic hypoxia. Drug Testing and Analysis, 13(2), 360-368.

Bieber, E. (2001). Erythropoietin, the biology of erythropoiesis and epoetin alfa: An overview. The Journal of Reproductive Medicine, 46(5 Suppl), 521–530.

Bornø, A., Aachmann-Andersen, N. J., Munch-Andersen, T., Hulston, C. J., & Lundby, C. (2010). Screening for recombinant human erythropoietin using [Hb], reticulocytes, the OFF (hr score), OFF (z score) and Hb (z score): status of the Blood Passport. European Journal of Applied Physiology, 109(3), 537-543.

Brzeziańska, E., Domańska, D., & Jegier, A. (2014). Gene doping in sport - perspectives and risks. Biology of Sport, 31(4), 251-259.

Büttner, P., Mosig, S., Lechtermann, A., Funke, H., Mooren, F. C., Northoff, H., & Wolfarth, B. (2007). Exercise affects the gene expression profiles of human white blood cells. Journal of Applied Physiology, 102, 26-36.

CAS 2016/O/4463 International Association of Athletics Federations (IAAF) v. All Russia Athletics Federation (ARAF) & Kristina Ugarova. (2016). Retrieved from https://www.tas-cas.org/fileadmin/user_upload/Award_4463__internet_.pdf.

Connolly, P. H., Caiozzo, V. J., Zaldivar, F., & et al. (2004). Effects of exercise on gene expression in human peripheral blood mononuclear cells. Journal of Applied Physiology, 97, 1461-1469.

Debeljak, N., & Sytkowski, A. J. (2012). Erythropoietin and erythropoiesis stimulating agents. Drug Testing and Analysis, 4(11), 805–812.

Durussel, J., Daskalaki, E., Anderson, M., Chatterji, T., Wondimu, D. H., Padmanabhan, N., Patel, R. K., McClure, J. D., & Pitsiladis, Y. P. (2013). Haemoglobin mass and running time trial performance after recombinant human erythropoietin administration in trained men. PLoS ONE, 8(2), e56151.

Durussel, J., Haile, D. W., Mooses, K., Daskalaki, E., Beattie, W., Mooses, M., Mekonen, W., Ongaro, N., Anjila, E., Patel, R. K., Padmanabhan, N., McBride, M. W., McClure, J. D., & Pitsiladis, Y. P. (2016). Blood transcriptional signature of recombinant human erythropoietin administration and implications for antidoping strategies. Physiological Genomics, 48(3), 202-209.

Durussel, J., McClure, J. D., McBride, M. W., & et al. (2014). Validation of blood gene expression profiles post rHuEpo administration. Paper presented at the 61st American College of Sports Medicine Annual Meeting, Orlando, US.

Gore, C. J., Parisotto, R., Ashenden, M. J., Stray-Gundersen, J., Sharpe, K., Hopkins, W., Emslie, K. R., Howe, C., Trout, G. J., Kazlauskas, R., & Hahn, A. G. (2003). Second-generation blood tests to detect erythropoietin abuse by athletes. Haematologica, 88(3), 333-344. PMID: 12651273

Haile, D. W., Durussel, J., Mekonen, W., Ongaro, N., Anjila, E., Mooses, M., Daskalaki, E., Mooses, K., McClure, J. D., Sutehall, S., & Pitsiladis, Y. P. (2019). Effects of EPO on blood parameters and running performance in Kenyan athletes. Medicine & Science in Sports & Exercise, 51(2), 299-307.

Holt, R. I., Erotokritou-Mulligan, I., & Sönksen, P. H. (2009). The history of doping and growth hormone abuse in sport. Growth Hormone & IGF Research, 19(4), 320-326.

Jelkmann, W. (2013). Physiology and pharmacology of erythropoietin. Transfusion Medicine and Hemotherapy, 40(5), 302-309.

Lippi, G., & Guidi, G. (1999). Doping e sport [Doping and sports]. Minerva Medica, 90(9), 345-357. Italian. PMID: 10719440.

Lippi, G., & Sanchis-Gomar, F. (2019). Epidemiological, biological and clinical update on exercise-induced hemolysis. Annals of Translational Medicine, 7(12), 270.

Lombardi, G., Colombini, A., Lanteri, P., & Banfi, G. (2013). Reticulocytes in sports medicine: an update. Advances in Clinical Chemistry, 59, 125-153.

Lundby, C., Robach, P., & Saltin, B. (2012). The evolving science of detection of 'blood doping'. British Journal of Pharmacology, 165(5), 1306-1315.

Mørkeberg, J., Sharpe, K., Belhage, B., Damsgaard, R., Schmidt, W., Prommer, N., Gore, C. J., & Ashenden, M. J. (2011). Detecting autologous blood transfusions: A comparison of three passport approaches and four blood markers. Scandinavian Journal of Medicine & Science in Sports, 21(2), 235–243.

Murphy, W. G. (2014). The sex difference in haemoglobin levels in adults - mechanisms, causes, and consequences. Blood Reviews, 28(2), 41-47.

Pitsiladis, Y. P., Bale, J., Sharp, C., & Noakes, T. (Eds.). (2006). East African Running – Towards a Cross-Disciplinary Perspective. Routledge Taylor & Francis Group. ISBN-10: 0415377889, ISBN-13: 978-0415377881.

Pottgiesser, T., Sottas, P. E., Echteler, T., Robinson, N., Umhau, M., & Schumacher, Y. O. (2011). Detection of autologous blood doping with adaptively evaluated biomarkers of doping: A longitudinal blinded study. Transfusion, 51(8), 1707-1715.

Sanchis-Gomar, F., Martinez-Bello, V. E., Domenech, E., Nascimento, A. L., Pallardo, F. V., & Gomez-Cabrera, M. C., et al. (2009). Effect of intermittent hypoxia on hematological parameters after recombinant human erythropoietin administration. European Journal of Applied Physiology, 107(4), 429-436.

Sanchis-Gomar, F., Martinez-Bello, V. E., Gomez-Cabrera, M. C., & Vina, J. (2011). Current limitations of the Athlete's Biological Passport use in sports. Clinical Chemistry and Laboratory Medicine, 49(9), 1413-1415.

Sawka, M. N., Convertino, V. A., Eichner, E. R., Schneider, S. M., & Young, A. J. (2000). Blood volume: Importance and adaptations to exercise training, environmental stresses, and trauma/sickness. Medicine and Science in Sports and Exercise, 32(2), 332-348.

Sawka, M. N., Joyner, M. J., Miles, D. S., et al. (1996). American College of Sports Medicine position stand: The use of blood doping as an ergogenic aid. Medicine & Science in Sports & Exercise, 28.

Schumacher, Y. O. (2014). The athlete biological passport: Haematology in sports. The Lancet. Haematology, 1(1), e8–e10.

Schumacher, Y. O., Ahlgrim, C., Ruthardt, S., & Pottgiesser, T. (2008). Hemoglobin mass in an elite endurance athlete before, during, and after injury-related immobility. Clinical Journal of Sport Medicine, 18(2), 172-173.

Schütz, F., & Zollinger, A. (2018). ABPS: An R package for calculating the abnormal blood profile score. Frontiers in Physiology, 9, Article 1638. https://doi.org/10.3389/fphys.2018.01638

Sharpe, K., Hopkins, W., Emslie, K. R., Howe, C., Trout, G. J., Kazlauskas, R., Ashenden, M. J., Gore, C. J., Parisotto, R., & Hahn, A. G. (2002). Development of reference ranges in elite athletes for markers of altered erythropoiesis. Haematologica, 87(12), 1248-1257. PMID: 12495898.

Sottas, P. E., Robinson, N., & Saugy, M. (2010). The athlete's biological passport and indirect markers of blood doping. Handbook of Experimental Pharmacology, 195, 305-326.

Sottas, P., Robinson, N., Saugy, M., & Niggli, O. (2008). A forensic approach to the interpretation of blood doping markers. Law, Probability and Risk, 7(3), 191-210.

Sottas, P.-E., Robinson, N., Giraud, S., Taroni, F., Kamber, M., Mangin, P., & Saugy, M. (2006). Statistical classification of abnormal blood profiles in athletes. The International Journal of Biostatistics, 2(1).

Sutehall, S., Malinsky, F., Shurlock, J., Wang, G., Bosch, A., & Pitsiladis, Y. P. (2022). Whole-blood and peripheral mononuclear cell transcriptional response to prolonged altitude exposure in well-trained runners. Clinical Journal of Sport Medicine.

Varlet-Marie, E., Audran, M., Ashenden, M., Sicart, M. T., & Piquemal, D. (2009). Modification of gene expression: help to detect doping with erythropoiesis-stimulating agents. American Journal of Hematology, 84(11), 755-759.

Walters, S. J., Campbell, M. J., & Machin, D. (2021). Medical Statistics: A Textbook for the Health Sciences (5th ed.). Wiley Blackwell.

Wang, G., Durussel, J., Shurlock, J., Mooses, M., Fuku, N., Bruinvels, G., Pedlar, C., Burden, R., Murray, A., Yee, B., Keenan, A., McClure, J. D., Sottas, P. E., & Pitsiladis, Y. P. (2017). Validation of whole-blood transcriptome signature during microdose recombinant human erythropoietin (rHuEpo) administration. BMC Genomics, 18(Suppl 8), Article 817.

Wang, G., Kitaoka, T., Crawford, A., Mao, Q., Hesketh, A., Guppy, F. M., Ash, G. I., Liu, J., Gerstein, M. B., & Pitsiladis, Y. P. (2021). Cross-platform transcriptomic profiling of the response to recombinant human erythropoietin. Scientific Reports, 11(1), Article 21705.

Wang, J., Zhao, S., Su, Z., & Liu, X. (2019). Analytical comparison between two hematological analyzer systems: Mindray BC-5180 vs Sysmex XN-1000. Journal of Clinical Laboratory Analysis, 33(8), e22955.

World Anti-Doping Agency. (2010, August). Guidelines for urine sample collection. Retrieved from https://www.wada-ama.org/sites/default/files/resources/files/WADA_Guidelines_Urine_Sample_Collection_v5.1_EN.pdf.

World Anti-Doping Agency. (2011, February). Retrieved from http://www.wada-ama.org/en.

World Anti-Doping Agency. (2014, October). Athlete Biological passport ABP operating guidelines. Retrieved from https://www.wadaama.org/sites/default/files/resources/files/wada_abp_operating_guidelines_2014_v5.0_en.pdf.

World Anti-Doping Agency. (2019). Athlete Passport Management Unit Requirements and Procedures. Retrieved December 14, 2019, from https://www.wada-ama.org/sites/default/files/resources/files/td2019apmu_final2.pdf.

World Anti-Doping Agency. (2023, July 10). Athlete Biological Passport (ABP) Operating Guidelines. Retrieved from https://www.wada-ama.org/en/resources/world-anti-doping-program/athlete-biological-passport-abp-operating-guidelines#resource-download.

Zorzoli, M. (2011). Biological passport parameters. Journal of Human Sport and Exercise, 6(2).

Appendix 1:
Participant information sheet

AN "OMICS" APPROACH TO THE DETECTION OF RECOMBINANT HUMAN ERYTHROPOIETIN (R-HUEPO) DOPING IN ENDURANCE ATHLETES

INFORMATION SHEET

What is the purpose of the study?

This study aims to recruit individuals who are not using erythropoietin to create a reference dataset that will be used in comparison to previous erythropoietin administration studies to develop enhanced anti-doping tests.

Why have I been chosen to participate?

You responded to our advertisement and have been selected as a possible participant for this investigation as you meet the criteria for one of our groups within the study.

Do I have to take part?

Participation in this research is completely voluntary; you are under no obligation to take part. If you do wish to take part, then you will be required to read and keep this information sheet and sign the consent form prior to commencing. Furthermore, if you decide to participate, then at any point feel distressed or uncomfortable, you can withdraw from participation without giving a reason or any notice.

What is expected from me if I participate?

You will be asked to complete and sign an informed consent form. It is a good idea to raise questions and concerns that may be unclear at this point. Following this, you will be provided with dates for the commencement of the study, choose ones that suit your commitments.

Once selected, you will be required to report, when suitable, to the university's clinic laboratory for sample collection. *A 3 mL whole blood sample from an antecubital vein will be collected into a Tempus™ Blood RNA Tube (Thermo Fisher Scientific, Waltham, MA, USA) to preserve RNA for subsequent extraction, and a 3 mL whole blood sample will be collected into a BD® EDTA (K2) tube for haematological analysis. You may be expected to attend the lab on 4-6 occasions throughout a period of 12 months. All samples will be solely used to contribute towards the creation of a reference dataset. No DNA/genetic testing will be carried out in this study and only extracted RNA (no blood samples) will be stored/biobanked. Commercial companies will not be given access to any biological material to emerge from the study for commercial exploitation/use.*

Please refer to the below regarding further information on the study.

Location – Sample collection will be conducted as per good clinical laboratory guidelines by a trained and certified phlebotomist (sample collector) who will explain to you about sample collection and the volume required.

Sample and data collection

Collected samples will be analysed first for complete blood count and then using Next Generation Sequencing technologies (RNA-Sequencing) to develop a "molecular signature" of gene expression for reference data that will aid in the validation of a novel gene expression anti-doping test.

What do I have to do?

You will be requested to fill in the sample and data collection log provided throughout the duration of the study.

You will be requested to visit the university's clinic laboratory every three weeks for four visits. We will provide you with a study card to remind you about the appointments.

Will I be paid for taking part?

Participation in this project is voluntary. However, reasonable travel expenses and loss of earnings occurred during the study will be compensated to the participants. This is set at KSh. 1,100 per visit as per the local regulatory requirements.

What are the possible side effects, disadvantages and risks of taking part?

There are only minimal risks associated with the data collection procedures. You will experience only minimal pain due to sample collection procedures involving blood draws.

The research will require the co-operation or permission of an individual or gatekeeper in order to gain access to you. These are coaches and university management. If you so wish, you may ask a gatekeeper to accompany you during your routine visits. This will enable you and your gatekeeper to understand that participating in this study will not interfere with your studies and lifestyle.

What are the possible benefits of taking part?

This study will enhance our understanding of the molecular changes that occur during training and also provide information on the role of gene expression of professional athletes from a variety of team and individual sports. This design will permit the effects of sex and age on "omics" profiles to be determined, new technologies will be used to measure the "molecular signature" of exercise that combine with existing indirect methods will improve Athlete Biological Passport (ABP) and promote clean sport.

What if something goes wrong?

If you are harmed due to someone's negligence, then you may have grounds for a legal action. The blood collection officers (phlebotomists) are fully trained in blood collection and familiar with dealing with relevant first aid situations. In the event of an untoward incident, the

investigators will provide first aid support in case needed. You may want to consult the university's medical doctor if you are experiencing any side effects from taking part in the study and should also inform the Principal Investigator. This study has been approved by the Maseno University Scientific and Ethics Review Committee (MUERC) and you are free to report any discomfort through the telephone number: **+25457351622** or email to: muerc-secretariate@maseno.ac.ke. You can also call the Principal Investigator (Prof. James Ombaka) on **+254721260279.**

Will my taking part in this study be kept confidential?

All your information and data obtained from testing will remain anonymous. A numeric participant ID will be assigned to you and will be used for tracking and recording throughout the study. All data will be used for research purposes only and confidentiality of the data will be respected and always maintained. Only RNA will be extracted from whole blood, analysed and remaining sample (if any) stored. Only de-identified RNA samples will be provided to other researchers for research with a favourable ethical opinion. The RNA obtained from the blood samples will be shipped to external laboratories for RNA sequencing as per regulatory requirements.

What will happen to the results of the research study?

No identifiable individual data will be provided to any third party for any purposes including for anti-doping purposes. Results will be published in a peer-reviewed scientific journal once the study is completed. The published data is available on request. You will not be identified in any publication.

Who is funding the research?

The research is being funded by the World Anti-Doping Agency (WADA).

Who has reviewed the study?

This study has been reviewed and approved by the Maseno University Scientific and Ethics Review Committee (MUERC)

If you wish to find out more about this investigation, you can contact the Principal Investigator:

Professor James H. Ombaka
Professor of Pharmacology
Department of Biomedical Sciences and Technology
School of Public Health and Community Development
Maseno University
Kisumu-Kenya.
Tel: **+25457351622**
e-mail: ombaka2000@yahoo.com

Informed consent form

INFORMED CONSENT FORM

Title of Project: AN "OMICS" APPROACH TO THE DETECTION OF RECOMBINANT HUMAN ERYTHROPOIETIN (R-HUEPO) DOPING IN ENDURANCE ATHLETES

Name of Researcher: Prof. James H. Ombaka

Contact Details: Department of Biomedical Sciences and Technology; School of Public Health and Community Development; Maseno University, Kisumu-Kenya. Tel: +25457351622; e-mail: ombaka2000@yahoo.com

Please initial box

- I confirm that I have read the Participant Information Sheet dated..... (version.....) for the above study. I have had the opportunity to consider the information, ask questions and have had these answered satisfactorily. ☐
- I consent to the use of my blood for the above study. I understand that only extracted RNA will be stored/biobanked for potential future use. ☐
- I give my consent to the storage of my RNA sample for future research project with a favourable ethical opinion (generic consent) ☐
- If generic consent if granted: I understand that my consent to the above study or future research projects with a favourable ethical opinion will continue for ten (10) years or until I withdraw my consent (whichever is sooner). ☐
- I understand that my participation is voluntary and that I am free to withdraw my consent at any time without giving a reason. Should I withdraw my consent, I understand that any of my stored sample will be disposed of however, any data already generated from the use of my sample will remain the property of the research team. ☐
- I am aware that my RNA samples may be used in a range of future studies with a favourable ethical opinion. ☐
- I agree to take part in the above study. ☐
- My biological sex is (mandatory):
  Female ☐
  Male ☐
  Other ☐
- My current sport is (mandatory):_____
- My chronological age in years (optional): _____years

Name of Participant_____ Signature: _____ Date: _____

Name of Person receiving consent _____ Signature: _____Date: _____

If you have any concerns or complaints, please direct these appropriately to Maseno University Scientific and Ethics Review Committee (MUERC) and you are free to report any discomfort through the telephone number: ++254721260279 or email to: muerc-secretariate@maseno.ac.ke

1 copy to participant, 1 copy to the researcher.

4

AN "OMICS" APPROACH TO THE DETECTION OF RECOMBINANT HUMAN ERYTHROPOIETIN (R-HUEPO) DOPING IN ENDURANCE ATHLETES

Participant ID: _____ Date of Enrollment: _____

Date of Birth: _____

Age: _____

Gender: Male_____ Female_____

Residence: _____

County: _____

College/University: _____

Do you participate in sports? Yes_____ No_____

If yes, please specify:

- ☐ Athletics
  - Track (100m, 200m, 400m, 800m, 1500m, 3000m SC, 5000m, 10000m) _____
  - Road Running (10km, Half Marathon, Marathon) _____
  - Field: (Javelin, Shortput, Discus) _____
- ☐ Soccer
- ☐ Rugby
- ☐ Tennis
- ☐ Basketball
- ☐ Badminton
- ☐ Other (Specify) _____

At what level do you compete?

- ☐ College/ University games
- ☐ National Events
- ☐ International competition (Commonwealth, World Championships, Olympics)
- ☐ Other (Specify)_____

Sample collection log

SAMPLE COLLECTION LOG

OMICS Sample schedule: STUDY-ROUTINE

Participant No:			Initials:	Visit 1 (DD/MM/YYYY; 00:00H)			
Time Point	Planned Date/Time of Collection	Actual Date/Time of Collection		Vacutainer Type	Vacutainer Number	Sample condition	Lab.Tech (Initials)
Visit 1	01/Jan/2023 00:00	__/__/____ __:__		Tempus (Blue) (Venous Blood)	1	Volume OK? Yes □ No □	
				EDTA (Purple)	1	Volume OK? Yes □ No □	
				Urine	1	Haemolysis? Yes □ No □	

OMICS Sample schedule: STUDY-ROUTINE

Participant No:			Initials:	Visit 2 (DD/MM/YYYY; 00:00H)			
Time Point	Planned Date/Time of Collection	Actual Date/Time of Collection		Vacutainer Type	Vacutainer Number	Sample condition	Lab.Tech (Initials)
Visit 2	01/Jan/2023 00:00	__/__/____ __:__		Tempus (Blue) (Venous Blood)	1	Volume OK? Yes □ No □	
				EDTA (Purple)	1	Volume OK? Yes □ No □	
				Urine	1	Haemolysis? Yes □ No □	

OMICS Sample schedule: STUDY-ROUTINE

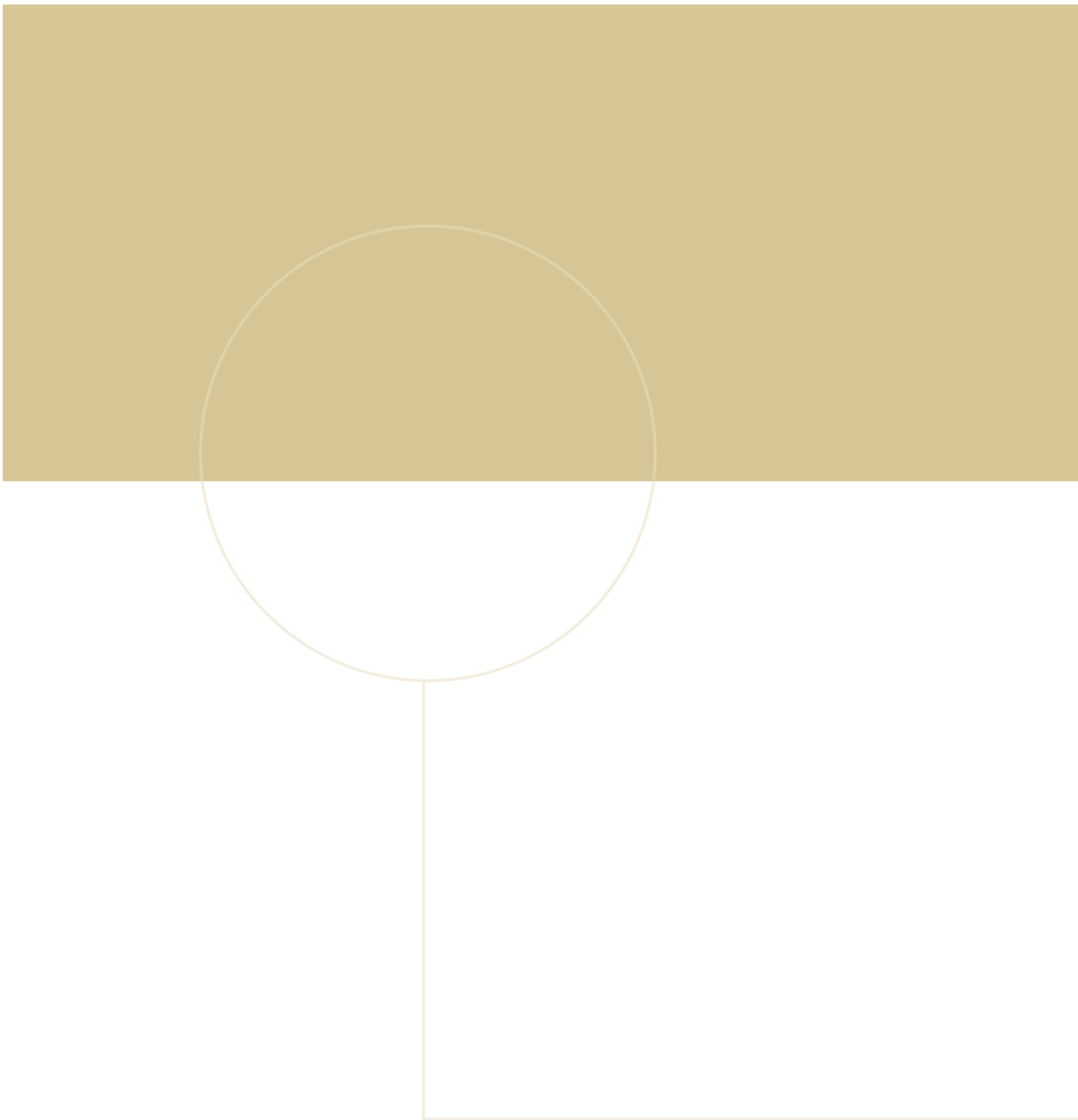
Participant No:			Initials:		Visit 3 (DD/MM/YYYY; 00:00H)		
Time Point	Planned Date/Time of Collection	Actual Date/Time of Collection		Vacutainer Type	Vacutainer Number	Sample condition	Lab.Tech (Initials)
Visit 3	01/Jan/2023 00:00	__/__/____ __:__		Tempus (Blue) (Venous Blood)	1	Volume OK? Yes □ No □	
				EDTA (Purple)	1	Volume OK? Yes □ No □	
				Urine	1	Haemolysis? Yes □ No □	

OMICS Sample schedule: STUDY-ROUTINE

Participant No:			Initials:		Visit 4 (DD/MM/YYYY; 00:00H)		
Time Point	Planned Date/Time of Collection	Actual Date/Time of Collection		Vacutainer Type	Vacutainer Number	Sample condition	Lab.Tech (Initials)
Visit 4	01/Jan/2023 00:00	__/__/____ __:__		Tempus (Blue) (Venous Blood)	1	Volume OK? Yes □ No □	
				EDTA (Purple)	1	Volume OK? Yes □ No □	
				Urine	1	Haemolysis? Yes □ No □	

Table of Figures

Table of Tables



NTNU

Norwegian University of
Science and Technology