DM

Urinary Metabolomic Fingerprints as a Powerful Tool for the Next Generation of Prostate Cancer Diagnosis

MASTER DISSERTATION

Cristina Viveiros Berenguer
MASTER IN APPLIED BIOCHEMISTRY

UNIVERSIDADE da MADEIRA
A Nossa Universidade
www.uma.pt

September | 2022

Urinary Metabolomic Fingerprints as a Powerful Tool for the Next Generation of Prostate Cancer Diagnosis

MASTER DISSERTATION

Cristina Viveiros Berenguer
MASTER IN APPLIED BIOCHEMISTRY

ORIENTATION
José de Sousa Câmara

CO-ORIENTATION
Jorge Augusto Machado Pereira

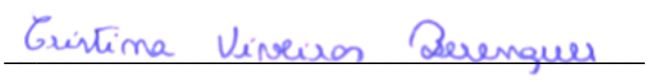
DECLARAÇÃO

O plágio consiste na apresentação de ideias, opiniões, frases/textos, resultados ou conclusões de outros como sendo suas/seus, mesmo que tenha havido tradução dos mesmos. A prática de plágio constitui uma grave violação da ética académica, além de poder levar à reprovação ou à retirada do grau, assim como à responsabilidade civil, criminal e disciplinar.

Assim, declaro por minha honra que a presente dissertação de mestrado é da minha exclusiva autoria, é original, e nela referenciei e citei todas as fontes utilizadas.

Funchal, 05 de setembro de 2022

Cristina Viveiros Berenguer



Urinary Metabolomic Fingerprints as a Powerful Tool for the Next Generation of Prostate Cancer Diagnosis

Cristina Viveiros Berenguer

Dissertação apresentada à Universidade Da Madeira para cumprimento dos requisitos necessários à obtenção do grau de Mestre em Bioquímica Aplicada, realizada sob a orientação científica do Professor Doutor José de Sousa Câmara e coorientação do Doutor Jorge Augusto Machado Pereira.

Faculdade de Ciências Exatas e de Engenharia, Universidade da Madeira
Centro de Química da Madeira

Setembro 2022

Funchal – Portugal

AGRADECIMENTOS

O apoio e a ajuda de muitas pessoas tornaram possível a concretização desta dissertação e gostaria de prestar-lhes o meu sincero agradecimento.

Ao meu orientador, Professor Doutor José de Sousa Câmara, e ao meu coorientador, Doutor Jorge A.M. Pereira, pela disponibilidade, oportunidade, aprendizagens, orientação e apoio ao longo deste trabalho. Obrigada pelo incentivo, paciência e ainda pela confiança que depositaram em mim e em todos os projetos ao longo do meu percurso académico.

Ao Centro de Química da Madeira (CQM) por proporcionar todas as condições que tornaram possível o desenvolvimento desta dissertação. Também gostaria de agradecer às instituições financiadoras envolvidas: **FCT** - Fundação para a Ciência e a Tecnologia, através dos projetos CQM *Base Fund* (UIDB/00674/2020) e *Programmatic Fund* (UIDP/00674/2020); **ARDITI** - Agência Regional para o Desenvolvimento da Investigação Tecnologia e Inovação, através dos projetos M1420-01-0145-FEDER-000005 - Centro de Química da Madeira - CQM+ (Programa Madeira 14-20) e M1420-09-5369-FSE-000001 (bolsa de pós-doutoramento concedida a JAMP); **FCT** e **programa Madeira 14-20** para a Rede Nacional de Espectrometria de Massa (RNEM) através do projeto PROEQUIPRAM (M14-20 M1420-01-0145-FEDER-000008).

Ao Doutor Ricardo Sousa, presidente do **Núcleo Regional da Madeira da Liga Portuguesa Contra o Cancro (NRM-LPCC)**, pelo apoio, suporte na recolha de amostras e pela oportunidade de concretizar este trabalho no âmbito da **Bolsa Rubina Barros**. Gostaria de agradecer ao **Centro de Investigação Doutora Maria Isabel Mendonça – SESARAM, EPERAM**, ao Serviço de Urologia do SESARAM, EPERAM, ao Centro de Saúde do Bom Jesus, e à equipa do projeto de investigação CANPRO, nomeadamente, Doutora Cristina Ornelas, Doutora Maria do Carmo Pinto, Doutor João Pedro A. Vital, Doutor Pedro Berenguer, Doutora Sónia Freitas e à Professora Doutora Ana Célia Sousa, pela oportunidade e cedência das amostras imprescindíveis para a concretização desta dissertação, e, em especial, ao Doutor Ferdinando Pereira por toda a colaboração, conhecimento e disponibilidade desde o início deste trabalho.

Aos meus colegas do CQM que colaboraram comigo e sempre estiveram disponíveis a ajudar-me: Beatriz Andrade, João Gonçalves, Jorge Freitas, Mariangie Castillo, Rosa Perestrelo e Vera Alves. À Filipa Jardim e à Carolina Andrade, as minhas companheiras de mestrado, obrigada por todo o incentivo, apoio, risadas, conselhos e por tornarem melhor o meu percurso académico. Às técnicas de laboratório, Paula Andrade e Paula Vieira, e à Engenheira Teresa Abreu por estarem sempre disponíveis para ajudar-me.

Por fim, agradeço a todas as pessoas que, de forma direta ou indireta, auxiliaram a execução deste trabalho. A todos os que estão no meu coração e sempre apoiaram e confiaram em mim, a minha família, e em particular aos meus pais que acompanharam todo o meu percurso. Ao João, pela disponibilidade, paciência, incentivo e todo o apoio.

Muito obrigada a todos!

“Twenty years from now you will be more disappointed by the things that you didn't do than by the ones you did do. So, throw off the bowlines. Sail away from the safe harbour. Catch the trade winds in your sails. Explore. Dream. Discover.”

Mark Twain

SCIENTIFIC WORK

The development of this Master's Dissertation was a great opportunity to acquire organizational and communication skills.

Article

Berenguer CV, Pereira F, Pereira JAM, Câmara JS. Volatilomics: an emerging and promising avenue for the detection of potential prostate cancer biomarkers. Cancers. 2022;14(16):3982. doi:10.3390/cancers14163982.

Oral communications

Berenguer C, Pereira F, Sousa R, Pereira JAM, Câmara JS. Establishment of urinary volatomic profiles for the identification of putative prostate cancer biomarkers. 8th CQM Annual Meeting; 2021 October 7-8; Funchal, Madeira. ISBN: 978-989-54090-5-1.

Berenguer C. Combate ao cancro - O poder que temos na prevenção contra o cancro. Bridging the Gap – Mulheres na Ciência: Passado, Presente e Futuro; 2022 May 17-18; Câmara de Lobos, Madeira.

Berenguer C, Pereira F, Pereira JAM, Câmara JS. Urinary Metabolomic Fingerprints as a Powerful Tool for the Next Generation of Prostate Cancer Diagnosis. 9th CQM Annual Meeting; 2022 September 28-30; Funchal, Madeira.

Poster communications

Berenguer C, Pereira F, Sousa R, Pereira JAM, Câmara JS. Estabelecimento de padrões volatómicos como plataformas inovadoras e integradas para a próxima geração de diagnóstico do cancro da próstata. 3º Encontro Nacional de Jovens Investigadores em Oncologia; 2021 September 24; Porto, Portugal.

Other publications

During my academic path, I had the opportunity to develop other research projects, which resulted in scientific publications in international peer-reviewed journals, as well as, in participation in national meetings, through oral and poster communications.

Articles

Berenguer C, Pereira JAM, Câmara JS. Fingerprinting the volatile profile of traditional tobacco and e-cigarettes: A comparative study. Microchem J. 2021;166(7):106196. doi:10.1016/j.microc.2021.106196.

Figueira JA, Porto-Figueira P, Berenguer C, Pereira JAM, Câmara JS. Evaluation of the Health-Promoting Properties of Selected Fruits. Molecules. 2021;26(14):4202. doi:10.3390/molecules26144202.

Berenguer C, Pereira JAM, Câmara JS. Urinary volatomic profile of traditional tobacco smokers and electronic cigarettes users as a strategy to unveil potential health issues. J Sep Sci. 2022;45:582-593. doi:10.1002/jssc.202100671.

Pereira JAM, Berenguer CV, Andrade CFP, Câmara JS. Unveiling The Bioactive Potential Of Fresh Fruit And Vegetable Waste In Human Health From A Consumer Perspective. Appl Sci. 2022; 12(5):2747. doi: 10.3390/app12052747.

Câmara JS, Perestrelo R, Berenguer CV, Andrade CFP, Gomes TM, Olayanju B, Kabir A, Rocha CMR, Teixeira JA, Pereira JAM. Green Extraction Techniques as Advanced Sample Preparation Approaches in Biological, Food, and Environmental Matrices: A Review. Molecules. 2022;27(9):2953. doi:10.3390/molecules27092953

Câmara JS, Perestrelo R, Olayanju B, Berenguer CV, Kabir A, Pereira JAM. Overview of Different Modes and Applications of Liquid Phase-Based Microextraction Techniques. Processes. 2022;10:1347. doi: 10.3390/pr10071347

Oral communications

Porto-Figueira P, Figueira J, Berenguer C, Pereira JAM, Câmara JS. Investigation of the health-promoting benefits from different food matrices: From extraction to *in vitro* evaluation. XX EuroFood Chem Conference; 2019 June 17-19; Porto, Portugal.

Porto-Figueira P, Figueira J, Baptista M, Berenguer C, Pereira JAM, Máximo V, Câmara JS. Selection of the best urine sampling period based on the volatomic profile. 11 National Meeting on Chromatography; 2019 December 9-11; Lisbon, Portugal. ISBN: 978-989-8124-29-6.

Berenguer C, Pereira JAM, Câmara JS. Fingerprinting the volatile profile of traditional tobacco and e-cigarettes: A comparative study. World Vape Day – APORVAP [online event] 2021 May 30.

Figueira J, Porto-Figueira P, Berenguer C, Pereira JAM, Câmara JS. Evaluation of the health-promoting properties of different fruits grown in Madeira Island. XV Encontro de Química dos Alimentos; 2021 September 5-8; Funchal, Madeira. ISBN: 978-989-8805-68-3.

Poster communications

Berenguer C, Jardim F, Gouveia M, Ribeiro JP, Câmara JS. Establishment of the volatile profile of traditional and industrial ciders by HS-SPME/GC-MS. XV Encontro de Química dos Alimentos; 2021 September 5-8; Funchal, Madeira.

ABSTRACT

Prostate cancer (PCa) remains the most frequent malignant tumour and a leading cause of oncological death in men, despite the spectacular advances in molecular medicine, including genomics, proteomics, transcriptomics, lipidomics, and personalized medicine. Apart from classical biomarkers, the study of endogenous volatile organic metabolites (VOMs) biosynthesized by different metabolic pathways and found in several biofluids is emerging as an innovative, efficient, and non-invasive source of data to establish the volatilomic biosignature of PCa.

In this context, the primary objective of this thesis was to establish the urinary volatilomic profile of PCa using headspace solid-phase microextraction combined with gas chromatography-mass spectrometry (HS-SPME/GC-MS). This non-invasive approach to set putative PCa biomarkers was applied to PCa patients (n = 29), men subjected to a radical prostatectomy (RP, n = 34), and cancer-free individuals (control group, n = 49). A total of 60 VOMs belonging to different chemical families were identified in the groups under study. This included phenolic compounds, terpenes, norisoprenoids, ketones, alcohols, and sulphur-containing and furanic compounds. The data matrix obtained was submitted to multivariant analysis, through partial least-squares discriminant analysis (PLS-DA). The results obtained did not show complete discrimination between the groups under study, since more than 50 % of the variability in the outcome data could not be explained by the models. The heatmap according to Pearson's correlation showed that 2-(1-cyclopentyl)furan, 2-pentanone and 2-bromophenol were more associated with the control group, while carvone, α-corocalene, 2-acetylfuran, and cyclohexanone showed high correlations with the PCa group. In contrast, 2-ethyl-5-methylfuran, methanethiol, *o*-methoxyphenol, *p*-cymenene, and 3,4-dehydro-β-ionene, were more associated with the RP group. An exhaustive study of the demographic and clinical characteristics of the patients recruited is crucial to elucidate the changes that occur in their volatilomic profile. Moreover, a larger cohort of samples is necessary to improve the predictive power and reliability of the statistical models developed.

Keywords: Prostate cancer; Volatilomics; Urine; Biomarkers.

RESUMO

O cancro da próstata (PCa) continua a ser o tumor maligno mais frequente e uma das principais causas de morte oncológica nos homens, apesar dos progressos importantes na medicina molecular, incluindo na genómica, proteómica, transcriptómica, lipidómica, e na medicina personalizada. Além dos biomarcadores clássicos, o estudo dos metabolitos orgânicos voláteis (VOMs) endógenos, biosintetizados por diferentes vias metabólicas e encontrados em vários biofluidos, está a emergir como uma abordagem inovadora, eficiente e não invasiva para o estabelecimento de uma bioassinatura volatilómica do PCa.

Neste contexto, o principal objetivo desta tese consistiu no estabelecimento do perfil volatilómico urinário do PCa, usando uma abordagem não invasiva, a microextracção de fase sólida por *headspace* combinada com cromatografia gasosa acoplada à espectrometria de massa (HS-SPME/GC-MS), com o objetivo de identificar possíveis biomarcadores para esta forma de cancro. Para isso, foram analisados indivíduos com PCa (n = 29), sujeitos submetidos a prostatectomia radical (RP, n = 34) e indivíduos saudáveis (grupo de controlo, n = 49). No total, foram identificados 60 VOMs nos grupos em estudo, pertencentes a diferentes famílias químicas, nomeadamente compostos fenólicos, terpenos, norisoprenoides, cetonas, álcoois, compostos contendo enxofre e compostos furânicos. Esta matriz de dados foi submetida a análise multivariada, através da análise discriminante dos mínimos quadrados parciais (PLS-DA), resultando numa discriminação incompleta dos grupos em estudo, uma vez que mais de 50 % da variabilidade dos dados obtidos não pode ser explicada pelos modelos gerados. O *heatmap* construído, segundo a correlação de Pearson, demonstrou que os metabolitos 2-(1-ciclopentil)furano, 2-pentanona e 2-bromofenol estavam mais associados ao grupo de controlo, enquanto que a carvona, α-corocaleno, 2-acetilfurano e a ciclohexanona apresentaram correlações elevadas com o grupo do PCa. Quanto ao grupo da RP, os metabolitos 2-etil-5-metilfurano, metanotiol, *o*-metoxifenol, *p*-cimeneno e 3,4-dehidro-β-ioneno, foram os que apresentaram correlações mais elevadas. Deste modo, o estudo exaustivo das características demográficas e clínicas dos pacientes recrutados, assim como grupos de pacientes recrutados maiores, serão cruciais para compreender as mudanças que ocorrem nos perfis volatilómicos obtidos, assim como para melhorar o poder preditivo e a fiabilidade dos modelos estatísticos desenvolvidos.

Palavras-chave: Cancro da próstata; Volatilómica; Urina; Biomarcadores.

CONTENTS

List of figures xv

List of tables xvii

Abbreviations xix

Chapter I. Introduction 1

1.1. Prostate cancer 1

1.2. Incidence and mortality rates worldwide 2

1.2.1. Prostate cancer burden in Europe 4

1.2.2. Prostate cancer burden in Portugal 4

1.3. Prostate cancer risk factors 5

1.3.1. Ethnicity 9

1.3.2. Family history and genetic factors 9

1.3.3. Dietary factors 10

1.3.4. Unhealthy lifestyles - obesity, insulin, and tobacco use 12

1.3.5. Occupation 12

1.4. Prostate cancer screening 13

1.4.1. Prostate-specific membrane antigen: a theranostic approach 15

1.4.1.1. Molecular imaging 16

1.4.1.2. Radioligand targeted therapy 17

1.4.2. Tumour biomarkers 18

1.4.2.1. Molecular biomarkers 18

1.4.2.2. Long non-coding RNAs 21

1.4.2.3. Liquid biopsy biomarkers 23

1.4.3. Active surveillance and risk-stratification algorithms 24

1.5. Contribution of the OMICs science 26

1.5.1. Volatilomics – an emerging yet challenging approach 27

1.5.1.1. Challenges of volatilomics 32

1.5.2. Volatilome of prostate cancer - a promising approach for biomarkers' detection 34

1.5.3. Chemometrics tools in volatilomics: supervised and unsupervised methods 39

1.6. Objectives of the thesis and experimental approach followed 41

Chapter II. Materials and methods **43**
2.1. Extraction and analysis conditions 43
2.1.1. Materials and reagents 43
2.1.2. Subjects and urine sampling 43
2.1.3. HS-SPME procedure 45
2.1.4. GC-MS analysis 45
2.2. Statistical analysis 46

Chapter III. Results and discussion **47**
3.1. Characterization of urinary volatile metabolites 47
3.2. Chemometric analysis of urine samples 54

Chapter IV. Conclusions and future perspectives **65**

References **67**

Appendices. Supplementary material **89**
Additional figures and tables 89
Supplementary information 95

LIST OF FIGURES

Figure 1 - A. Most diagnosed types of cancer among men worldwide, 2020. Nonmelanoma skin cancer was included in the calculations of top cancer per country. **B.** Leading cause of cancer deaths among men worldwide, 2020. Source: GLOBOCAN 2020 (1). …… 3

Figure 2 - Estimated number of new cases and deaths from PCa from 2020 to 2025. Source: GLOBOCAN, 2020 (1). ……...……...……... 3

Figure 3 – A. Number of PCa deaths from 2010 to 2019 in Portugal. **B.** Number of PCa deaths from 2010 to 2019 in the Autonomous Regions of Azores and Madeira. Source: Instituto Nacional de Estatística, 2022 (23). ……..... 5

Figure 4 – Modifiable (in yellow) and unmodifiable (in orange) PCa risk factors. ……..... 8

Figure 5 – PCa diagnosis pathway. ……...……..... 13

Figure 6 – Application of a risk stratification algorithm in the diagnosis pathway for men without a previous diagnosis of PCa (9,67). ……..……..... 26

Figure 7 - General flowchart of volatilomics approaches. ……..... 28

Figure 8 - Cancer-related metabolic and biochemical activities can be detected through VOMs. The establishment of a volatilomic biosignature can lead to the discovery of biomarkers for PCa diagnosis. ……..…….....…….....…….....…….....…….... 35

Figure 9 – Representative chromatograms of the groups analysed. Legend: CTRL: control group; RP: radical prostatectomy group; PCa: prostate cancer group. …….. 47

Figure 10 – Distribution of the chemical families identified in the studied groups, RP (*n* = 34) control group (*n* = 49), and PCa (*n* = 29). Legend: Ph: phenolic compounds; T: terpenes; N: norisoprenoids; K: ketones; A: alcohols; S: sulphur-contain compounds; F: furanic compounds; O: others. ……..…….....…….....…….....…..…..... 51

Figure 11 – Score scatter plots of the volatilomic fingerprint of urine samples from the three groups under study, control (0) *vs* PCa (1) *vs* RP (2). **A.** PCA PC1 × PC2 score scatter plot. **B.** PLS-DA score scatter plot. **C**. Heatmap illustrating the mean levels (normalized peak areas) of the 22 VOMs statistically significant by Pearson's distance analysis. Rows correspond to the mean normalized peak area of each metabolite with the sample groups in the columns.……... 56

Figure 12 - A: PLS-DA score scatter plots of the volatilomic fingerprint of urine samples from the three groups under study, control (0) *vs* PCa (1) *vs* RP (2). **B.** VIP scores **C:** 10-fold CV performance of control *vs* PCa *vs* RP (*means best Q^2 value). **D:** Model validation by permutation test based on 1000 permutations of VOMs obtained by GC-MS of urine samples from thr groups under study….. 57

Figure 13 - A-C: PLS-DA score scatter plots of the volatilomic fingerprint of urine samples from **A.** control (0) *vs* PCa (1), **B.** control (0) *vs* RP (2), and **C.** PCa (1) *vs* RP (2). **D-F:** Corresponding VIP scores. ……. 59

Figure 14 - A-C: 10-fold CV performance of **A.** control *vs* PCa, **B.** control *vs* RP, and **C.** PCa *vs* RP (*means best Q^2 value). **D-F:** Model validation by permutation test based on 1000 permutations of VOMs obtained by GC-MS of urine samples from the groups under study……..…….…... 60

Figure 15 - Score scatter plots of the volatilomic profile of urine samples from the control (0) *vs* PCa with combined therapy (1) *vs* PCa with one therapy (2) *vs* PCa with no therapy (3) *vs* RP (4). **A.** PCA PC1 × PC2 score scatter plot. **B.** PLS-DA score scatter plot. ……...…….. 62

Figure 16 – A-C: PLS-DA score scatter plots of the volatilomic fingerprint of urine samples from **A.** control (0) *vs* PCa with combined therapy (1) *vs* PCa with one therapy (2) *vs* PCa with no therapy (3) **B.** PCa with combined therapy (1) *vs* PCa with one therapy (2) *vs* PCa with no therapy (3), and **C.** PCa with combined therapy (1) *vs* PCa with one therapy (2) *vs* PCa with no therapy (3) *vs* RP (4). **D-F:** Corresponding VIP scores. …… 64

Figure S.1 – Representative chromatogram of the pure standards applied to confirm the VOMs identified using the NIST library. Legend: 1: 2-pentanone (9.31 min); 2: 3-hexanone (12.80 min); 3: dimethyl disulfide (13.94 min); 4: 4-heptanone (17.75 min); 5: 2-ethyl-1-hexanol (61.76 min); 6: octanoic acid (104.84 min). ……...……...…….....…… 89

Figure S.2 – Score scatter plots of the volatilomic profile of urine samples from the groups under study. **A-C:** PCA PC1 x PC2 score scatter plot of **A.** control (0) *vs* PCa (1), **B.** control (0) *vs* RP (3), and **C.** PCa (1) *vs* RP (3). **C.** PCA PC1 × PC2 score scatter plot. **D-F:** Corresponding PLS-DA score scatter plots. ……..…….....…….....…….....…….....…..... 90

Figure S.3 – Score scatter plots of the volatilomic profile of urine samples from the groups under study. **A-C:** PCA PC1 x PC2 score scatter plot of **A.** control (0) *vs* PCa with combined therapy (1) *vs* PCa with one therapy (2) *vs* PCa with no therapy (3), **B.** PCa with combined therapy (1) *vs* PCa with one therapy (2) *vs* PCa with no therapy (3), and **C.** PCa with combined therapy (1) *vs* PCa with one therapy (2) *vs* PCa with no therapy (3) *vs* RP (4). **D-F:** Corresponding PLS-DA score scatter plots…….....…….....…….....…….....…….....…….....….. 91

LIST OF TABLES

Table 1 – PCa risk factors and their role in the development of this cancer (articles from the last 5 years). ……...……...……...……. 6

Table 2 – Emerging diagnostic methods for PCa detection and management (articles from the last 5 years). ……...……...……...……. 14

Table 3 - Potential clinical utility, characteristics, and availability of PCa biomarkers. ……... 18

Table 4 - Studies on VOMs for the identification of cancer biomarkers in non-invasive matrices. ……...……...……...……... 29

Table 5 - Studies on VOMs for the identification of PCa biomarkers. ……. 36

Table 6 - Demographic and clinical data of the cancer-free controls, PCa patients and RP subjects included in this study. ……..…….....…….....……..... 44

Table 7 – Identification, chemical family, possible origin, and mean relative peak area of the identified VOMs in the control group, PCa patients and RP subjects (n = 3; RSD < 20 %)…….. 48

Table 8 - Important features identified by one-way ANOVA with post-hoc Turkey test (p<0.05) for the three groups under study, control (0), PCa (1), and RP (2)……... 55

Table 9 – PLS-DA CV details of the comparison of control *vs* PCa *vs* RP……...…..…… 58

Table S.1 – Identification, chemical family, possible origin, and frequency of occurrence of the identified VOMs in the control group, PCa patients and RP subjects (n = 3; RSD < 20 %)…. 92

Table S.2 - PLS-DA CV details of the models generated in the three comparisons, control *vs* PCa, control *vs* RP, and PCa *vs* RP. ……..…… 94

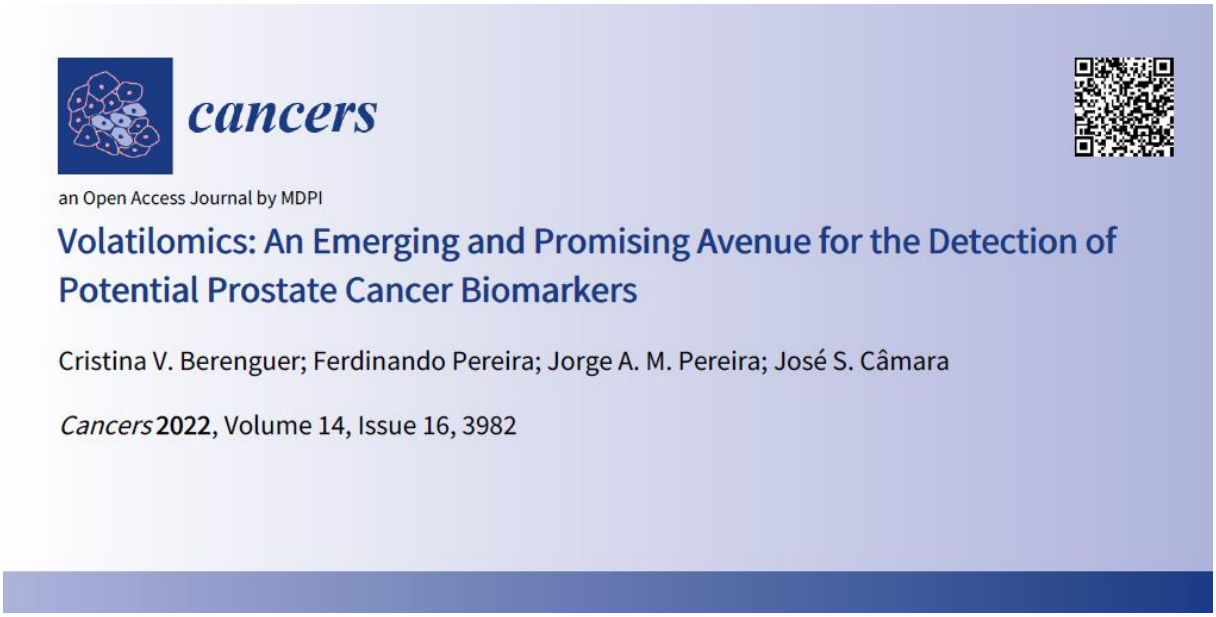
ABBREVIATIONS

^{225}Ac	Actinium-225
^{211}At	Astatine-211
^{68}Ga	Gallium-68
^{177}Lu	Lutetium-117
4KScore	Four-kallikrein score
^{223}Ra	Radium-223
^{186}Re	Rhenium-186
^{153}Sm	Samarium-153
^{89}Sr	Strontium-89
^{99m}Tc	Technetium-99m
^{227Th}	Thorium-227
ANN	Artificial neural analysis
AUC	Area under the receiver operating characteristic curve
Bact	Bacterial
BC	Bladder cancer
BMI	Body mass index
BPH	Benign prostate hyperplasia
BTC	Breast cancer
CAR	Carboxen
CC	Colon cancer
cfDNA	Cell-free deoxyribonucleic acid
cfRNA	Cell-free ribonucleic acid
CLIA	Clinical Laboratory Improvement Amendments
CRPCa	Castration-resistant prostate cancer
CT	Computed tomography
CTCs	Circulating tumour cells
ctDNA	Circulating tumour deoxyribonucleic acid
CTRL	Control
CYP	Cytochrome P
CV	Cross-validation
DFA	Discriminant function analysis
dHS-SPME	Dynamic headspace solid-phase microextraction
Diet	Dietary

DNA	Deoxyribonucleic acid
DRE	Digital rectal examinations
DVB	Divinylbenzene
EAU	European Association of Urology
Endo	Endogenous
Env	Environmental
E-nose	Electronic noses
EPI	ExoDx Prostate IntelliScore
EVs	Extracelular vesicles
Exo	Exogenous
FDA	Food & Drug Administration
FDR	False discovery rate
GC	Gas chromatography
GC-IMS	Gas chromatography-ion migration spectroscopy
GC-MS	Gas chromatography-mass spectrometry
GC-qMS	Gas chromatography-quadrupole mass spectrometry
GC-TOF-MS	Gas chromatography coupled to time-of-flight mass spectrometry
HC	Healthy control
HCC	Hepatocellular cancer
HCl	Hydrochloric acid
HS	Headspace
HS-SPME	Headspace solid-phase microextraction
hK2	Human Kallikrein 2
IGF	Insulin-like growth factor
IS	Internal standard
LDA	Linear discriminant analysis
lncRNAs	Long non-coding ribonucleic acids
LC	Liquid chromatography
LUC	Lung cancer
Mp	Multiparametric
MRI	Magnetic resonance imaging
mRNA	Messenger ribonucleic acid
MPS	MyProstateScore
MS	Mass spectrometry
m/z	Mass-to-charge ratio
NA	Not analysed

NaCl	Sodium chloride
NIST	National Institute of Standards and Technology
OPLS-DA	Orthogonal partial least-squares discriminant analysis
PAHs	Polycyclic aromatic hydrocarbons
PCa	Prostate cancer
PCA	Principal component analysis
PCA3	Prostate cancer antigen 3
PCs	Principal components
PDMS	Polydimethylsiloxane
PET	Positron emission tomography
PLS	Partial least squares
PLS-DA	Partial least-squares discriminant analysis
PG	Prostaglandins
PHI	Prostate health index
PI-RADS	Prostate imaging-reporting and data system
PNN	Probabilistic neural network
Ppm	Parts per million
PTR-TOF-MS	Proton-transfer-reaction time-of-flight mass spectrometry
PSA	Prostate-specific antigen
PSMA	Prostatic specific membrane antigen
PVC	Polyvinyl chloride
p2PSA	Precursor isoform uncomplexed [-2]pro-prostate-specific antigen
Q^2	Goodness of prediction
R^2	Goodness of fit
RAM	Autonomous Region of Madeira
RCs	Risk calculators
REC	Renal cancer
RNA	Ribonucleic acid
ROC	Receiver operating characteristic
ROS	Reactive oxygen species
RP	Radical prostatectomy
RT	Retention time
SBSE	Stir bar sorptive extraction
SD	Standard deviation
Sens	Sensitivity
Spec	Specificity

SPME	Solid-phase microextraction
Syst	Systemic
TD-GC-MS	Thermal desorption gas chromatography-mass spectrometry
TD-GC-TOF-MS	Two-dimensional gas chromatography coupled to time-of-flight mass spectrometer
TDN	1,1,6-Trimethyl-1,2-dihydronaphthalene
TRUS	Transrectal ultrasound
Unk	Unknown
UV	Ultraviolet
Versus	*Vs*
VIP	Variable importance in projection
VOMs	Volatile organic metabolites
wASR	World age-standardized incidence rates
WHO	World Health Organization



CHAPTER I.

INTRODUCTION

1.1. Prostate cancer

Prostate cancer (PCa) is the second most frequent malignant tumour and the fifth leading cause of cancer death among men worldwide (the leading cause of cancer death among men in 46 countries) (1, 2). In 2020, almost 1.4 million new cases (15.1 % of the total malignant tumours among men) and about 0.4 million deaths (6.8 % of the total tumour deaths) were estimated (GLOBOCAN data) (1, 2). PCa burden was very dramatic until the beginning of the XXI century, due to the increased use of the prostate-specific antigen (PSA) test for screening. From this date onwards, different innovations increasing the efficacy of the therapeutic methods, along with earlier diagnoses, led to a significant reduction in the number of deaths and a less pronounced downward trend in the incidence of PCa.

PCa is very heterogeneous in terms of grade and genetics, displaying complex biological, hormonal, and molecular features (1). This cancer presents diverse phenotypes, ranging from indolent asymptomatic, a non-life-threatening form, to metastatic, very aggressive, rapidly progressive, and lethal forms (3, 4). Unlike diseases such as breast and colon cancer, no major predisposition genes for PCa have been detected. Instead, multiple chromosomal loci of susceptibility genes have been identified and most of the genomic regions remain poorly studied, which explains this cancer's heterogeneity (5). Furthermore, epigenetic factors play an important role in its clinical phenotypes (6).

Epidemiological studies have shown that the geographical and racial distribution differences in PCa trends, incidence, and mortality rates (7), reflect differences in the distribution of populations, with varying degrees of genetic susceptibility. Another difference is in the availability, use, and access to medical care (8), especially regional differences in the diagnosis of latent cancers through PSA screening (7, 9). PCa screening is frequent in areas with a high or very high human development index, but most individuals seem to be unaware of its main objectives and potentially harmful effects. For cancer control, it is of utmost importance to build a sustainable platform for the dissemination of cancer prevention and provision of cancer care, namely in low-income and transitioning countries. These results

highlight the need to increase health literacy and ensure that opportunistic screening is preceded by a thorough discussion about its potential benefits and risks (10).

PCa and subsequent treatments have a high impact on the functional and psychological status of patients, significantly affecting their quality of life (11). The current diagnostic methods are based on the measurement of PSA blood levels, transrectal ultrasound, digital rectal examinations (DRE), and prostate biopsies (7). However, these methods are invasive, expensive, and unpleasant to patients, with consequent risks of unnecessary complications, and can lead to both false-positive and false-negative results (12). The PSA test has a low selectivity to detect PCa, which often leads to the overdiagnosis and overtreatment of relatively indolent tumours with a low potential for morbidity or death if left untreated (13). Hence, it is urgent to find specific diagnostic tools for non-invasive detection of PCa, preferentially able to stratify patients by cancer aggressiveness and consequent choice of therapy, resulting in personalized and targeted therapies.

1.2. Incidence and mortality rates worldwide

The prevalence of PCa varies among different racial groups and the vast disparity has been associated with socioeconomic conditions, as well as environmental and biological factors, which play an important role in the aetiology of PCa. Variations in the incidence rates may be due to underdiagnosis, differences in screening methods, and disparities in healthcare access (1). Requesting the PSA test directly influenced the incidence values around the world. In more developed countries, the use of the PSA test resulted in a reduction in the mortality rates, while in less developed countries it showed an increase, reflecting the access to early detection and available therapies through the PSA result (2, 14).

In 2020, PCa was the most frequently diagnosed cancer among men in 121 of 185 world countries (1, 2, 7) (Figure 1). The world age-standardized incidence rates (wASR) are three times higher in areas with high or very high human development index (2, 7), when compared with less developed countries (37.5 and 11.3/100000, respectively), while the mortality rates are almost constant (8.1 and 5.9/100000, respectively). The variations among countries with a higher human development index might be due to differences in diagnostic testing and increased life expectancy (1). In developing countries, medical care and assistance are not widely accessible, which can explain the high mortality despite the lower incidence (1).

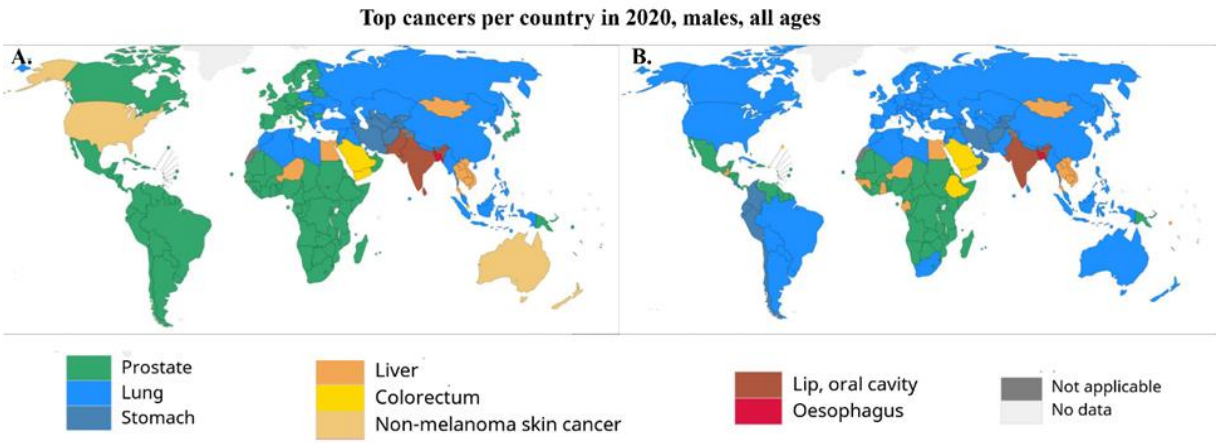


Figure 1 - A. Most diagnosed types of cancer among men worldwide, 2020. Nonmelanoma skin cancer was included from calculations of top cancer per country. **B.** Leading cause of cancer deaths among men worldwide, 2020. Source: GLOBOCAN, 2020 (1).

Overall, in the last 5 years, the mortality rates have declined, most probably due to improved access to treatments and dissemination of therapies, such as surgery and hormonotherapy. However, the mortality rates continue to increase in several countries in Asia, South America, and Eastern Europe, which may be related to increases in obesity and unhealthy lifestyles (poor diet, physical inactivity, and tobacco and alcohol consumption) (7, 15, 16). The projections for the next 5 years show an increasing trend in the estimated number of new cases and deaths (Figure 2), for all continents. Furthermore, in the upcoming years, the number of cases may increase, because the detour of resources to the COVID-19 pandemic has delayed the diagnosis and treatment of many cases.

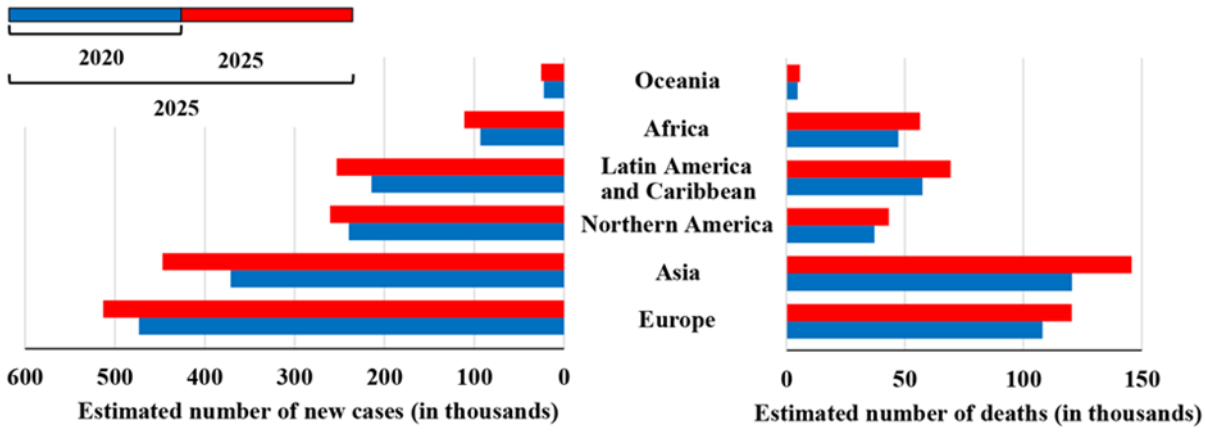


Figure 2 - Estimated number of new cases and deaths from PCa from 2020 to 2025. Source: GLOBOCAN, 2020 (1).

1.2.1. Prostate cancer burden in Europe

PCa incidence in Europe is high when compared with other geographical areas, such as Africa or Asia, due to the use of PSA for early detection (17). Consequently, this leads to the detection of many indolent cases and potentially reduces mortality. More developed regions exhibit higher incidence rates owing to increased awareness and improvements in the health care systems (1, 2, 7), and the highest estimated rates were found in Western and Northern countries (7). These regional differences are related to environmental risk factors, and differences in healthcare policies across individual countries, like the access and availability of costly targeted therapies, in addition to heterogeneity in health and socioeconomic status (17, 18). In Western Europe, the incidence rates decreased or stabilized during the most recent years in Austria, France, Germany, Switzerland, and the Netherlands. The mortality rates decreased in Austria, France, Germany, and the Netherlands, and stabilized in Belgium, Hungary, Luxembourg, and Switzerland (7). Concerning the Northern European countries, the incidence rates were stable during the last 5 years, while the mortality rates decreased in the United Kingdom, Finland, Ireland, and Sweden, and stabilized in the remaining countries. In Eastern and Southern Europe, the incidence rates have been increasing, because of a reported increase in the use of the PSA test for screening (17). Likewise, PCa incidence and mortality rates continue to increase in Eastern European countries, particularly in Bulgaria. In Southern Europe, Slovenia has the highest incidence rate followed by Spain. Concerning the remaining countries, both incidence and mortality rates have stabilized, except for Italy, which has shown decreasing trends (7).

1.2.2. Prostate cancer burden in Portugal

In Portugal, PCa was the most diagnosed tumour in 2020 among men, and each year 5.000 new patients and 1.880 deaths are estimated. The most recent data indicates 5741 new cases (21.2 % of the total of malignant tumours in men) and a wASR of 53.2/100000 only in 2018 (19). The geographical variability in the incidence and mortality rates in this country reflects differences in the access to health infrastructures, namely Oncological Institutes (20). Evidence shows that because of the higher literacy levels and easier access to medical services, men living in the coastal regions of the Portuguese mainland have a lower risk of advanced PCa, better survival, and lower mortality rates than men living in rural areas (7, 21). For instance, the highest incidence rates were found in the Metropolitan Areas of Lisbon and

Alentejo (151.6 and 135.5/100000, respectively) (19). In 2018, the Autonomous Region of Madeira (RAM) recorded 160 new PCa cases (26.8 % of the total malignant tumours in men) and one of the highest incidence rates in the country (134.9/100000) (22). Despite the high incidence, the mortality rates recorded between 2010 and 2019 were almost constant (an average of 24.2 ± 3,7/100 000) and lower than the country's average (an average of 36.5 ± 1.5/100 000) (23) (Figure 3). In 2020, RAM recorded 139 new cases (10.8 % of the total diagnoses). This tumour is rare among younger men (< 65 years old) and the highest incidence rate is seen in men between 75 and 84 years old. Moreover, 65 % of the tumours with a known disease status correspond to clinical stage II, while 25 % of these cases are advanced PCa (stages III or IV). The statistics also show that 80.9 % of men diagnosed with stage IV died because of the tumor itself.

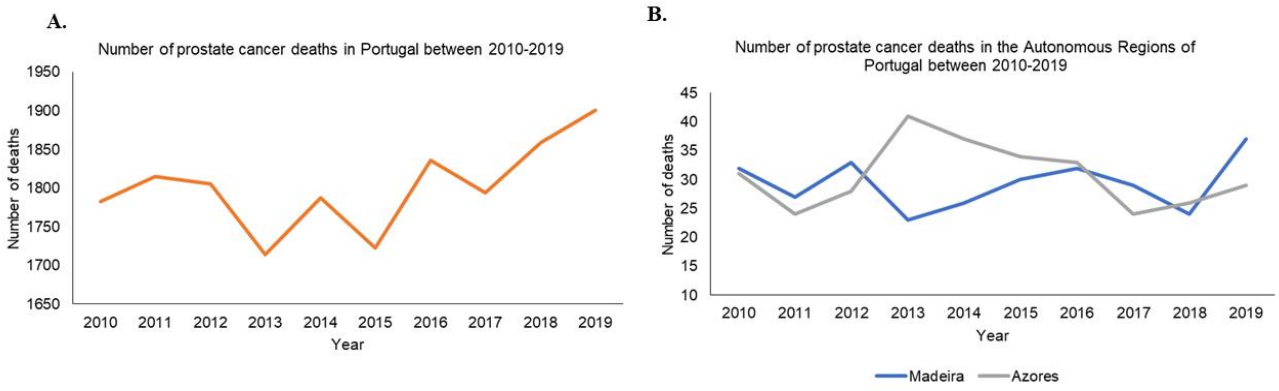


Figure 3 – A. Number of PCa deaths from 2010 to 2019 in Portugal. **B.** Number of PCa deaths from 2010 to 2019 in the Autonomous Regions of Azores and Madeira. Source: Instituto Nacional de Estatística, 2022 (23).

1.3. Prostate cancer risk factors

The well-established PCa risk factors are advancing age, ethnicity (Black race), certain genetic mutations, insulin-like growth factors (IGF), and family history of this malignancy (Table 1) (7). Lifestyle including diet, tobacco and alcohol consumption, obesity and physical inactivity, and environmental factors, such as exposure to chemicals or ionizing radiation, may also increase the risk of advanced PCa (1, 7, 16).

Table 1 – PCa risk factors and their role in the development of this cancer (articles from the last 5 years).

Risk factor	Role in PCa	Reference
Ethnicity	PCa incidence, morbidity, and mortality rates vary significantly by race and ethnicity. African-American, Black, and Caribbean men show the highest PCa rates worldwide. These disparities are mostly related to differences in access to screening and treatment, exposure to PCa risk factors, and variation in genomic susceptibility (e.g.: risk loci found at chromosome 8q24), among other biological factors.	Tonon *et al.*, 2019 (8) Rebbeck, 2018 (24) Taitt, 2018 (25) McAllister, 2019 (26) Brown *et al.*, 2018 (27) Grossman *et al*., 2018 (28)
Family history and genetic factors	The proportion of PCa attributed to hereditary factors has been estimated to be around 5 to 15 %. About 170 susceptibility loci for hereditary PCa (about 33 % of familial PCa risks) have been identified in genome-wide association studies. Many genes have a clear association with hereditary PCa risk, including *BRCA1*, *BRCA2*, *ATM*, *CHEK2*, and *PALB2*, and Lynch syndrome *MLH1*, *MSH2*, *MSH6*, and *PMS2* genes. However, other genes have a poorly defined cancer risk with unknown clinical significance.	Vietri *et al*., 2021 (29) Brandão *et al*., 2020 (30) Bhanji *et al*., 2021(31) Bree *et al*., 2021 (32) Coughlin *et al*., 2021 (33) Grossman *et al*., 2018 (28) Ventimiglia *et al.*, 2017 (34)
Obesity, overweight and physical inactivity	Obesity is implicated in the dysregulation of various hormonal pathways, leading to higher levels of insulin and IGF, oxidative stress, and inflammatory cytokines, and lower levels of adiponectin, testosterone, and sex hormone-binding globulin. Obesity is associated with an increased risk of PCa mortality and recurrence, worsened treatment-related adverse effects, development of obesity-related comorbidities, and earlier progression and development of metastatic disease. Nevertheless, the physiological mechanisms associated between obesity and poor PCa outcomes remain unknown.	Vidal *et al.*, 2017 (35) Wilson *et al*., 2022 (36) Adesunloye 2021 (37) Bandini *et al*., 2017 (38) Fujita *et al.*, 2019 (39) Wilson & Mucci, 2019 (40) Pernar *et al*., 2018 (41) Kaiser *et al.*, 2019 (42)
Tobacco use	Smoking increases the risk of death from PCa, which increases with obesity, specifically for advanced PCa. Moreover, tobacco smoking increases the risk of biochemical recurrence and metastasis. Nevertheless, the association between tobacco smoking and PCa prognosis needs to be explored.	Darcey & Boyle, 2018 (43) Sato *et al*., 2020 (44) Khan *et al*., 2019 (45) Jochems *et al.*, 2022 (46) Foerster *et al*., 2018 (47) Wilson & Mucci, 2019 (40) Pernar *et al*., 2018 (41)

Table 1 – *Continuation.*

Risk factor	Role in PCa	Reference
Lycopene and tomato-based products	Epidemiologic studies have focused on tomatoes as a specific source of lycopene with more consistent findings supporting a protective effect of a higher intake of tomatoes on PCa risk. Furthermore, studies have shown a reduced risk of advanced PCa with the consumption of cooked tomatoes, since these products have more available lycopene. Current epidemiologic evidence is not definitive but suggests that a higher intake of tomato-based products is associated with a reduced risk of PCa and a potentially lower risk of progression. Further studies are required to determine whether the effect is because of lycopene or other components of tomatoes.	Wilson & Mucci, 2019 (40) Pernar *et al*., 2018 (41) Fraser *et al.*, 2020 (48) Li *et al.*, 2021 (49) Soares *et al.*, 2019 (50) Puah *et al*., 2021 (51) Oczkowski *et al.*, 2021 (52) Beynon *et al.*, 2019 (53)
Calcium, dairy products, and vitamin D	A high intake of dairy products, higher than the daily recommended dose, has been positively associated with PCa risk. A potential mechanism underlying the association with calcium is through suppressing circulating levels of dihydroxyvitamin D, which seems to have a protective role against PCa. The mechanisms behind this association are not yet fully understood, but researchers suggest reducing dairy intake while increasing consumption of fish and tomato products for PCa prevention.	Wilson & Mucci, 2019 (40) Pernar *et al*., 2018 (41) Maksymchuk & Kashuba, 2020 (54) Capiod *et al*., 2018 (55) Grant, 2020 (56) Ardura *et al.*, 2020 (57)
Cruciferous, soy, and green tea	Cruciferous, soy, and green tea seem to have a role in decreased risk of PCa due to compounds with anticarcinogenic properties in their composition. Asian populations consume soy foods as a part of their regular diet, which might contribute to the lower PCa incidence found in these countries. However, the preventive action of these compounds needs to be further explored.	Applegate *et al*., 2018 (58) Wilson & Mucci, 2019 (40) Tsugane, 2020 (59) Oczkowski *et al.*, 2021 (52) Rogovskii *et al.*, 2019 (60) Musial *et al.*, 2020 (61) Miyata *et al.*, 2019 (62) Ferreira *et al.*, 2018 (63)

Lifestyle factors are modifiable and may provide an effective method for reducing cancer risk (Figure 4). According to the World Health Organization (WHO), 30 to 50 % of cancers are preventable by healthy lifestyle choices, such as avoidance of tobacco and alcohol consumption, and public health measures, like immunization against cancer-causing infections (7, 16, 40, 41). Modifiable risk factors influence the risk of developing PCa and of dying from the disease (18). Changes in the metabolic profile caused by metabolic disorders such as obesity, insulin resistance, and changes in the hormonal profile are often associated with PCa and some conditions can lead to more aggressive tumours. For instance, smoking, obesity, and some nutritional factors seem to increase the risk of advanced PCa. Furthermore, epidemiological studies suggest that obesity and excessive calorie intake are associated with a higher PCa incidence, whereas a low-fat diet inhibits the growth of cancer cells, due to decreased plasma insulin, hormones, IGF-1 levels, and increased apoptosis (40-43). Some nutrients seem to have a role in PCa pathogenesis and progression, through inflammation, antioxidants, and sex hormones. However, the research about nutrient intake and PCa needs to be further elucidated and extensively studied to better understand how some populations can change their dietary habits to prevent cancer growth (40, 41, 64).

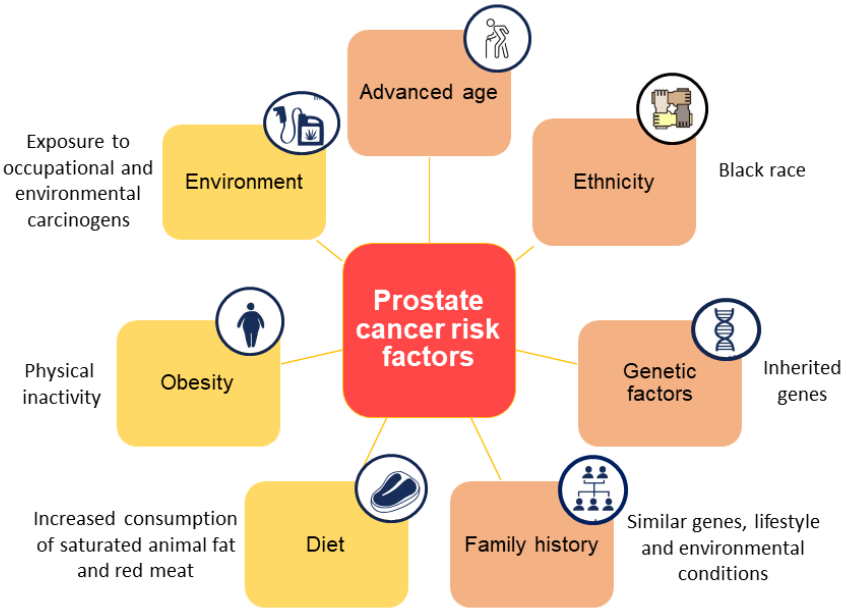


Figure 4 – Modifiable (in yellow) and unmodifiable (in orange) PCa risk factors.

1.3.1. Ethnicity

African-American, Caribbean, and Black men in Europe have the highest incidence of PCa and are more likely to develop the disease earlier in life when compared to other racial and ethnic groups (25, 65, 66). Studies have shown that individuals with African ancestry have a higher chance of dying from PCa than any other racial/ethnic group (7, 8, 26). For instance, men with a family history of PCa and of African descent are more likely to be diagnosed with this cancer and die of PCa at an earlier age (67). These individuals possess a common genetic background more prone to the development of cancer, such as specific genes (e.g.: chromosome 8q24) that are more susceptible to mutation (Table 1) (1, 24, 68). Moreover, African-American men have the highest incidence rates (about twice as White men) and a more aggressive type of PCa when compared to White men (1, 14, 69).

1.3.2. Family history and genetic factors

The prevalence of family PCa is estimated to be around 20 %, while the rate of inherited PCa is about 5 % to 15 % (18, 29). The presence of similar genes, similar lifestyles, and similar environmental conditions are among the reasons associated with family PCa. In men over 75 years old with a family history, the chances of having PCa are 30 to 60 % (70). Inherited PCa occurs when a gene mutation is transmitted from one generation to the next, when at least 3 of their first-degree relatives are affected by PCa, or when three or two generations of a family, or more close relatives (such as father, brother, son, grandfather, uncle, and nephew), are affected by this cancer (29, 30). The presence of first-degree family members younger than 60 years old at diagnosis increases the risk of PCa by 2.5, while for two or more relatives the risk is 5.7 (18, 34). Some cancer predisposition genes have been identified to affect the risk of PCa, including hereditary mutation of *HOXB13* as well as *BRCA1*, *BRCA2*, *ATM*, *CHEK2*, and *PALB2*, and Lynch syndrome *MLH1*, *MSH2*, *MSH6*, and *PMS2* genes (Table 1) (29). Other genes have a poorly defined cancer risk with unknown clinical significance. Nevertheless, the genetics behind family and hereditary PCa remains complex (18, 29, 30).

1.3.3. Dietary factors

Dietary factors are associated with worldwide and ethnic differences in PCa incidence rates (1, 40, 41). Dietary factors are modifiable, and the identification of nutrients that regulate the risk for advanced PCa represents a practical strategy for its primary prevention (40, 71).

Diets containing vegetables, such as tomatoes, cruciferous, and soybeans, have been associated with a lower risk of developing PCa (40, 52, 58). Cruciferous or *Brassica* vegetables (broccoli, Brussels sprouts, cauliflower, and cabbage) have anticancer properties mediated by phenylethyl isothiocyanate, sulforaphane, phytochemicals, and indole-3-carbinol (63). However, further research is still needed to deeply understand the anticancer capabilities of *Brassica* vegetables (1). The potential chemo-preventive action of soy and green tea, part of the normal diet in Asia, has also been studied (Table 1). A decreased risk of PCa and several other cancers has been seen with the consumption of soy and green tea (58-60). Catechins found in green tea and isoflavones in soybeans have anticarcinogenic properties, with the potential to inhibit different phases of carcinogenesis and metastasis (59-62). Additionally, lycopene found in red fruits and vegetables, mainly tomatoes, has been described to possess potent antioxidant properties, as well as cancer-preventive effects, through decreased lipid peroxidation and inhibition of cell proliferation (48-50, 71). Lycopene acts on the androgen receptors and reverses the effects of dihydrotestosterone (72). Even though tomato consumption and lycopene intake both seem to be associated with a decreased risk of PCa (50, 51, 53), evidence of a connection between tomatoes and PCa risk requires further investigation, since some studies could not show any clinical benefit of lycopene intake (1, 73).

Several epidemiological studies have shown a positive correlation between PCa mortality and *per capita* intake of meat, fat, and dairy products (40-42, 71). Saturated fat intake increases oxidative stress and reactive oxygen species (ROS) levels, causing lipid peroxidation and deoxyribonucleic acid (DNA) damage. Aronson *et al.* (74) found that a high intake of total fat, particularly saturated fat, was associated with a statistically significant increase in PCa risk. A few biological mechanisms that are thought to be involved between saturated animal fat intake and PCa risk include the promotion of prostate carcinogenesis through androgen signalling, increased levels of ROS, leukotrienes, and prostaglandins from lipid metabolism, along with increased basal metabolism, IGFs levels, and tumour proliferation (1). Rohrmann *et al.* (75) showed that men consuming five or more servings of processed meat per week had a higher risk of PCa when compared with men who consumed one or fewer servings (76). Consumption of trans and saturated fatty acids, as well as processed meat, contributes to

prostate carcinogenesis through the alteration of prostate hormonal regulation, induction of oxidative inflammation and stress, alteration of growth factor signalling, and lipid metabolism (77). Furthermore, cooking at high temperatures (125 – 300 °C) causes the formation of aromatic hydrocarbons and mutagenic heterocyclic amines, while grilled or barbecued meat can lead to the formation of N-nitroso compounds, and result in lipid peroxidation and DNA damage through the production of free radicals (1, 78). Yet, there are some inconsistent associations between the intake of meat and meat products and the risk of advanced PCa (71). Low fats like corn oil may also promote cancer growth via linoleic acid, the most abundant Omega-6 fat in the oil. Arachidonic acid, a metabolite of linoleic acid, gives rise to the formation of pro-inflammatory prostaglandins (PG), including PGE2, involved in the promotion of cell proliferation, and 5-hydroxyeicosatetraenoic acid, which is found to be increasingly expressed in malignant PCa (40-42). This suggests that Omega-6 fats have a pro-inflammatory effect and a decrease in their intake can decrease cancer growth. On the contrary, Omega-3 fats are beneficial against cancer growth (1, 79).

Vitamin D (dihydroxyvitamin D3) is an essential vitamin that belongs to the family of steroid hormones. Human sources include both dietary intake and sunlight exposure (14, 40, 41). Vitamin D and its analogues seem to protect from PCa, through inhibition of cell proliferation and invasion (Table 1). For instance, several epidemiological studies suggest that PCa occurs more frequently in older men with vitamin D deficiency (1, 56, 80). Vitamin D deficiency may also have a role in the incidence of PCa in African-American men since ultraviolet (UV) radiation is blocked in darkly pigmented skin due to high melanin levels, which in turn inhibits the conversion to vitamin D3 (81). Moreover, a high dietary intake of dairy products rich in calcium, higher than the daily recommendation, also increases PCa risk, due to decreased serum levels of vitamin D (54, 55, 57, 81).

Folate and vitamin B12 are essential vitamins that participate in DNA methylation, synthesis, and repair. Low availability of these vitamins can lead to altered methylation patterns and, consequently, influence cancer development. Higher concentrations of folate and vitamin B12 have been associated with an increased risk of PCa. For instance, elevated serum concentrations of folate were associated with an increased proliferation of PCa cells, in some prostate samples collected from patients who underwent radical prostatectomy (RP) (82). *In vitro* (83), *in vivo* (84), and genetic studies (85) showed the role of folate in the development of an aggressive form of PCa. However, due to the heterogeneous nature of PCa, some studies could not show the effect of consuming folate on disease progression or survival (1, 86), and

thus the association of folate and vitamin B12 with PCa remains unclear and requires further investigation (87).

1.3.4. Unhealthy lifestyles - obesity, insulin, and tobacco use

Obese men have a greater circulating plasma volume and clinical studies have demonstrated that obesity might have clinical implications for disease detection and management. Obese men show alterations in circulating levels of metabolic and sex steroid hormones, both known to be involved in prostate development and oncogenesis. In turn, obesity makes it difficult to interpret the PSA results (35, 36, 88). A high body mass index (BMI) is associated with worse outcomes and obesity is often linked to advanced and aggressive PCa forms (36, 89). Insulin is a risk factor to promote PCa initiation and/or progression. This hormone promotes the growth and proliferation of cancer cells, given the high expression of insulin receptors in various types of cancer cells. For instance, elevated circulating insulin concentrations were found in aggressive PCa tumours, supporting the role of insulin in PCa growth (90). Tobacco consumption is another PCa risk factor (43, 45, 46) (Table 1). The incidence and mortality rates of PCa have increased significantly with the increase in tobacco use, due to exposure to carcinogens and alterations in circulating levels of hormones (91). Functional polymorphisms in genes involved in the polycyclic aromatic hydrocarbons (PAHs) metabolism, one of the carcinogenic chemicals of cigarette smoke, may affect cancer onset and progression (1). Researchers found that smoking increases the metabolism of serum estrogen, which is involved in a more aggressive tumour phenotype, resulting in increased PCa-related deaths (92). Moreover, cigarette smoking has been associated with adverse pathological features and worse oncological control (18).

1.3.5. Occupation

Exposure to chemicals, such as acetic acid, arsenic compounds, PAHs, petrol and diesel engine emissions, mononuclear aromatic hydrocarbons, organochloride pesticides, chromium, and lubricating oils, were found to significantly increase the risk of PCa (93). A study in 2016 showed that the risk among farmers was four times higher (94), because of the exposure to chemicals found in pesticides, when compared to other occupations, such as forestry workers, police officers, and administrative and clerical workers (95).

1.4. Prostate cancer screening

Screening for PCa is based on the PSA biomarker values in blood serum (> 4.0 ng/mL), and DRE. After suspicion, a magnetic resonance imaging (MRI) scan is usually performed, which indicates whether a prostate biopsy should be performed, considering the prostate imaging-reporting and data system (PI-RADS) value (PI-RADS > 3). Following histological confirmation (biopsy) of malignant neoplasia, staging tests are performed, through imaging techniques such as computed tomography (CT) or positron emission tomography (PET). In turn, the result of these tests dictates the patient's therapy based on a combination of surgical strategies, hormone therapy, radiotherapy, and chemotherapy (Figure 5) (96).

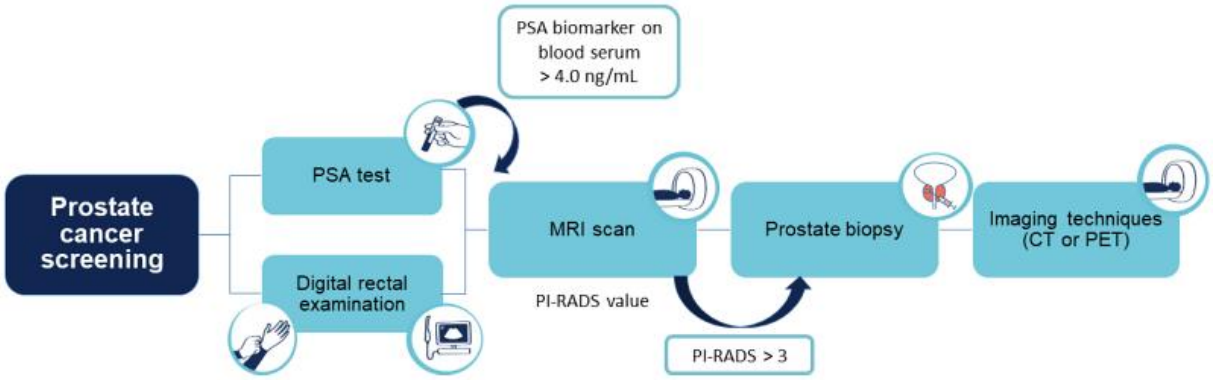


Figure 5 – PCa diagnosis pathway.

PSA is a glycoprotein normally expressed by the prostate tissue with a cut-off of 4.0 ng/mL (7). However, this test shows low selectivity to detect PCa and monitor the disease's progression (97), due to its limited sensitivity (20.5 %) (98), accuracy (62-75 %) (99), and specificity (51-91 %) (100). PSA screening cannot differentiate patients in terms of the aggressiveness of the tumour (101) and cannot distinguish between benign prostatic hyperplasia or prostatitis (11). Furthermore, PSA levels may be affected by several other factors, such as age, BMI, and urinary tract infection, leading to false-positive results (98, 102). Due to concerns about overdiagnosis and overtreatment, along with the high rate of false-positive results, the United States Preventive Services Task Force made recommendations against PSA testing among men over 70 years old (10, 97). This decision resulted in a decline in the incidence of PCa from 2007 to 2014. Between 2013 and 2017 the mortality rates flattened, most likely because of a decline in the use of PSA, which consequently resulted in the diagnosis of more men with metastatic PCa (97). Therefore, it has become very important that men are fully informed of the potential benefits and harms of PSA screening (103).

A decisive diagnosis of PCa is based on a prostate biopsy when PSA and DRE show abnormal results (104, 105). Besides being an invasive, unpleasant, and potentially harmful procedure (106), prostate biopsies also show the risk of severe infection, due to the introduction of rectal commensal or other bacteria through the needle into the sterile prostate (107). Moreover, this procedure can still lead to both false-positive and false-negative results (1, 108, 109). False-negative results may occur when the tumour is small, when the cancer cells are distributed heterogeneously, and in early PCa stages when, histologically, the tumour appears benign. Accordingly, the samples obtained during the biopsy may not be representative of cancer. Another issue is the overdiagnosis and overtreatment of relatively indolent tumours with low potential for morbidity or death if left untreated (13, 110). Hence, serum PSA levels and prostate biopsy histology have a very limited accuracy in predicting the clinical behaviour of individual tumors, especially the ones prone to become aggressive at a later stage. Several studies have focused on the development of new methods to overcome these limitations and provide more accurate tools for PCa detection and management (Table 2).

Table 2 – Emerging diagnostic methods for PCa detection and management (articles from the last 5 years).

Method	Evidence/Aim	Reference
PSMA radioligand targeted therapy and molecular imaging	Evidence: Molecular imaging techniques detect PCa lesions that are occult on anatomic imaging. PSMA radioligand therapy shows promising response rates with low toxicity in extensively pre-treated patients with PCa. Aim: Theragnostic applications: diagnosis, management, and treatment of metastatic PCa.	Parsi *et al.*, 2021 (111) Uijen *et al*., 2021 (112) Seifert *et al*., 2021 (113) Hofman *et al*., 2018 (114) Hofman *et al*., 2019 (115) Calais *et al.*, 2019 (116) Moradi *et al.*, 2021 (117) Mena *et al.*, 2020 (118) Thomas *et al.*, 2017 (119)
Evs	Evidence: EVs can mediate PCa progression and metastasis. EVs have great potential to be used as liquid biopsy biomarkers in the diagnosis of PCa. EVs can be used in risk stratification and to predict the response to hormonal, chemo-, immune- and targeted-therapy. Aim: Diagnosis and treatment. Can be used to personalize and guide treatment decisions.	Campos-Fernández *et al*., 2019 (109) Kim *et al*., 2021 (97) Gaglani *et al*., 2021 (107) Oey *et al*., 2021 (120) Ludwig *et al.*, 2021 (121) Lorenc *et al.*, 2020 (122) McKiernan *et al*., 2016 (123) Ji *et al.*, 2021 (124)

Table 2 – *Continuation.*

Method	Evidence/Aim	Reference
lncRNAs (*PCA3, MALAT1, SChLAP1, BDNF-AS, FALEC)*	Evidence: lncRNAs provide new insights into cancer signalling networks, along with novel strategies and methods for PCa diagnosis and treatment. lncRNAs analysis has the potential to improve the specificity and sensitivity of existing biomarkers. Aim: Novel biomarkers (predictive, diagnostic, prognostic) and therapeutic targets.	Xu *et al*., 2019 (125) Li *et al*., 2018 (126) Zhao *et al*., 2017 (127) Hu *et al.*, 2021 (128) Li *et al.*, 2021 (129) Kidd *et al.*, 2021 (130) Huang *et al.*, 2018 (131)

Legend: EVs: extracellular vesicles; lncRNAs: long non-coding ribonucleic acids (RNAs); PSMA: prostate-specific membrane antigen.

1.4.1. Prostate-specific membrane antigen: a theranostic approach

Imaging methods are used to define the stage of PCa and so guide its management. However, PCa's more aggressive forms can manifest rapid growth with progression to adjacent organs and spread to lymph nodes and bones (1, 3, 4), and CT, bone scan, and MRI have limited performance in the detection of lymph node metastasis (111). Patients with castration-resistant PCa (CRPCa) have a 90 to 95 % probability of developing bone metastases, which leads to severe morbidity, including bone pain, pathological fractures, spinal cord compression, and haematological consequences of bone marrow infiltration (132-134). Due to the importance of bone metastases in the overall disease progression, bone-targeted therapy constitutes an essential part of the treatment of CRPCa (135). A possible therapy can be based on the use of radiopharmaceuticals systemically administered to slow or reverse the bone metastatic progression (134).

Current research is focused on molecular targeting of prostate-specific membrane antigen (PSMA) as a theranostic approach, to diagnose, monitor, and treat PCa (111). PMSA is a transmembrane enzymatic protein found on most PCa cells and its overexpression correlates to adverse factors, such as androgen independence, metastasis, and progression, making PSMA an antigenic marker for PCa progression (111, 112, 134, 135). Hence, PMSA can be used for diagnostic and therapeutic purposes, and several clinical trials have been investigating its effectiveness as a diagnostic tool and for direct radioligand therapy (Table 2) (111).

1.4.1.1. Molecular imaging

PSMA scans can detect metastatic lesions that are missed by conventional imaging techniques (111), and so, small molecules, antibodies, and antibody fragments that target PSMA have been created, radiolabeled, and used for molecular imaging (117).

PET is emerging as a highly sensitive molecular imaging technique in the detection and localization of primary PCa. PET uses a positron emitter to label key molecules that are intravenously injected, and their distribution and uptake images provide insights into metabolic changes associated with cancer (136). This technique has been reported as a valuable tool in the diagnosis of PCa patients with negative MRI and systematic biopsies (117). Recently, ligands of PSMA were introduced in PET to diagnose and manage PCa (reviewed by Mena *et al.*, 2020 (118)). This approach can improve PCa detection by identifying lesions that are not visible on MRI and also provides better estimates of tumour volume (117). PSMA-PET can be used in the diagnosis, staging, and management of PCa patients (118). PSMA-PET has an important role in the initial staging of PCa, superior diagnostic performance to anatomical imaging, and enhanced sensitivity to detect node metastasis (reaching 99 % (136)), outperforming other molecules imaging techniques, including PET-CT (117, 118). Furthermore, PSMA-PET can be combined with anatomical CT (PET/CT) and MRI (PET/MRI) images for the detection of bone metastases (118, 119) (Table 2). PSMA-PET/MRI consistently outperforms multiparametric MRI (mpMRI) for detection or localization of PCa in intermediate or high-risk PCa patients (reviewed by Moradi *et al.*, 2021 (117) and Mena *et al.*, 2020 (118)). PSMA-PET/CT has greater sensitivity in the detection of bone metastasis when compared to whole-body bone scans (119), and has shown the most utility in biochemical recurrence (136). PSMA-PET/CT was first coupled with gallium-68 (^{68}Ga) and is considered the most sensitive and specific method of staging high-risk PCa and imaging recurrent PCa (111, 117). Moreover, ^{68}Ga-labeled ligands have shown higher sensitivity and specificity in the diagnosis of primary and recurrent PCa (119). In a retrospective analysis, Maurer *et al.* (137) investigated the diagnostic efficacy of ^{68}Ga-PSMA-PET for lymph node staging in patients with PCa and compared it to CT and MRI imaging. In their analyses, ^{68}Ga-PSMA-11 showed sensitivity, specificity, and accuracy levels of 65.9 %, 98.9 %, and 88.5 %, respectively, for the detection of nodal metastases, compared with 43.9 %, 85.4 %, and 72.3 % for morphological imaging (137). In another study, Thomas *et al*. (119) investigated the difference between technetium-99m (^{99m}Tc)-MDP bone scan and ^{68}Ga-PSMA-PET/CT for the detection of bone metastases in PCa. The authors compared the number of identified lesions and found that the

PSMA-PET/CT method detected twice the number of lesions, especially in the thorax and pelvis. Their results suggested that when patients go through ^{68}Ga-PSMA-PET/CT, the bone scan is not mandatory (119).

1.4.1.2. Radioligand targeted therapy

Recent studies suggest that newer molecular theragnostic approaches, based on PSMA radioligands, have the potential to provide even more effective and personalized treatment options for both diagnostic, prognostic, and therapeutic applications in patients with CRPCa, with fewer toxicities and adverse effects (111-113). This approach has been developed to select patients and delivers irradiation to all tumour sites, including osseous, nodal, and visceral metastases (111). PSMA radioligand therapy uses small-molecule inhibitors of PSMA, usually labelled with beta and alpha-emitting radionuclides that emit cytotoxic radioactive decay (111, 112). Alpha and beta radionuclides differ in energy, tissue range, linear energy transfer, and the number of DNA hits needed for cell destruction (134). These radiopharmaceuticals deliver targeted irradiation to the active bone turnover sites, where metastatic infiltration and destruction are happening. This approach can simultaneously treat multiple sites of disease, ease administration, and can be integrated or combined with other treatments. Alpha-emitters include actinium-225 (^{225}Ac), thorium-227 (^{227}Th), radium-223 (^{223}Ra), and astatine-211 (^{211}At). Recently, ^{223}Ra was approved to treat bone metastases from PCa. This authorization follows the symptomatic relief and significant improvement in the overall survival of CRPCa with predominant bone metastases that ^{223}Ra was shown to elicit (138). Beta-emitting radiopharmaceuticals, including lutetium-117 (^{177}Lu), strontium-89 (^{89}Sr), samarium-153 (^{153}Sm), and rhenium-186 (^{186}Re), have been used for bone palliation. ^{177}Lu is the most used beta-emitter, due to its favourable safety profile, short range of emission, and relatively long half-life, allowing the delivery of a high degree of radiation to specific lesions (111). For instance, [^{177}Lu]Lu-PSMA-617 shows a favourable safety profile due to reduced kidney uptake and has demonstrated promising results in prospective trials with high response rates, low toxic effects, and reduction of pain in men with metastatic CRPCa who progressed after standard treatments (114-116). In general, radioligand therapy shows promising response rates with low toxicity in extensively pretreated patients with PCa (111). While most of these studies remain experimental and the effects of this therapy on overall survival and safety are yet to be determined, their clinical observations are very promising (114, 135, 139-141).

PSMA targeted imaging and therapy have proven to be excellent diagnostic and therapeutic options for metastatic PCa, but further studies are still required to determine the effect of this approach on overall survival and safety. Moreover, current research is still ongoing regarding the exact role of PSMA in various stages of PCa care (111).

1.4.2. Tumour biomarkers

In recent years, advances in molecular medicine, including the OMICs genomics, proteomics, transcriptomics, and lipidomics contributed to the discovery of new potential biomarkers to aid in PCa screening and management. Common liquid biopsies biomarkers include extracellular vesicles (EVs), circulating tumour cells (CTCs) and DNA (ctDNA), and cell-free DNA (cfDNA) (142). Long noncoding ribonucleic acids (lncRNAs) in addition to molecular biomarkers for urine, serum, and tissue samples, also emerged as promising PCa biomarkers (Table 2).

1.4.2.1. Molecular biomarkers

Molecular biomarkers for urine, serum, and tissue samples have been developed (Table 3), based on the combination of imaging techniques with other methodologies such as gene or protein profiling, to enhance cancer detection, pre-biopsy decision, determination of cancer risk, and therapeutic management of PCa (143). Abnormalities in these tests indicate the performance of a prostate biopsy. Moreover, risk calculators (RCs) are combined with these tests to help determine the risk of cancer for each individual, reducing the number of unnecessary biopsies. The guidelines recommend using these tests in combination with the current PCa screening methods (98).

Table 3 - Potential clinical utility, characteristics, and availability of PCa biomarkers.

Biomarker test	Molecular markers	Potential clinical utility	Characteristics	Availability
		Serum biomarkers		
PSA	PSA	Treatment monitoring	Sens: 60 % (144) Spec: 79 % (144) AUC: 0.55 (145)	
4Kscore	Total PSA, free PSA, intact PSA, hK2	Unnecessary biopsy reduction of 43 % (146)	Sens: 75 % (147) Spec: 63 % (147) AUC: 0.71 (148)	CLIA-certified

Table 3 – *Continuation.*

Biomarker test	Molecular markers	Potential clinical utility	Characteristics	Availability
		Risk prediction of PCa metastases Previous negative biopsy		
PHI	Total PSA, free PSA, p2PSA isoform	Unnecessary biopsy reduction of 40 % (149) Prediction of high-grade PCa Active supervision monitoring	Sens: 82 % (150) Spec: 80 % (150) AUC: 0.71 (147)	FDA-approved
		Urinary biomarkers		
Progensa (PCA3)	lncRNAs (ratio of PCA3 mRNA:PSA mRNA)	Unnecessary biopsy reduction of 23-38 % (151) PCa detection, staging and prognosis Previous negative biopsy	Sens: 69 % (152) Spec: 65 % (152) AUC: 0.73 (152)	FDA-approved
SelectMDx	HOXC6 and DLX1 mRNA	Unnecessary biopsy reduction of 53 % (153) Prediction of high-grade PCa	Sens: 91 % (153) Spec: 36 % (153) AUC: 0.71-0.81 (153)	CLIA-certified
MPS	PCA3 and TMPRSS2-ERG mRNA	Unnecessary biopsy reduction of 35-47 % (154) Predict the risk of PCa and high-grade PCa	Sens: 93 % (155) Spec: 33 % (155) AUC: 0.69 (154)	CLIA-certified
EPI	Exosomal RNA (SPDEF, PCA3, ERG)	Unnecessary biopsy reduction of 27 % (123) Improved identification of high-grade PCa	Sens: 92 % (123) Spec: 34 % (123) AUC: 0.70 (123)	CLIA-certified
		Tissue biomarkers		
ConfirmMDx	DNA hypermethylation GSTP1, APC, and RASSF1	Prediction of true negative prostate biopsies	Sens: 68 % (153) Spec: 64 % (153) AUC: 0.74 (153)	Not FDA-approved yet
OncotypeDX	mRNA expression (17 genes)	Monitoring of tumour aggressiveness	AUC: 0.73 (155)	Not FDA-approved yet
Prolaris	mRNA expression (31 genes)	Monitoring of tumour aggressiveness	AUC: 0.78 (155)	FDA-approved
Decipher	mRNA expression (22 genes)	Treatment monitoring	Sens: 73 % (156) Spec: 74 % (156) AUC: 0.79 (156)	CLIA-certified
ProMark	Protein biomarker test (8 proteins)	Monitoring of tumour aggressiveness	Sens: 90 % (153) Spec: 85 % (153) AUC: 0.72 (153)	CLIA-certified

Legend: AUC – area under the receiver operating characteristic (ROC) curve; CLIA - Clinical Laboratory Improvement Amendments; FDA – Food & Drug Administration; Sens: Sensitivity; Spec: Specificity.

Biomarkers such as the non-coding RNA (PCA3) and TMPRSS2-ERG fusion gene, have shown increased sensitivity and specificity (Table 3), potentially reducing PCa overdiagnosis. The Progensa prostate cancer antigene 3 (PCA3) assay detects lncRNAs in urine samples by detecting the PCa gene 3 transcript levels. This test was approved by the Food and Drug Administration (FDA) in 2012 and calculates the ratio of PCA3 messenger RNAs (mRNAs) *versus* (*vs*) PSA mRNA in the first urine post-DRE and is approved for patients with a previous negative biopsy (11, 98, 157, 158). The Prostate Health Index (PHI) test is an algorithm approved by the FDA, which combines total PSA, free PSA, and p2PSA isoform ([-2]proPSA). PHI calculates PCa probability and is recommended for men with PSA levels between 2 and 10 ng/mL and no abnormalities in their DRE. This blood test is also able to assess the likelihood of PCa progression during active surveillance, being used to monitor patients (159). The Four-Kallikrein (4KScore) test is a diagnostic algorithm that combines four kallikreins in blood plasma, including human kallikrein 2 (hK2), total PSA, free PSA, and intact PSA, in addition to the patient's clinical information (age, DRE results, and prostate biopsy history). This test assesses the probability of high-grade PCa and is recommended for patients undergoing initial and repeated biopsy. Moreover, 4KScore also predicts the risk of occurrence and development of aggressive PCa (159). ExoDx Prostate IntelliScore (EPI) is a pre-biopsy RNA-based assay that uses the expression of PCA3, ERG, and SPDEF, isolated from urinary exosomes to predict the probability of high-grade PCa (Gleason score ≥ 7) on diagnostic biopsy. This is the only test that is not based on any other parameters, related to PSA or a PSA derivative in the test algorithm, to calculate the result, but clinicians can use it in conjugation with other clinical variables (123, 159). SelectMDx and MyProstateScore (MPS) tests are based on the combination of multiple gene analyses. SelectMDx is a non-invasive test that measures mRNA transcripts from the genes HOXC6 and DLX1 in urine samples post-DRE and relates them with clinical risk factors such as age, family history, and PSA levels. This test is used to evaluate the presence of any PCa during biopsy and the risk of high-grade PCa. It also avoids unnecessary biopsies in the case of low-risk PCa (153). The MPS assay also requires the collection of urine post-DRE and is based on combinations of multiple gene analyses, including total serum PSA, the PCA3 assay, and the expression of the TMPRSS2:ERG fusion gene (160, 161). Serum and urine biomarkers are used for consideration of initial biopsy, while tissue biomarkers are used to confirm test results. Tissue biomarkers tests have been developed to aid the clinical practice to decide what kind of therapy should be applied for different PCa diagnoses. ConfirmMDx is based on pronounced epigenetic changes, indicative of the presence of cancer, in the benign prostate tissue that is near the focus of PCa. This test determines the

level of methylation of the promoter regions of the genes GSTP1, APC, and RASSF1 in benign prostate tissue, identifying high-grade PCa in patients with negative biopsies (153). Prolaris is a prognostic test that measures tumour biology to improve the accuracy of risk stratification in men with localized PCa. This test combines the RNA expression levels of 31 genes involved in cell cycle progression and 15 housekeeping genes to generate a Prolaris Score. Prolaris can be used to guide patient selection for active surveillance or definitive treatment (159, 162, 163). The ProMark test is a protein-based prognostic assay of eight protein markers (DERL1, CUL2, SMAD4, PDSS2, HSPA9, FUS, phosphorylated S6, and YBOX1) that predicts the aggressiveness of cancer in patients with biopsy Gleason scores of 3 + 3 and 3 + 4. Moreover, ProMark predicts adverse pathology during RP and also predicts if the tumour can be managed with or without aggressive treatment (153). The Decipher is a genomic classifier prediction model for metastasis, which measures the levels of RNA expression of 22 different genes on post-prostatectomy tissue samples. This test calculates the likelihood of clinical metastases within 5 years of prostatectomy in men with adverse pathological features. It could be a useful tool for diagnosis and local therapy planning for new PCa patients (156, 159). Despite the recent progress in the discovery of new biomarkers, gene mutations, and genomic signatures, some challenging obstacles must be overcome to develop effective biomarkers. These limitations include tumour heterogeneity, tumour-host interplay, complexity, multiplicity, and redundancy of tumour-cell signalling networks involving genetic, epigenetic, and microenvironmental effects (11). Additionally, the technologies associated with these approaches are often expensive, unavailable in many medical facilities, and time-consuming (98).

1.4.2.2. Long non-coding RNAs

LncRNAs consist of RNA transcripts longer than 200 nucleotides that do not encode proteins. LncRNAs have been reported to exhibit abnormal expression in various types of cancer, including PCa. Most lncRNAs associated with PCa are overexpressed in tumour tissues and cancer cells, contributing to tumour proliferation, invasion, and metastasis. In contrast, only a few lncRNAs are downregulated and may act as tumour suppressors, in addition to their potential activity as transcriptional regulators and oncogenes (125). All these unique features make lncRNAs promising predictive biomarkers and therapeutic targets, for diagnosis, screening, prognosis, and progression of PCa (125) (Table 2). Recent studies have shown that lncRNAs such as PCA3, GAS5, and HOTAIR are related to the occurrence and progression of PCa (125). Currently, PCA3 is one of the most well-studied lncRNAs, given its higher

specificity and sensitivity than the PSA blood test. Moreover, its combination with PSA testing or other biomarkers will significantly improve the sensitivity, specificity, and accuracy of PCa screening and diagnosis. For example, PCA3 combined with TMPRSS2-ERG assays can reduce the number of unnecessary biopsies and increase diagnostic accuracy (125). MALAT1 is another putative diagnostic marker of PCa and its higher expression has been correlated with high PSA levels and Gleason scores as well as with tumour stage and CRPCa (125). Hu *et al.* (128) studied single-nucleotide polymorphisms of *MALAT1* and found that rs619586 and rs1194338 were significantly associated with the susceptibility to both advanced Gleason grade and nodal metastasis in PCa. Li *et al.* (129) developed and validated a noninvasive post-DRE urine assay based on the combination of the lncRNAs PCA3 and MALAT1 for early diagnosis of PCa and high-grade tumors. However, some researchers reported that the PCA3 assay is limited due to intra-individual variability and cannot distinguish high-grade from low-grade diseases. Hence, PCA3 needs further research to determine its optimal application in PCa diagnosis (125). Other lncRNAs may also be promising biomarkers for PCa diagnosis and progression, such as TMPO-AS1 and FALEC (125, 131). Zhao *et al.* (127) characterized the expression profile of lncRNA FALEC in PCa and paired histologically normal tissues and analysed the biological function of FALEC in PCa cell lines. In 85 patients, FALEC expression was significantly increased in clinical PCa tissues compared to adjacent normal tissues. Moreover, the downregulation of FALEC could inhibit cell proliferation, migration, and invasion *in vitro*. The results of this study suggested that FALEC may be a potential diagnostic and therapeutic target in patients with PCa (127). Li *et al.* (126) investigated the expression, prognostic value, and functional role of lncRNA BDNF-AS in PCa and also correlated the expression of BDNF-AS with the clinicopathological factors of patients. The results of this study showed that BDNF-AS was downregulated in PCa patients with poor prognoses and shorter overall survival, demonstrating its potential as a prognostic biomarker of these cases. Furthermore, lncRNAs can be used as predictive of biochemical recurrence. The high expression levels of SChLAP1 in PCa tissue were significantly associated with biochemical recurrence, clinical progression, and PCa-specific mortality (130). Additionally, SChLAP1 can be easily detected in urine, an important feature for the development of a SChLAP1 assay for guide therapy (as reviewed by Xu *et al.* (125)). Given the roles of lncRNAs in PCa, it will be important to develop specific drugs that interfere with malignant signaling networks in which lncRNAs participate, particularly in PCa cells. However, the molecular mechanisms of action of lncRNAs are not very clear yet. Hence, it will be important to fully understand and investigate the roles and mechanisms of lncRNAs in prostate carcinogenesis (125).

1.4.2.3. Liquid biopsy biomarkers

Liquid biopsy has emerged as a complement to invasive tissue biopsy to guide cancer diagnosis and treatment (97). Liquid biopsies rely on the detection of specific biomarkers in readily accessible body fluids, such as blood, serum, or urine (109). The common liquid biopsy biomarkers are EVs, CTCs, ctDNA, and cfDNA, which provide specific information based on their intrinsic characteristics. CTCs are cancer cells from primary and metastatic tumours that are released into the vasculature and circulate through the body to form metastatic niches in other tissues, being detectable in cancer patients only (142). Similarly, ctDNA is a tumour-derived short, fragmented DNA found in the bloodstream, which reflects cancer-related genetic changes. cfDNA or RNA (cfRNA) are cell-free circulating small nucleic acid fragments that are released after the lysis of apoptotic or necrotic cells. cfDNA is detectable in blood and urine samples from patients with cancer, and their analyses improve the evaluation of mutations, polymorphism, methylation, and loss of DNA integrity (97, 109, 164). Numerous studies have shown the relevance of liquid biopsies in PCa screening. cfDNA and EVs seem to have a better application in the diagnosis and prognosis of PCa than CTCs (97, 107, 109, 120) (Table 2). This occurs because early-stage or localized PCa patients have very few CTCs and their use is more effective in the later stages of this cancer (109). The only FDA-approved liquid biopsy test for PCa, CellSearch, is based on the detection of CTC's, and there is no evidence of wide clinical implementation of this technology in medical practice. EVs are nano-sized, double-lipid membrane vesicles, including exosomes and microvesicles, that are secreted from cells and shed into biofluids, including blood and urine (123). EVs are involved in intercellular communication and immune function, through proteins, lipids, mRNA, microRNAs (miRNAs), and DNA, which have been correlated to the presence of cancer for diagnostic purposes (Table 2) (97, 120, 165, 166). Cells exchange proteins, nucleic acids, sugars, and lipids through EVs to induce changes in the recipient cells, which makes EVs potential carriers of cancer biomarkers from tumour cells to other tumour or non-tumour cells (109). EVs can also be used as a vehicle for drugs or nucleic acids with antineoplastic effects (107, 121). The EVs approach may improve the sensitivity of PCa biomarkers, given the protective role of EVs' lipid layer over biomolecules, meaning that the concentration of PCa biomarkers will be higher in EVs (109). Urine is the most used body fluid for the detection of biomarkers in EVs from liquid biopsies of PCa. Moreover, exosomal miRNAs are emerging as promising prognostic biomarkers for metastatic CRPCa patients (109). The concentration of RNA-based biomarkers, particularly miRNA, is higher in EVs than in CTCs from urine samples. Nevertheless, the

application of miRNA as a diagnostic marker has been limited due to a lack of specificity, and in turn, many studies have emerged to investigate EV-mRNA as a diagnostic and prognostic biomarker for PCa management (97). McKiernan *et al*. (123) developed an exosome-derived gene expression signature from normalized PCA3 and ERG RNA from urine predictive of initial biopsy results. Exosomes in post-DRE urine of PCa patients contain both PCA3 and TMPRSS2: ERG mRNA. In their study, the authors were able to develop a molecular signature predictive of PCa combined with serum PSA in a diagnostic test, which was able to discriminate between benign disease, and high- and low-grade tumours, reducing the total number of unnecessary biopsies (123). Ji *et al.* (124) developed a strategy for exosomal mRNA detection based on features of mRNA of circulating exosomes and identified a PCa exosomal mRNA signature for PCa screening and diagnosis. With this strategy, the authors were able to distinguish PCa patients from healthy controls (124). Despite the beneficial properties of EVs for the diagnosis of PCa, their clinical application still presents a few challenging issues (97). EVs are released from all cells in the body, which makes it difficult to determine which EVs are tumour-derived, meaning that new technologies for the specific detection and isolation of tumour-derived EVs need to be developed (97). Recent EVs isolation technologies have been developed to improve isolation performance, yield, purity, usability, hands-on procedures, and processing time (97). However, EVs' isolation is still difficult, especially in EVs from blood plasma, due to the purity and efficiency achieved by laboratory procedures. Moreover, there is no wide clinical application of liquid biopsies of PCa with EVs (109), and automated analysis platforms are yet to be developed for large-scale clinical studies (97). In general, a few issues are preventing the effective use of CTCs and EVs as biomarkers in liquid biopsies for the diagnosis of PCa, such as the need for specific guidelines for the biomarker's isolation and detection. Moreover, the microfluidic devices used for the development of liquid biopsies have not been fully validated and standardized yet (164).

1.4.3. Active surveillance and risk-stratification algorithms

PCa is very heterogeneous in terms of grade, phenotypes, and genetics, displaying complex features (1). This tumour often has indolent growth, not compromising the patient's quality of life, but its diagnosis and subsequent treatments have a high impact on the physical and mental status of patients, significantly affecting their quality of life (11). The main goal of early detection is to identify PCa in a phase where it needs less aggressive treatments with fewer side effects and has a higher chance of cure, even in the cases of locally advanced and metastatic

PCa. Many early diagnoses can be safely managed by active surveillance, preventing overtreatment, and thereby improving or maintaining the patient's quality of life and. avoiding adverse outcomes (167).

Active surveillance consists of serial monitoring of the disease progression, through PSA tests, DRE, and biopsies to track cancer growth. This has become the preferred approach for men with low-grade PCa (1, 168), as men can avoid immediate treatment and prospective side effects (1). When discussing therapy choices and in the selection criteria for active surveillance programs (169), external factors, like obesity, BMI, and the hormonal profile (e.g.: testosterone levels), should be considered by the clinical practice, since all these factors influence the PSA levels (170, 171). Recent studies suggest that the conjugation of PSA screening with other methodologies, such as risk RCs, biomarkers, and imaging techniques like MRI, can attenuate overdiagnosis and underdetection issues (67) (Figure 6). Van Poppel *et al.* (67) proposed a risk-stratified algorithm, combining MRI, RC, and PSA tests, that improves the efficiency of the "PSA-only" screening and reduces unnecessary biopsies and overdiagnosis. The combination of these tools improves the individual balance between the harms and benefits of early detection in well-informed men who are at risk of having PCa (67). It is important to encourage men to have a PSA test when appropriate and after counselling. Based on the initial PSA test result and age, different time intervals for repeated PSA testing are purposed, reflecting the likelihood of a future diagnosis of clinically significant cancer. This strategy helps to avoid false-positive biopsies. For instance, low-risk men can go through individualized PSA tests and, if necessary, repeated MRIs to track cancer growth. Then, RCs seem to be the most appropriate approach to assess the risk of developing PCa after PSA testing. RCs are accessible to every clinician, easy to use, inexpensive, and non-invasive. Moreover, the combined data from RCs and MRI results can be used to determine the need for a prostate biopsy. The European Association of Urology (EAU) guidelines strongly recommends performing a mpMRI before biopsy to modify the management approach accordingly. mpMRI presents preferable detection rates and reduces the number of biopsy procedures due to its capacity to differentiate between significant and insignificant tumours (167). However, the implementation of MRI in the risk assessment of PCa is not yet fully realized in the whole of Europe (67), which in turn reflects the geographical differences in the incidence rates between European countries.

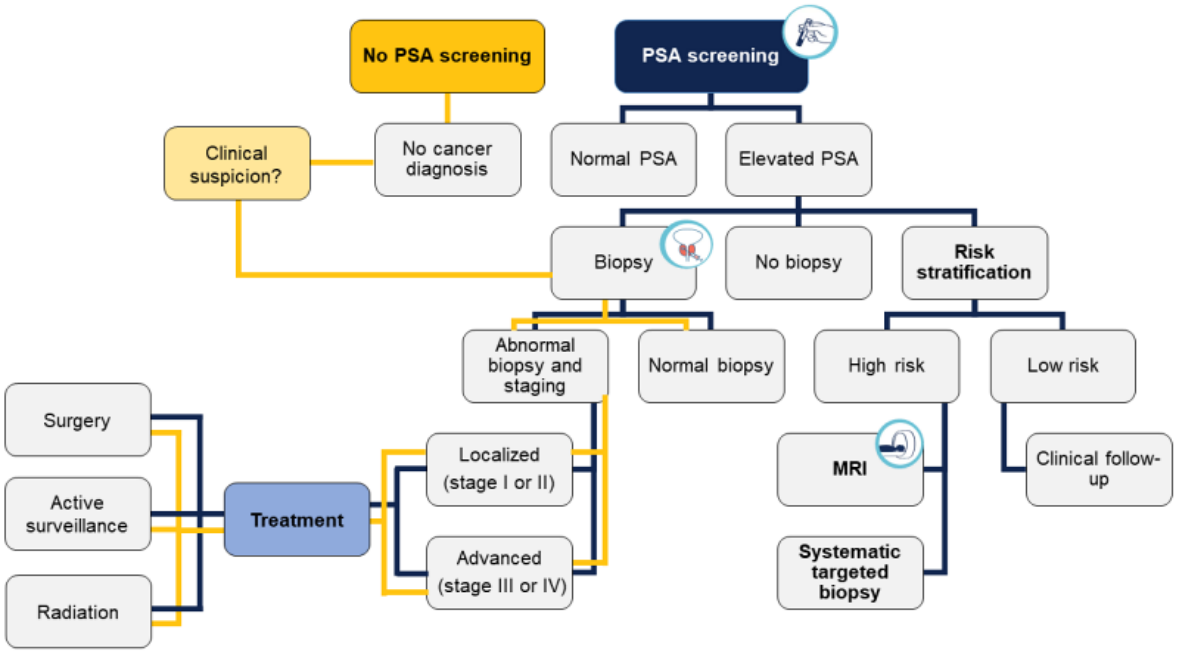


Figure 6 – Application of a risk stratification algorithm in the diagnosis pathway for men without a previous diagnosis of PCa (9,67).

1.5. Contribution of the OMICs science

The OMICs science comprises the dataset of genomics (DNA), transcriptomics (RNA), proteomics (proteins), and metabolomics (metabolites), and is intended to discover cancer-specific biomarkers, useful for its diagnosis and prognosis. In recent years, metabolomics emerged as a promising tool to offer novel insights into disease aetiology and etiological pathways (172). Metabolomics is complementary to genomics, transcriptomics, and proteomics, as it represents the integration of genetic regulation, enzyme activity, and metabolic reactions (173). This science studies the complex interaction of small molecules in biological systems providing comprehensive and detailed information on the phenotype and molecular physiological changes resultant from the interactions between environmental factors, genetic, and both exogenous and endogenous factors, such as age, diet, drugs, chronobiological variations, among others (11). Metabolites represent the end-products of physiological processes and the altered levels of certain metabolites can be measured using metabolomics, to establish a correlation with the disease status (106). Neoplastic cells have a unique metabolic phenotype, related to cancer development and progression, resulting in changes in the production, use, and levels of metabolites (106, 172, 174). Hence, metabolomics has become a powerful tool for the discovery of new cancer biomarkers and therapeutic monitoring, through

the analysis of biomarkers indicative of disease progression and therapeutic response (108, 175).

Volatilomics is a subset of metabolomics based on the study of volatile organic metabolites (VOMs). VOMs are low-molecular-weight metabolites (< 500 Da) with high volatility and a carbon-based chemical group. These metabolites are generated through the metabolism of cells, reflecting their biological activity, and can be released in the blood and excreted through exhaled breath, sweat, urine, or saliva (176). Cancer cells can be distinguished from normal cells, by alterations in normal metabolic rates, apoptotic pathways, and protein expression patterns (98, 176). Metabolic shifts and different responses of the immune system may consist of some of the earliest and detectable changes in cancer, which may become more pronounced as the disease develops. Since VOMs are produced and emitted through the metabolism of cells, cancer development and progression can lead to changes in the volatilomic profile, which can be used to define a volatilomic biosignature for diagnostic purposes (98, 106).

1.5.1. Volatilomics – an emerging yet challenging approach

Empirical data have confirmed the potential of VOMs analysis for cancer screening, characterization of disease progression, and follow-up of the treatment's success, as well as for the discrimination between different types of cancer. The volatilomics approach is based on highly sensitive analytical techniques and does not require invasive procedures, since VOMs can be found in readily accessible biofluids.

Volatilomics studies can go from targeted analysis of one or a small number of metabolites associated with a specific biological pathway to the fingerprinting of a large subset of metabolites associated with a particular phenotype or stimulus (173, 177). Untargeted approaches are more appropriate to detect unexpected changes in the concentrations of specific metabolites (11), meaning that the use of a multi-biomarker panel provides a better evaluation of the cancer progression (98). Moreover, studying VOMs from different chemical families is a much more sensible method, because small genetic or proteic alterations, together with different epigenetic factors, are reflected in differences in the metabolites' levels. VOMs detection requires precise, reliable, and effective instrumentation (178) (Figure 7). Mass spectrometry (MS) is the most used analytical platform for the identification of the volatilomic profile of biological matrices. MS requires an initial separation of metabolites by gas or liquid chromatography (GC or LC, respectively), followed by ionization and resolution according to

the mass-to-charge (m/z) ratio. MS methods have a high sensitivity in more global metabolic profiling applications and can detect secondary metabolites at low concentrations, being more suitable for high throughput methods (98, 173). The most used method for VOMs analysis is the headspace (HS) solid-phase microextraction (SPME) combined with GC-MS, due to its reliable and reproducible results. SPME, developed by Arthur and Pawliszyn in the early 1990s (179, 180), involves the partitioning of analytes from the sample solution into the sorbent coating of the SPME fibre due to the intermolecular interaction with the sorbent material (181). This technique combines sampling, extraction, concentration, and sample preparation in a single step (179, 182). It is a highly efficient technique, with increased sensitivity, automation, and portability, that does not require any concentration step before analysis, preventing the production of interferents (183, 184). The availability of extraction materials and basic equilibration mechanisms makes this methodology very selective, fast, cost-efficient, and with high performance (179, 180). SPME benefits from the constant development of new sorbent coatings (181). For instance, the Divinylbenzene/Carboxen/Polydimethylsiloxane (DVB/CAR/PDMS) fibre, indicated for polar and non-polar volatile metabolites, has been described to enhance the analytical performance when compared to other types of fibres (178).

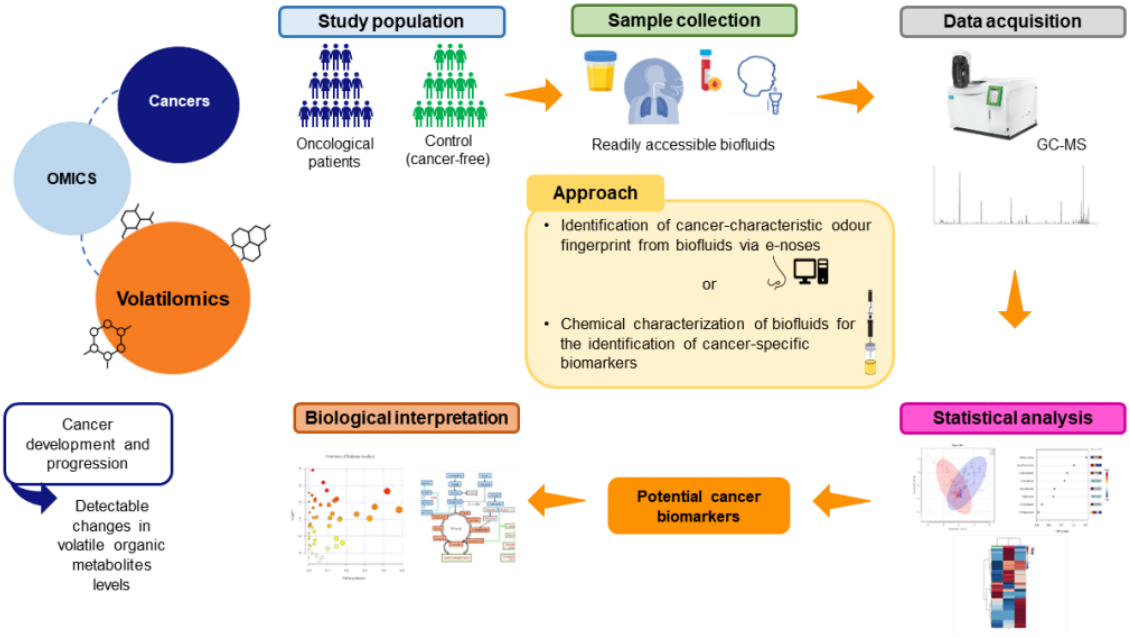


Figure 7 - General flowchart of volatilomics approaches.

Different approaches have been proposed concerning volatilomic studies aiming to find a relationship between VOMs' signature of the body and cancer, based on the comparison of the VOMs pattern in biological samples from cancer patients and healthy individuals. Many

studies have focused on the identification of a cancer-characteristic odour fingerprint from biological fluids through the application of sensorial analyses via electronic noses (e-noses). E-noses consist of non-specific sensors that interact differently with VOMs. Each VOM generates a characteristic fingerprint due to the interaction with the sensor array, which is analysed by an appropriate pattern recognition system to investigate its nature and origin (reviewed by Capelli *et al.* (185)). Urine analysis via an e-nose, for instance, has been shown to distinguish between different types of cancer. This is a non-invasive method based on the findings that dogs can be trained to smell urine and, thus, recognize several types of cancers (186, 187). Other studies have suggested a chemical characterisation of biological fluids for the identification of cancer-specific biomarkers through analytical techniques such as MS-based techniques or SPME (188). These studies are often performed in readily accessible biofluids, mainly exhaled breath, urine, and saliva, since the sampling procedure of these matrices is non-invasive, painless, easy, does not cause any discomfort to the patients, and does not require any specific expertise or specialized staff (175, 189). Blood has also been used in VOMs studies, but obtaining blood samples is an invasive, costly, and time-consuming procedure. Moreover, changes in the temperature or pH of blood samples can change the VOMs profile (176). Table 4 shows a few studies on the volatilomic composition of oncological patients for the identification of discriminative biomarkers for different cancer types in urine, exhaled breath, and saliva samples.

Table 4 - Studies on VOMs for the identification of cancer biomarkers in non-invasive matrices.

Cancer type	Analytical approach	Biomarker's candidates	Prediction model	Validation characteristics	Ref.
		Urine			
Head and neck	HS-SPME/ GC-MS	*m*-cresol, benzene, nonanal, acetone	PLS-DA	NA	(190)
Head and neck	HS-SPME/ GC-MS	2,6-dimethyl-7-octen-2-ol, 1-butanol, *p*-xylene, 4-methyl-2-heptanone	PLS-DA, ROC	NA	(191)
Leukaemia, colorectal, lymphoma	dHS-SPME/ GC-qMS	16 VOMs were found statistically significant	PCA	NA	(192)
Breast	dHS-SPME/ GC-qMS	heptanal, dimethyl disulfide and 2-methyl-3-phenyl-2-propenal	PCA	NA	(183)
Renal cell carcinoma	HS-SPME/ GC-MS	11 VOMs	PCA, PLS-DA	NA	(193)
Pancreatic	TD-GC-TOF-MS GC-IMS	2,6-dimethyl-octane, nonanal, 4-ethyl-1,2-dimethyl-benzene, 2-pentanone	Repeated 10-Fold CV	NA	(194)

Table 4 – *Continuation.*

Cancer type	Analytical approach	Biomarker's candidates	Prediction model	Validation characteristics	Ref.
		Exhaled breath			
Lung	HS-SPME/ GC-MS	acetone, methyl acetate, isoprene, methyl vinyl ketone, cyclohexane, 2-methylheptane, cyclohexanone	DFA, ANN	Sens: 80 % Spec: 91.2 % AUC: NA	(195)
Lung	HS-SPME/ GC-MS	caprolactam and propanoic acid	PCA, OPLS-DA, PLS-DA	NA	(196)
Pancreatic	TD-GC-MS	formaldehyde, acetone, acetoin, undecane, isopropyl alcohol, pentane, n-hexane, 1-butanol, 1-(methylthio)-propane, benzaldehyde, tetradecane, amylene hydrate	ROC	Sens: 81 % Spec: 58 % AUC: 0.736	(197)
Colorectal	TD-GC-MS	15 specific VOMs	PNN	Sens: 86 % Spec: 83 % AUC: 0.852	(198)
Gastric	PTR-TOF-MS	propanal, aceticamide, isoprene, 1,3-propanediol	ROC	Sens: 61 % Spec: 94 % AUC: 0.842	(199)
		Saliva			
Head and neck	HS-SPME/ GC-MS	1,4-dichlorobenzene, 1,2-decanediol, 2,5-bis1,1-dimethylethylphenol, E-3-decen-2-ol	ROC, OPLS-DA	NA	(200)
Colorectal/ stomach	GC-FID	acetaldehyde, acetone, 2-propanol, ethanol, methanol	ROC	Sens: 95.7 % Spec: 90.9 % AUC: 0.857/ 0.839	(201)
Breast	HS-SPME/ GC-MS	3-methyl-pentanoic acid, 4-methyl-pentanoic acid, phenol, p-tert-butyl-phenol (Portuguese samples) and acetic, propanoic, benzoic acids, 1,2-decanediol, 2-decanone, decanal (Indian samples)	PLS-DA, OPLS-DA	NA	(202)
Breast	dHS-SPME/ GC-qMS	phenol, 2-ethyl-1-hexanol	PCA	NA	(203)
Oral	HS-SPME/ GC-MS	1-octen-3-ol, hexanoic acid, E-2-octenal, heptanoic acid, octanoic acid, E-2-nonenal, nonanoic acid, 2,4-decadienal, 9-undecenoic acid	PCA	Sens: 100 % Spec: 100 % AUC: 1	(204)

Legend: ANN: artificial neural networks; AUC: area under the receiver operating characteristic (ROC) curve; CV: cross-validation; DFA: discriminant function analysis; dHS-SPME: dynamic headspace solid-phase microextraction; GC-IMS: gas chromatography-ion migration spectroscopy; GC-MS: gas chromatography-mass spectrometry; GC-qMS: gas chromatography-quadrupole mass spectrometry; HS-SPME: headspace solid-phase microextraction; PCA: principal component analysis; OPLS-DA: orthogonal partial least-squares discriminant analysis; PLS-DA: partial least-squares discriminant analysis; PNN: probabilistic neural network; PTR-TOF-MS: proton-transfer-reaction time-of-flight mass spectrometry; ROC: receiver operating characteristic; TD-GC-MS:

thermal desorption gas chromatography-mass spectrometry; TD-GC-TOF-MS: two-dimensional gas chromatography with time of flight mass spectrometer.

VOMs can originate from endogenous metabolic pathways and/or from external sources, such as diet, drugs, and environmental exposure (exogenous). VOMs constitute a rich source of information on health or disease status, reflecting biochemical and metabolic activities, along with environmental effects, caused by biological activities, including cell death, oxidative stress, or inflammation (176).

VOMs in urine are considered intermediate or end products of metabolic pathways and may include ketones, alcohols, furans, and sulfides (189, 205). The human urinary profile changes over time due to bacterial activity, metabolism, pH variations, decomposition of urine constituents, health status, or physical stress. All these factors are important sources of VOMs produced endogenously. In contrast, dietary habits and environmental exposure to contaminants are two important exogenous sources of many VOMs found in our organism (183). Urine is the preferred biological fluid for a volatilomic approach, due to the enrichment of volatile components, ranging in polarity and complexity (178), caused by their concentration in the kidney before excretion (183). Moreover, its sampling can be performed as often as needed (192), it is easier to obtain in large quantities and to handle, samples can be stored for long periods, it needs less sample preparation, and it contains high amounts of metabolites and low protein content (108).

Most studies on cancer-related VOMs have been performed using exhaled breath, especially for the detection of lung cancer (195, 196) (Table 4). Exhaled breath reflects the volatilomic composition of the bloodstream and airways and, thus, the status and condition of the whole metabolism. Hence, it has the potential to assess the diagnosis, severity, and progression of diseases (206). Exhaled breath contains VOMs from endogenous sources, as well as a large number from exogenous origins. The endogenous metabolites are blood-borne compounds released to the environment through the lungs or compounds made from all classes of symbiotic bacteria. The exogenous VOMs include compounds inhaled from the external environment, including compounds produced after the oral ingestion of food and compounds derived from smoking cigarettes and exposure to pollutants and chemicals (189). VOMs in exhaled breath can be easily detected through biological sensors, such as e-noses (189, 207); however, such an approach often targets a limited number of VOMs previously defined as having a discriminative potential for a given condition. A comprehensive analysis of the volatilomic composition of breath often requires laboratory layouts similar to those used for

other volatilomics analyses. Proton transfer reaction time-of-flight mass spectrometry (PTR-TOF-MS) allows for the real-time identification of the breath volatilomic composition, representing the most potent and sophisticated approach currently available (208). Nevertheless, it involves high acquisition and maintenance costs, as well as specific adaptations to be used in the clinical environment and highly trained personnel to operate it (reviewed by Pereira *et al.* (206)). Given its informative potential to guide clinical decisions, continuous technological improvements will certainly drive the real-time monitoring of exhaled breath in the clinical environment, making it an important tool to add to conventional medical diagnostics (206).

Saliva is the easiest way of sampling biofluids to obtain relevant metabolic information. It is readily available in large amounts all day and contains fewer proteins than blood, thus decreasing any potential risk of non-specific interference and hydrostatic interactions (202, 203). Saliva is considered a mirror of the metabolic interaction with the environment, as it integrates both endogenous and exogenous contributions (209). The volatilomic composition of saliva reflects the oral composition, biochemical, and metabolic blood information, constituting a valuable source of VOMs for cancer biomarkers (210-212). This approach is still relatively new, and it is not as popular in medical diagnosis as urine and serum, but a few studies have demonstrated its potential for head and neck (200) and breast cancer (202, 203) diagnosis.

1.5.1.1. Challenges of volatilomics

Volatilomic studies comprise different approaches, such as the identification of a cancer-characteristic odour fingerprint from biological fluids through e-noses or the chemical characterisation of biological fluids for the identification of cancer-specific biomarkers through MS-based techniques combined with multivariate statistical analysis (188). These approaches are still in progress, and there are a few limitations that prevent them from being used in clinical practice.

E-noses typically use non-specific sensors and can only detect specific molecular patterns based on the differentiation of the odours' fingerprints, which can vary substantially in different biological fluids. Consequently, it is possible to have different sets of VOMs in different biological fluids that are related to the same disease (186, 187). The discrimination between different odours is not operated based on the identification of their chemical composition, since e-noses cannot identify and quantify every VOM found in a sample; thus, other analytical techniques are also used, such as GC-MS. Hence, e-noses do not allow the

identification of specific biomarkers associated with a certain disease (185). Further research is needed to develop and improve specific e-noses for different diseases, along with the need to discover and analyse the connections between specific diseases and body fluids' odours.

MS-based techniques, such as GC-MS, also show some limitations that prevent them from being used in real-time diagnostic applications, such as low sample throughput, high costs, and the requirement for trained personnel and sophisticated software (213). Moreover, these techniques are not portable, and differences in sample preparation and the lack of standardized analytical procedures and statistical treatment of data can compromise the comparison of results among different studies, which is a common problem and debated subject in the field of VOMs assessment (108, 188, 214). In the future, standardization of procedures from collection to data treatment might revolutionize the volatilomics field. For instance, given these concerns, Aggarwal *et al.* (178) proposed an optimized method for urine sample preparation followed by HS-SPME/GC-MS analysis.

The exhaustive comprehension of metabolic pathways and VOMs' origin, in addition to a better evaluation of confounding factors' influence, is another important point in the volatilomics approach. VOMs profile may vary according to the patient's stage of cancer, and potential biomarker candidates must consist of endogenous VOMs, linked to disease-related changes in metabolism. Thus, it is crucial to select and separate endogenous VOMs from exogenous ones before the VOMs analyses. This selection avoids contamination from exogenous and uncontrolled sources, such as diet, medications, environmental factors, smoking, or alcohol consumption, which can lead to abnormal metabolism with subsequent excretion of differing concentrations of VOMs in the biofluids. Additionally, epigenetic factors play a very important role in the clinical phenotypes of cancer, meaning that the volatilomic biosignature and the possible biomarkers will differ between different regions of the world, due to genetic, environmental, and toxicological factors, in addition to the different eating habits practised around the world, that can be related to the development of cancer. Gaining knowledge of the biochemical pathways involved in the VOMs formation is highly desirable to understand the formation of these metabolites and thus, determine their source since some of these metabolites can be originated from both endogenous and exogenous sources (189). Therefore, all these confounding factors make VOMs determination analytically challenging (188, 214). The progress in volatilomics studies will allow an exhaustive comprehension of metabolic pathways and an elucidation of the mechanisms of cancers and how they affect VOMs' production. Moreover, this information will allow researchers to determine the VOMs'

origin in more detail, in addition to leading to a better evaluation of the confounding factors' influence.

Despite these limitations, the tremendous informative potential of volatilomics will allow researchers to gain more in-depth knowledge of cancer development and progression. The volatilomics findings from GC-MS analyses will allow for the discovery of cancer-specific biomarkers and, consequently, the development of highly specific, fast, inexpensive, easy-to-use, and portable sensors (176). Sensors do not require invasive procedures and can be easily implemented in clinical practice without the need for specialized staff (213). Moreover, different types of sensors have been developed to detect cancer-related VOMs, including metal oxide and nanomaterial-based chemiresistive sensors, piezoelectric sensors, colourimetric sensors, and metal-organic frameworks, with very promising diagnostic accuracy in terms of specificity and sensitivity (176). Given these advantages, this approach can be easily disseminated to countries where economic resources and advanced infrastructures are not available (98, 108, 175, 189). Therefore, despite the previously mentioned concerns, the standardization of methods, along with the development of highly specific sensors, will allow for the detection and quantification of specific metabolites for the definition of cancer biomarkers, thus proving the importance of the volatilomics approach (185, 188).

1.5.2. Volatilome of prostate cancer - a promising approach for biomarkers' detection

Prostate cells have a distinct metabolic profile, reflecting the production of citrate, PSA, and polyamines (spermine and myo-inositol), the major components of prostate fluid (173). Studies on the metabolic alterations associated with PCa have demonstrated characteristic decreases in citrate and polyamines and increases in choline glycerophospholipids, lactate, and components of several pathways of amino acid metabolism (173), as well as in the synthesis and oxidation of fatty acids (98, 215). In the advanced stages of PCa, metastases formation is associated with an increase in the glycolytic pathway, also known as the Warburg effect. This phenomenon is characterized by a shift in energy production due to increased aerobic glycolysis and lactate secretion (98, 216). Dysregulations in 14 metabolic pathways mainly related to valine, leucine, and isoleucine biosynthesis; the pentose phosphate pathway; and glycine, serine, and threonine metabolism denoted that PCa development and progression are deeply connected to alterations in amino acids metabolism, energy metabolism, and membrane metabolism (98, 188, 217).

Cancer growth is promoted by the progressive accumulation of genetic, epigenetic, and post-translational changes and by a high metabolic demand, leading to cellular oxidative stress. In turn, these changes increase the liver's production of Cytochrome P (CYP) 450 oxidase enzymes to deal with stress. CYP 450 enzymes catalyze the metabolism of many compounds of both exogenous and endogenous origin, including steroids, vitamins, fatty acids, PG, and leukotrienes. These enzymes may activate exogenous compounds to toxins or carcinogens, and mutations in the *CYP* genes can cause serious health problems. Given their ability to activate or deactivate most carcinogens, CYP 450 enzymes play an important role in cancer formation and are involved in tumour initiation or promotion (176). For instance, mutations in *CYP17A1* lead to mineral corticoid excess syndromes, glucocorticoid, and sex hormone deficiencies, increasing the risk of PCa and benign prostatic hypertrophy (54, 218). CYP 450 enzymes can also mediate the generation of ROS, which are known to be overexpressed in cancer cells (176). Oxidative stress also leads to lipid peroxidation, caused by disease processes or immune responses such as inflammation, and all of these changes can be detected through VOMs. VOMs are end products of carbohydrate and lipid metabolism, as well as oxidative stress and CYP 450 enzymes, and these processes can lead to different volatilomic profiles in cancer patients compared to cancer-free individuals (Figure 8).

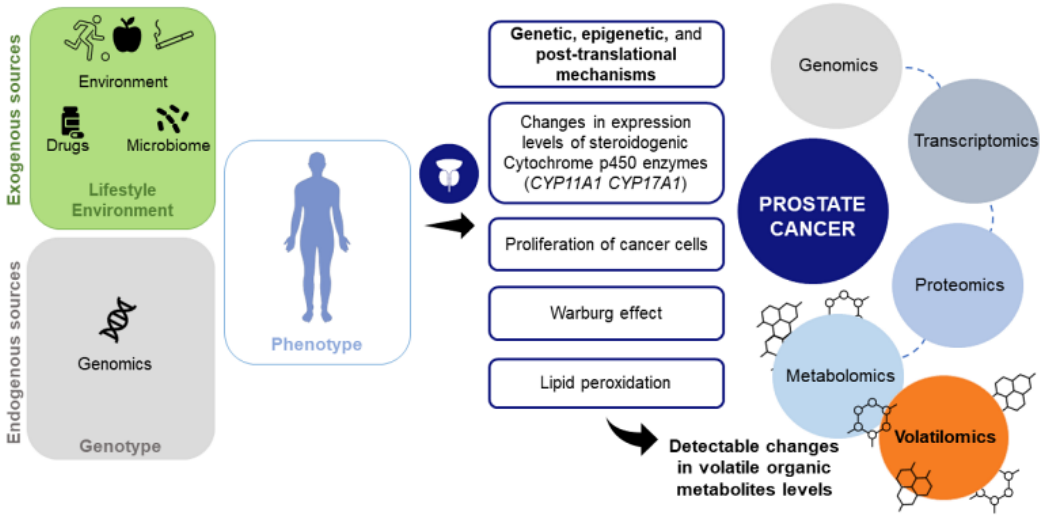


Figure 8 - Cancer-related metabolic and biochemical activities can be detected through VOMs. The establishment of a volatilomic biosignature can lead to the discovery of biomarkers for PCa diagnosis.

The volatilomics approach is still relatively new in PCa when compared to other cancers (Table 4). Table 5 describes volatilomic studies for the definition of PCa biomarkers in urine, exhaled breath, and saliva. Most studies found were performed in urine, as this biofluid contains compounds coming directly from the prostate gland and does not require cross-over blood-

tissue barriers, having fewer confounding elements (11). The identification of specific biomarkers for cancer diagnosis consists of the chemical characterisation of liquid urine or its HS, through MS-based techniques or SPME, aiming at the detection of PCa biomarkers and quantification of their amounts (188). Most research is based on the comparative analysis of samples from PCa patients and controls, such as the studies of Lima *et al.* (106), Khalid *et al*. (219), Stuck-Lewicka *et al.* (220), Gao *et al*. (221), and Jiménez-Pacheco *et al*. (222) (Table 5). These studies developed their methods by combining different analytical techniques, mostly HS-SPME/GC-MS, for the detection and quantification of changes in VOMs levels in PCa samples compared to healthy ones. Interestingly, the studies of Lima *et al.* (213) and Peng *et al.* (223) reported biomarkers for the differential diagnosis of PCa when compared with other cancers. Lima *et al.* (213) developed a urinary 10-biomarker panel for the diagnosis of PCa, with a higher accuracy level than the PSA test. This panel of biomarkers was able to discriminate PCa patients from controls and other urological cancers, including bladder and renal cancers. Peng *et al.* (223) focused on the volatilomic composition of PCa patients' exhaled breath. In this study, the authors used a nanosensor array to discriminate the exhaled breath profile of lung, breast, prostate, and colorectal cancers. The sensor developed was able to differentiate between healthy controls and cancerous patients and between different cancer types. Nevertheless, no exhaustive results have been published until now, since many different VOMs have been proposed as PCa biomarkers, and divergent opinions upon the same metabolites emerged in different studies.

Table 5 - Studies on VOMs for the identification of PCa biomarkers.

Sample Groups	Analytical approach	Biomarker's candidates	Prediction model	Validation Characteristics	Ref.
		Urine			
PCa: 59 HC: 43	HS-SPME/ GC-MS	2,6-dimethyl-7-octen-2-ol, pentanal, 3-octanone, 2-octanone	Repeated 10-Fold CV, Repeated Double CV	NA	(219)
PCa: 58 HC: 60	HS-SPME/ GC-MS	hexanal, 2,5-dimethylbenzaldehyde, 4-methylhexan-3-one, dihydroedulan IA, methylglyoxal, 3-phenylpropionaldehyde	PLS-DA, ROC	Sens: 89 % Spec: 83 % AUC: 0.904	(106)
PCa: 32 HC: 32	GC-MS	VOMs involved in amino acids, purine, glucose, urea, and Krebs cycle biochemical pathways	PCA, PLS-DA	NA	(220)

Table 5 – *Continuation.*

Sample Groups	Analytical approach	Biomarker's candidates	Prediction model	Validation Characteristics	Ref.
PCa: 20 BC: 20 REC: 20 HC: 20	HS-SPME/ GC-MS	methylglyoxal, hexanal, 3-phenylpropionaldehyde, 4-methylhexan-3-one, 2,5-dimethylbenzaldehyde, dihydroedulan IA, ethylbenzene, heptan-2-one, heptan-3-one, 4-(2-methylpropoxy)butan-2-one, methyl benzoate, 3-methyl-benzaldehyde	PLS-DA	Sens: 76 % Spec: 97 % AUC: 0.90	(213)
PCa: 108 HC: 75	SBSE/ TD-GC-MS	11 VOMs	ROC	Sens: 87 % Spec: 77 % AUC: 0.86	(221)
PCa: 29 BPH: 21	HS-SPME/ GC-MS	furan, *p*-xylene	-	NA	(222)
BC: 15 PCa: 55 HC: 36	GC-TOF-MS and GC-IMS	35 VOMs	ROC, Repeated 10-Fold CV	GC-IMS Sens: 87 % Spec: 92 % AUC: 0.95 GC-TOF-MS Sens: 78 % Spec: 88 % AUC: 0.94	(187)
PCa: 88 HC: 86	Urine HS conditioning, followed by e-nose analysis	The study tested the ability of urinary volatilome profiling to distinguish patients with PCa from HC	ROC	Sens: 85 % Spec: 79 % AUC: 0.82	(224)
PCa: 133 HC: 139	Urine HS conditioning, followed by e-nose analysis (Cyranose C320)	The study tested the ability of urinary volatilome profiling to distinguish patients with PCa from HC	PCA, ROC	Sens: 83 % Spec: 88 % AUC: 0.90	(186)
PCa: 132 HC: 60	Urine HS conditioning, followed by e-nose analysis (Cyranose C320)	The study tested the ability of urinary volatilome profiling to distinguish patients with PCa from HC	PCA	Sens: 82 % Spec: 87 % AUC: NA	(225)
HCC: 31 PCa: 62 BC: 29 HC: 18	SPME, followed by analysis with polymer tabs sensor	The study tested the ability of urinary volatilome profiling to distinguish patients with PCa from HC	PCA, ROC	PCa detection Sens: 70 % Spec: NA AUC: 0.70	(226)
		Exhaled breath			
LUC: 30 CC: 26 BTC: 22 PCa: 18 HC: 22	HS-SPME/ GC-MS	6 VOMs for LC, 6 VOMs for CC, 5 VOMs for BTC, 4 VOMs for PCa	PCA	NA	(223)

PCa: 32 HC: 53	E-nose analysis (Cyranose C320)	The study tested the ability of exhaled breath volatilome profiling to distinguish patients with PCa from HC	ANN	Sens: 84 % Spec: 70 % AUC: 0.79	(227)

Legend: ANN: artificial neural network; AUC: area under the ROC curve; BC: bladder cancer; BPH: benign prostate hyperplasia; BTC: breast cancer; CC: colon cancer; CV: cross-validation; GC-IMS: gas chromatography-ion mobility spectrometry; GC-MS: gas chromatography-mass spectrometry; GC-TOF-MS: gas chromatography coupled to time-of-flight mass spectrometry; HC: healthy control; HCC: hepatocellular cancer; HS: headspace; HS-SPME: headspace solid-phase microextraction; LUC: lung cancer; NA: not analysed; PCA: principal component analysis; PLS-DA: partial least-squares discriminant analysis; REC: renal cancer; ROC: receiver operating characteristic; Sens: sensitivity; Spec: specificity; SBSE: stir bar sorptive extraction; TD-GC-MS: thermal desorption gas chromatography-mass spectrometry.

Urine analysis via an e-nose has been shown to distinguish between PCa patients and healthy controls, according to their volatilome profiling. Filianoti *et al*. (186), Taverna *et al.* (224), Capelli *et al.* (225), and Bannaga *et al.* (226) used different methodologies based on urine analysis through e-noses (Table 5) and proved that urine HS and its modification are connected to cancer. In turn, Waltman *et al.* (227) used an e-nose to distinguish between PCa patients and healthy controls, according to the volatilome profiling of exhaled breath. However, to our best knowledge, so far, no study has analysed the VOMs found in the saliva of PCa patients.

Despite the many advantages of using a volatilomics approach to screen PCa, it is challenging to find robust biomarkers given the disparities in results between studies. Potential biomarker candidates must consist of endogenous VOMs linked to disease-related changes in metabolism. Additionally, epigenetic factors play a very important role in the clinical phenotypes of PCa, meaning that the volatilomic biosignature and the possible biomarkers will differ between different regions of the world, as previously explained (189). Hopefully, the progressive increase in studies involving the VOMs composition of PCa patients will help to unveil biomarkers suitable for its diagnosis, as a complementary tool to the current methods (164). Recent studies have shown that the conjugation of PSA screening with other methodologies, such as RCs, biomarkers, and imaging tests, e.g., MRI or fusion biopsies, can attenuate overdiagnosis and under-detection issues (67). This means that the combination of volatilomics with other methodologies could be extremely valuable for the classification and screening of cancer, being beneficial in the active surveillance of patients (11). In the future, the identification of the PCa volatilomic biosignature, through the differentiation of clinical phenotypes in a group of patients, along with the use of specific sensors in clinical practice, will allow for the stratification of individuals into subgroups on which outcomes and treatments are based, thus providing more personalized treatments (11).

1.5.3. Chemometrics tools in volatilomics: supervised and unsupervised methods

Data analysis in volatilomics can be challenging given the high dimensionality and complexity of datasets under analysis. To cope with this problem, the dataset obtained is subject to statistical analysis to find significant variations and trends. These features may allow the discrimination of patients with a specific disease from a control group and eventually the determination of putative cancer biomarkers.

Usually, two types of chemometric tools are used in exploratory studies of volatilomic datasets, univariate and multivariate analysis. Univariate analysis, such as one-way ANOVA, provides a preliminary overview of which variables are potentially significant in discriminating the conditions under study. This approach tests individual variables against a null hypothesis and is also useful to measure correlations between sample groups. The multivariate analysis is divided into unsupervised and supervised methods, in which dataset patterns are obtained based on the relationships between groups (175). Visualizing methods, such as principal component analysis (PCA) and partial least squares-discriminant analysis (PLS-DA), reduce the dimensionality of volatilomics datasets and provide insights into the connections between variables and observations. The main goal is to find the most relevant metabolic alterations that reflect the physiological or pathological state under study (175, 228).

In exploratory studies, the initial application of an unsupervised method like PCA provides an informative first look at the relationships between groups, by showing clusters according to the separation between groups of samples (228). PCA reduces the number of associated factors and produces a scores plot that expresses the data's variation into principal components (PCs), where each sample is expressed as a point (score). PCA allows the visualization of relationships between observations, detects outliers, and predicts sample classification. Each PC comprises the decreasing proportions of the total variance of the variables, where the first components are the ones that explain the majority of the variation contained in the samples (229, 230). However, PCA does not show which variables are directly responsible for the formation of the clusters. Consequently, unsupervised methods do not allow the establishment of a classification model of the different variables (175, 228). Then, the application of predictive and supervised methods such as partial least squares (PLS) regression helps verify and test in more detail the initial conclusions formulated in the PCA, since these methods use the sample information to guide the classification, maximizing class discrimination (228). Furthermore, supervised methods are commonly applied in volatilomic

fingerprinting experiments, as they allow the establishment of important relationships between the variables (VOMs) and the cases (samples) (231). According to the complexity of the sample and the final goal of the analysis, it is possible to use linear methods, such as linear discriminant analysis (LDA). LDA aims to determine a linear function based on the matrix of metabolites (VOMs) obtained and to establish differences between the groups under study. PLS-DA is a type of LDA method commonly used for the analysis of biological samples to understand which metabolites are responsible for the separation between the control group and the patients in the studied groups. The PLS-DA reduces the size of the data matrix through the elimination of redundant variables. In turn, this allows the determination of a set of metabolites that best defines the discrimination between the groups under study (control *vs* patients), according to the existence of an altered metabolite pattern (175). Additionally, PLS-DA provides several features, including loading weight, variable importance in projection (VIP), and regression coefficient that can be used to identify the most important variables (VOMs) that explain the variance between the groups under study. The VIP is a summary of the most relevant variables (VIP score superior or equal to 1) in the PLS-DA model, where the higher the VIP, the more important the metabolite. This test also provides a visual interpretation of complex datasets through a scores plot that illustrates the discrimination between different groups (231). However, as a supervised method, PLS-DA is prone to overfitting. Thus, the outcomes of the PLS-DA should be validated to assess if the model can predict correctly the hypothesized relationships between variables and responses, using a set of independent samples (validation set), cross-validation (CV) or permutation tests (231). CV is used to test the significance and ability of the multivariate model to predict new independent samples without collecting additional information (175, 228). The parameters R^2 (the goodness of fit) and Q^2 (the goodness of prediction) evaluate the quality and reliability of the model generated. Thus, an R^2 value closer to 1, indicates a perfect description of the data by the model, whereas a Q^2 value around 1 indicates perfect predictability of the training model (original data) to correctly classify new samples. Moreover, the difference between R^2 and Q^2 should be minimal (typically $R^2 - Q^2 < 0.03$) to guarantee that the model is not overfitted. The permutation test evaluates the robustness of the PLS-DA model based on prediction accuracy to classify the groups under study compared to any other randomly groups formed by samples permutation (229, 230).

1.6. Objectives of the thesis and experimental approach followed

The primary objective of this thesis was to establish the urinary volatilomic profile of oncological patients and healthy individuals, using a non-invasive approach to set putative biomarkers for PCa diagnosis and management. Urine samples of cancer-free individuals (control group (n = 49)), PCa patients (n = 29), and men subjected to RP (n = 34) were analysed using HS-SPME/GC-MS, a reliable and sensitive methodology. Then, the data matrix obtained was submitted to advanced statistical tools, through multivariant analysis by PCA and PLS-DA, for the selection and definition of potential molecular biomarkers.

Thus, the specific aims to achieve are to:

- Define the populations under study, the population with the pathology characterized from the clinical point of view, and the control population;
- Use the HS-SPME methodology followed by GC-MS analysis, already optimized in previous works by the research group of the University of Madeira, for the study of volatile biomarkers in other types of cancer;
- Establish a methodological, chemical, and biochemical base to contribute to the development of a rapid, inexpensive, and non-invasive diagnostic tool for the detection of PCa.

CHAPTER II.

MATERIALS AND METHODS

2.1. Extraction and analysis conditions

2.1.1. Materials and reagents

Sodium chloride (NaCl, 99.5 %) was acquired from Panreac AppliChem ITW Reagents (Barcelona, Spain) to promote VOMs' salting-out. Hydrochloric acid (HCl, 37 %) and 3-octanol (internal standard (IS), 99 %) from Sigma-Aldrich (St. Louis, MO, USA) were used to prepare the solutions HCl 5 M and 3-octanol 5 parts per million (ppm), respectively, in ultrapure water obtained from a Milli-Q water purification system (Millipore, Bedford, PA, USA). The helium of purity 5.0 (Air Liquide, Algés, Portugal) was used as GC carrier gas. The glass vials, SPME holder for manual sampling, and fibre were purchased from Supelco (Merck KGaA, Darmstadt, Germany). The SPME device included a fused silica fibre coating partially cross-linked with 50/30 µm DVD/CAR/PDMS, which was conditioned at 270 ºC for 30 min before its use, according to the manufacturer's guidelines. Pure standards were used to confirm the VOMs identified through the National Institute of Standards and Technology (NIST) library and were acquired in their maximum purity available, which included octanoic acid (99 %), 2-ethyl-1-hexanol (98 %), and 2-pentanone (98 %) from Acros Organics (NJ, USA), dimethyl disulfide (> 98 %) and 4-heptanone (96 %) from Fluka Analytical (Honeywell Specialty Chemicals Seelze GmbH, Hannover, Germany), and 3-hexanone (98 %) from Aldrich Chemistry (St. Louis, MO, USA).

2.1.2. Subjects and urine sampling

This study involved 49 healthy individuals without any known pathology (control group), 29 PCa patients (PCa group), and 34 subjects who went through an RP (RP group) (Table 6). The control group consisted of non-smoker men with no history of prostate malignancy. These individuals also did not take any medication for age-related comorbidities or metabolic diseases, such as hypertension or diabetes. The urine samples from the control group were collected during General and Family Medicine consultations in *Centro de Saúde do Bom Jesus*. The urine samples from oncological patients were collected in the Urology Unit of

SESARAM, EPERAM before a prostatic biopsy and from patients already under treatment. The research protocol was approved by the local Ethics Committee (CES18/2022) and all participants were fully informed of the objectives of this study and signed an informed consent before being sampled (Appendice 1). Participants were instructed to collect the urine into a 100 mL sterile polyvinyl chloride (PVC) container that was provided to them. Upon collection, each sample was individually homogenized, aliquoted in 8 mL vials, and stored at -80 ºC to avoid any kind of degradation, such as enzymatic activity or oxidation of metabolites, until analysis. All the data collected from the participants was processed to respect confidentiality, privacy, and the ethical principles inherent to any research study involving human subjects.

Table 6 - Demographic and clinical data of the cancer-free controls, PCa patients and RP subjects included in this study.

Characteristics	**Control**	**Prostate cancer**	**Radical prostatectomy**
Number of subjects	49	29	34
Mean Age ± SD (years)	68.4 ± 10.7	71.60 ± 8.46	71.94 ± 7.11
BMI (kg/m^2) ± SD	27.32 ±3.70, *n*=19	28.35 ± 3.76, *n* = 27	27.53 ± 3.16, *n* = 30
Smokers			
Ever smokers	9	13	7
Never smokers	29	15	22
Unknown	11	1	5
PSA (ng/mL), *n* (%)			
< 4	49 (100 %)	4 (13.79 %)	1 (3.03 %)
4-10	-	14 (48.28 %)	17 (51.52 %)
> 10	-	11 (37.93 %)	16 (48.48 %)
Clinical stage, *n* (%)			
I-II	-	25 (86. 21 %)	34 (100 %)
III-IV	-	4 (13.79 %)	-
Treatment*			
Brachytherapy	-	2 (6.90 %)	-
Hormonotherapy	-	3 (10.34 %)	-
Radiotherapy	-	5 (17.24 %)	-
Brachytherapy + hormonotherapy	-	2 (2.60 %)	-
Radiotherapy + hormonotherapy	-	6 (20.69 %)	-

*Eight samples were collected before the confirmatory prostate biopsy.
Legend: BMI: body mass index; SD: standard deviation.

2.1.3. HS-SPME procedure

The HS-SPME extraction was performed according to previously optimized conditions for the analysis of the volatilomic composition of urine samples of other malignant tumours (183, 184). Instead of the CAR/PDMS fibre, a DVB/CAR/PDMS fibre was used to extract a wider range of VOMs. Briefly, 4 mL aliquots of urine sample adjusted to pH 1-2 with 500 µL of HCl (5 M) were transferred to an 8 mL sampling glass vial, and 0.8 g of NaCl and 5 µL of 3-octanol (5 ppm) were added, with stirring at 800 rpm. The vial was placed in a thermostat bath adjusted to 50.0 ± 0.1 ^{o}C and then the SPME fibre was inserted in the HS for 60 min. After sampling, the SPME fibre was withdrawn into the needle, removed from the vial, and inserted in the injector port (250 ^{o}C) of the GC-MS system for 7 min, for desorption of the analytes. Each sample was analysed in triplicate. Blanks, corresponding to the analysis of the coating fibre not submitted to any extraction procedure, were run between sets of three analyses, with a desorption time of 10 min. The absence of 3-octanol in the samples of all studied groups was confirmed before its use as an IS.

2.1.4. GC-MS analysis

The GC-MS analysis was performed with a Perkin Elmer Clarus SQ 8S GC-MS. The gas chromatograph was equipped with a 60 m × 0.25 mm ID × 0.25 µm film thickness, BP-20 (SGE, Dortmund, Germany) fused silica column. After extraction, the SPME coating fibre containing the VOMs was manually introduced into the GC injection port at 250 ^{o}C and kept for 7 min for thermal desorption of VOMs. The oven temperature conditions were: 45 °C for 2 min; 0.7 °C min^{-1} gradient up to 60 °C, held 1 min; 1.0 °C min^{-1} gradient up to 130 °C, held 1 min; and finally, a 15 °C min^{-1} gradient up to 220 °C, held 10 min, giving a total run time of 113.43 min. The column flow was constant at 1.1 mL min^{-1} using helium as carrier gas. The injection port was operated in spitless mode and held at 250 °C. The operating temperatures of the transfer line, quadrupole, and ionization source were 220, 150, and 230 °C, respectively, while electron impact mass spectra were recorded at 70 eV ionization voltage, and the ionization current was 10 µA. Data acquisition was performed in the scan mode (30 – 300 m/z, 0.2 scans per second). The electron multiplier was set to the auto-tune procedure. A transfer line heated to 220 °C carried the compounds from the GC to the MS. VOMs were identified by comparison of the obtained mass spectra with the data system library (NIST, 2014 software, Mass Spectral Search Program v. 2.2; Nist 2014, Gaithersburg, MD), considering a minimum percentage match of 80 %. VOMs identification was also confirmed with the injection of

available pure standards under the same conditions (Figure S.1). The results were expressed as relative peak areas.

2.2. Statistical analysis

The statistical analysis was carried out using MetaboAnalyst 5.0 (232), which comprises the data pre-processing to eliminate VOMs with missing values and data normalization. The data matrix was normalized by cubic root transformation and auto-scaled. The normalized data was processed using the one-way ANOVA followed by Tukey's test for post-hoc multiple comparisons of means (p-values < 0.05). The statistically significant VOMs obtained were then subjected to multivariate statistical analysis, which included PCA to detect trends and possible outliers, followed by PLS-DA. A hierarchical cluster analysis was performed to build a heatmap according to Pearson's correlation to recognize possible clustering patterns between the significantly altered VOMs in the groups under study. The important variables of the PLS-DA model were verified according to the VIP score and were used to validate the PLS-DA models by 10-fold CV and permutation tests (1000 random permutations of Y-observations). The triplicates of each sample were included in the data matrix and are represented in the scores scatter plot as individual variables.

CHAPTER III.

RESULTS AND DISCUSSION

3.1. Characterization of urinary volatile metabolites

VOMs have been described as a promising class of biomarkers for specific diseases, through the definition of volatilomic biosignatures. These sets of VOMs have the potential to be used in early detection, as diagnostic tools, and to monitor disease progression and therapy efficacy (98, 106). This study aimed to establish the urinary volatilomic profile of PCa to set putative biomarkers for PCa diagnosis. The volatilomic composition of urine samples from PCa patients (n = 29), healthy subjects without any known pathology (control group, n = 49), and individuals subjected to RP (n = 34) (Table 6) was established by HS-SPME/GC-MS, according to the experimental procedure described in Chapter II. Following the HS-SPME/GC-MS analysis of the urine samples of the 112 recruited subjects, different chromatographic profiles were obtained for healthy individuals (control group), PCa cancer patients, and RP subjects (Figure 9).

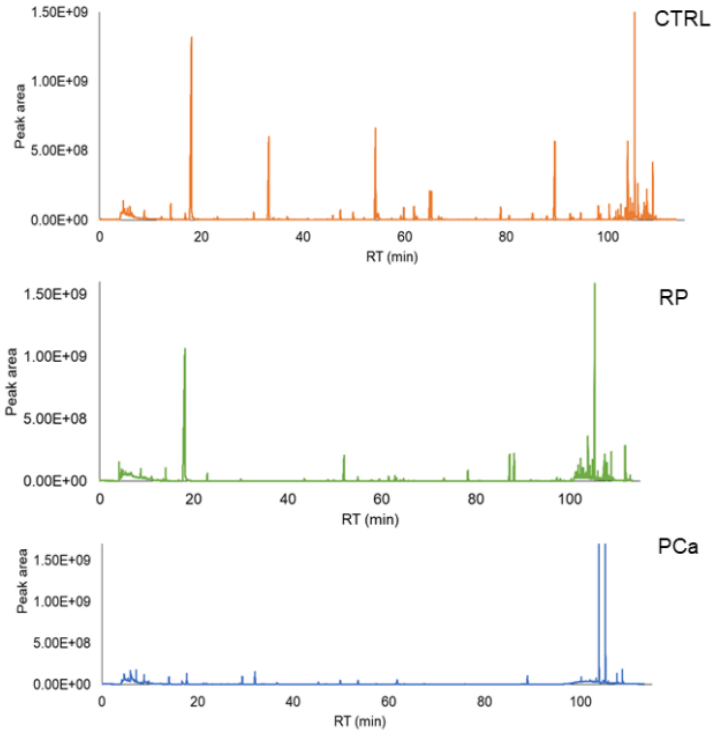


Figure 9 – Representative chromatograms of the groups analysed. Legend: CTRL: control group; RP: radical prostatectomy group; PCa: prostate cancer group.

Overall, 60 VOMs were identified in the samples analysed, belonging to different chemical families, which included 16 terpenes (including terpenoids, mono- and sesquiterpenes), 9 furanic compounds, 7 norisoprenoids, 6 volatile phenols, 5 ketones, 4 alcohols, 4 sulphur-containing compounds, and 9 others (including 1 aldehyde, 1 alkene, 3 benzene derivatives, 1 carboxylic acid, 1 ester, 1 fatty acid, and 1 lactone) (Table 7, frequencies of occurrence available in Table S.1).

Table 7 – Identification, chemical family, possible origin, and mean relative peak area of the identified VOMs in the control group, PCa patients and RP subjects (n = 3; RSD < 20 %).

RT (min)	Metabolites	Formula	CAS nr	Possible Origin	Mean relative peak area		
					CTRL	PCa	RP
Alcohols							
23.65	2-Hexanol[a]	$C_6H_{14}O$	26549-24-6	Endo/Exo (diet)	0.079	0.356	0.221
59.31	2,6-Dimethyl-7-octen-2-ol[a]	$C_{10}H_{20}O$	18479-58-8	Exo (diet)	7.624	0.737	1.637
61.84	2-Ethyl-1-hexanol[a,b]	$C_8H_{18}O$	104-76-7	Endo/Exo (diet)	0.457	1.234	0.938
80.89	2,6-Dimethyl-5,7-octadien-2-ol[a]	$C_{10}H_{18}O$	5986-38-9	Unk	-	-	0.034
Furanic compounds							
5.48	Furan[a]	C_4H_4O	110-00-9	Endo/Exo (diet, env)	0.431	0.317	0.926
6.51	3-Methylfuran[a]	C_5H_6O	930-27-8	Exo (diet, env)	0.114	-	0.128
8.60	2,5-Dimethylfuran[a]	C_6H_8O	625-86-5	Exo (diet, env)	0.005	-	2.593
11.78	2-Acetylfuran[a]	$C_6H_6O_2$	1192-62-7	Exo (diet)	-	0.094	-
11.89	2-Ethyl-5-methylfuran[a]	$C_7H_{10}O$	1703-52-2	End/Exo (diet)	-	0.015	0.106
29.02	2-Pentylfuran[a]	$C_9H_{14}O$	3777-69-3	End/Exo (diet)	-	-	0.011
30.19	2,2-dimethyl-5-(1 -methyl-1-propenyl)-tetrahydrofuran[a]	$C_{10}H_{18}O$	7416-35-5	Unk	0.565	0.389	0.833
85.15	4,7-Dimethylbenzofuran[a]	$C_{10}H_{10}O$	28715-26-6	Endo (syst)	0.066	0.024	-
107.68	2-(1-Cyclopentenyl)furan[a]	$C_9H_{10}O$	115754-78-4	Unk	0.233	-	0.025
Ketones							
9.31	2-Pentanone[a,b]	$C_5H_{10}O$	107-87-9	Exo (diet)	0.229	0.057	0.052
12.78	3-Hexanone[a]	$C_6H_{12}O$	589-38-8	Endo (syst)/ Exo (diet)	<0.001	0.006	-
17.88	4-Heptanone[a,b]	$C_7H_{14}O$	123-19-3	Unk	4.225	1.744	9.071
23.28	Acetone[a]	C_3H_6O	67-64-1	Endo (syst, bact)	0.081	-	0.003
34.59	Cyclohexanone[a]	$C_6H_{10}O$	108-94-1	Endo (syst)/ Exo (diet)	-	0.040	-

Table 7 – *Continuation.*

RT (min)	Metabolites	Formula	CAS nr	Possible Origin	Mean relative peak area		
					CTRL	PCa	RP
Norisoprenoids							
60.49	Theaspirane[a]	$C_{13}H_{22}O$	36431-72-8	Exo (diet)	0.139	0.011	1.072
63.70	3,4,4a,5,6,7-Hexahydro-1,1,4a-trimethyl-2(1H)-naphthalenone[a]	$C_{13}H_{20}O$	4668-61-5	Unk	0.252	0.304	0.559
64.10	γ-Ionone[a]	$C_{13}H_{20}O$	79-76-5	Exo (diet)	0.039	0.154	0.152
89.38	1,1,6-Trimethyl-1,2-dihydronapththalene (TDN) [a]	$C_{13}H_{16}$	30364-38-6	Exo (diet)	4.236	5.577	10.652
98.08	β-Damascenone[a]	$C_{13}H_{18}O$	23726-93-4	Exo (diet)	0.383	0.044	0.488
101.60	α-Ionene[a]	$C_{13}H_{18}$	475-03-6	Exo (diet)	0.010	0.006	0.071
103.86	3,4-Dehydro-β-ionone[a]	$C_{13}H_{18}O$	1203-08-3	Unk	0.522	0.228	2.096
Sulphur-containing compounds							
4.66	Methanethiol[a]	CH_4S	74-93-1	Endo (syst, bact)	1.241	1.562	3.016
14.05	Dimethyl disulfide[a,b]	$C_2H_6S_2$	624-92-0	Endo (bact)/ Exo (diet)	1.344	1.564	1.748
45.81	Dimethyl trisulfide[a]	$C_2H_6S_3$	3658-80-8	Endo (bact)/ Exo	0.276	0.309	0.337
47.28	2-Methyl-5-(methylthio)furan[a]	C_6H_8OS	13678-59-6	Exo (diet)	0.136	0.247	0.258
Terpenes							
12.24	β-Citronellene[a]	$C_{10}H_{18}$	10281-56-8	Exo (diet)	2.291	0.553	1.570
27.12	α-Phellandrene[a]	$C_{10}H_{16}$	99-83-2	Endo/Exo (diet)	-	-	0.001
27.56	γ-Terpinene[a]	$C_{10}H_{16}$	99-85-4	Exo (diet)	0.067	0.092	-
33.06	*o*-Cymene[a]	$C_{10}H_{14}$	13877-91-3	Exo (diet)	4.997	7.689	6.897
35.47	Isoterpinolene[a]	$C_{10}H_{16}$	586-63-0	Endo (syst)/ Exo (diet)	0.039	-	0.009
54.39	*p*-Cymenene[a]	$C_{10}H_{12}$	1195-32-0	Endo/Exo (diet)	4.502	3.484	6.877
58.37	Linalool oxide[a]	$C_{10}H_{18}O_2$	5989-33-3	Exo (diet)	0.172	0.112	0.063
67.65	Pinocarvone[a]	$C_{10}H_{14}O$	30460-92-5	Exo (diet)	-	0.047	-
78.86	Menthol[a]	$C_{10}H_{20}O$	89-78-1	Exo (diet)	4.682	2.722	2.683
86.15	Phellandral[a]	$C_{10}H_{16}O$	21391-98-0	Exo (diet)	0.068	-	-
87.89	Carvone[a]	$C_{10}H_{14}O$	6485-40-1	Exo (diet)	0.039	0.342	1.322
102.05	α-Calacorene[a]	$C_{15}H_{20}$	21391-99-1	Endo/Exo (diet)	0.550	-	0.210
102.78	Dehydro-Ar-α-himachalene[a]	$C_{15}H_{20}$	78204-62-3	Exo (env)	0.004	-	-
105.01	α-Corocalene[a]	$C_{15}H_{20}$	20129-39-9	Endo/Exo (diet)	-	0.322	-
105.65	*p*-Cymen-7-ol[a]	$C_{10}H_{14}O$	536-60-7	Exo (diet)	0.002	-	-
107.94	1,6-Dimethyl-4-isopropylnaphthalene[a]	$C_{15}H_{18}$	483-78-3	Exo (diet)	0.345	0.146	0.294

Table 7 – *Continuation.*

RT (min)	Metabolites	Formula	CAS nr	Possible Origin	Mean relative peak area		
					CTRL	PCa	RP
Volatile phenols							
100.27	*o*-Methoxyphenol[a]	$C_7H_8O_2$	90-05-1	Unk	0.390	0.279	1.012
103.37	2-Bromophenol[a]	C_6H_5BrO	95-56-7	Endo/Exo (diet)	0.188	0.048	0.012
103.94	Phenol[a]	C_6H_6O	108-95-2	Endo/Exo (env)	1.446	2.051	2.678
105.28	*p*-Cresol[a]	C_7H_8O	106-44-5	Endo (bact)/ Exo	6.520	7.860	23.242
106.75	Eugenol[a]	$C_{10}H_{12}O_2$	97-53-0	Exo (diet, env)	-	-	0.298
108.83	*m*-tert-Butylphenol[a]	$C_{10}H_{14}O$	585-34-2	Endo/Exo (diet)	1.516	1.716	2.492
Others							
11.65	2,7-Dimethylocta-1,7-diene	$C_{10}H_{18}$	59840-10-7	Unk	-	-	0.034
12.25	Toluene[a]	C_7H_8	108-88-3	Exo (env)	0.012	-	0.044
41.00	1,2,4-Trimethylbenzene[a]	C_9H_{12}	95-63-6	Endo/Exo (diet)	0.006	0.021	0.044
53.65	Acetic acid[a]	$C_2H_4O_2$	64-19-7	Endo (syst)/ Exo (env)	0.017	0.001	0.006
59.86	1,2,3,5-Tetramethylbenzene[a]	$C_{10}H_{14}$	527-53-7	Exo (env)	0.128	-	0.296
80.57	4-Methyl-4-vinylbutyrolactone[a]	$C_7H_{10}O_2$	1073-11-6	Exo (diet)	-	0.009	-
92.59	Methyl salicylate[a]	$C_8H_8O_3$	119-36-8	Endo/Exo (diet)	0.018	-	0.008
93.13	3-Isopropylbenzaldehyde[a]	$C_{10}H_{12}O$	34246-57-6	Exo (diet)	0.035	0.185	0.017
104.86	Octanoic acid[a,b]	$C_8H_{16}O_2$	124-07-2	Endo (bact)/ Exo	0.086	0.263	0.241

[a]Metabolite identification using mass spectra of the NIST library search.

[b]Metabolite identification using a chemical reference standard, with all analyses performed under identical analytical conditions

Legend: Bact – bacterial; Diet – dietary; Endo – endogenous; Env – environmental; Exo - exogenous; Unk – unknown; Syst – systemic. CTRL – control group; PCa – prostate cancer; RP – radical prostatectomy; RT – retention time; "–" - not identified.

The detailed analysis of each sample group shows differences in terms of areas for the different chemical families (Figure 10, Table 7). The human urinary profile changes over time because of bacterial activity, metabolism, pH variations, or decomposition of urine constituents, and it is affected by different external factors, such as health status, dietary habits, physical stress, and environmental exposure, which contribute with exogenous compounds to the volatilomic profile of individuals. Due to these factors, the human metabolism is very complex and cancer development and progression make it even more difficult to understand all the metabolic processes that may contribute to an increase or decrease of certain metabolites (183, 192, 233). Thus, it is crucial to establish a relationship between the VOMs identified and their

potential endogenous origin. Currently, the origin of many VOMs has not been clearly defined (176). In many cases, the possible origin is still unknown and in other cases, the same metabolite can be obtained from both endogenous and exogenous sources, being very challenging to completely understand its metabolic pathway (192, 234, 235). Hence, gaining knowledge of the metabolic pathways that lead to the production or elimination of VOMs will provide a better understanding of the biochemical changes that occur in cancer (176).

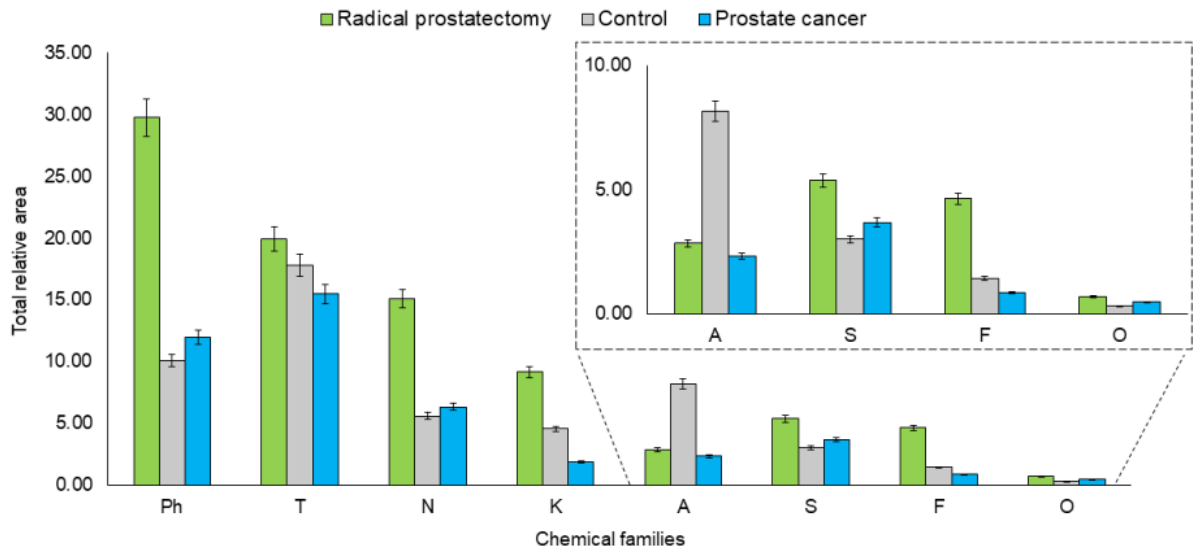


Figure 10 – Distribution of the chemical families identified in the studied groups, RP (*n* = 34), control (*n* = 49), and PCa (*n* = 29). Legend: Ph: phenolic compounds; T: terpenes; N: norisoprenoids; K: ketones; A: alcohols; S: sulphur-contain compounds; F: furanic compounds; O: others.

Terpenes, volatile phenols, and norisoprenoids, were the chemical families that contributed the most to the volatilomic profile of the three studied groups (Figure 10). Norisoprenoids, phenolic, and terpenic compounds can be easily found in different exogenous sources like food (236, 237). Nevertheless, many metabolites belonging to these chemical families originate from endogenous metabolomic processes in our organism, namely *p*-cymenene, α-calacorene, 2-bromophenol, phenol, and *p*-cresol (238). Terpenes come from the mevalonic acid pathway (183, 192) and can also result from the consumption of foods and beverages (238). *p*-Cresol, *o*-cymene, 1,1,6-trimethyl-1,2-dihydronapththalene (TDN), and *p*-cymenene were the most abundant metabolites in the PCa group (Table 7). *p*-Cresol is produced by human intestinal microflora during the aromatic amino acid metabolism and its levels in

urine strongly correlate with the levels of proteins in the diet (233, 238). *o*-Cymene, *p*-cymenene, and menthol are dietary products from vegetables, cereals and fruit (233, 238). *o*-Cymene is typically found in citrus fruits and has been proposed as a putative biomarker of citrus ingestion (233, 238). It was detected in the urine samples from the individuals belonging to the three studied groups (Table S. 1).

According to the literature, ketones are one of the most abundant chemical families in the volatilomic profile of urine (192, 239). However, the contribution of ketones to the urinary volatilomic profile of the recruited individuals was lower than expected (Figure 10). They are products of different metabolic pathways, namely carbohydrate metabolism and lipid oxidation processes (240, 241). A few studies have proposed that a considerable fraction of ketones in urine arises from the action of gut bacteria, but ketones can also come from exogenous sources, like food (beverages, foods, and flavouring ingredients) or environmental pollution (176). Acetone, the simplest ketone, results from either the metabolism of glucose or from the oxidation of fatty acids, through the decarboxylation of acetoacetate, which is formed by both lipolysis and the breakdown of ketogenic amino acids. Acetone might also come from diet or environmental contamination (176). Cyclohexanone has been found as a volatile component of human urine and it is obtained through oxidation of cyclohexane or dehydrogenation of phenol. This metabolite was only detected in urine samples of PCa patients (Table S.1). 3-Hexanone has been detected in different foods and is associated with several diseases such as non-alcoholic fatty liver disease, autism, and inborn metabolic disorder celiac disease (176). The levels of 3-hexanone were higher in PCa patients than in the control group (Table 7). 4-Heptanone was found in almost every sample (Table S.1) and was one of the most abundant metabolites in the RP group, along with *p*-cresol, TDN, and *o*-cymene (Table 7).

Sulphur-containing compounds are another chemical family identified with high expression in the volatilomic profile of the PCa group (Figure 10). Similarly to ketones, this chemical family has been described to possess a high expression in the human urinary volatilomic profile (183, 242). Most of these metabolites are formed by the incomplete metabolism of methionine and cysteine, through the transamination pathway (183, 243-245). During transamination, methionine and cysteine are transformed into methanethiol (219), which is easily oxidized to dimethyl sulfide, dimethyl disulfide, and dimethyl trisulfide (244), for example. It has been described that gram-negative bacteria may also produce considerable amounts of methanethiol and dimethyl disulfide (246). Furthermore, these compounds can also

result from dietary sources since dimethyl disulfide and dimethyl trisulfide are present in many foods and beverages.

Alcohols, along with terpenes and volatile phenols, contributed the most to the volatilomic profile of the control group, in which 2,6-dimethyl-7-octen-2-ol, *p*-cresol, *o*-cymene, and menthol were the most abundant metabolites (Figure 10, Table 7). Alcohols can have different origins, including the reduction of the respective fatty acid in the gastrointestinal tract, the pyruvate, the citrate, or glycolysis pathways (247), and even the metabolism of hydrocarbons (176). Similarly, the metabolism of microorganisms, such as bacteria, can also be a source of these metabolites (248). Another source of alcohols is the diet, through the ingestion of food and beverages (176). 2,6-Dimethyloct-7-en-2-ol was previously detected in urine samples of PCa patients (219) and reported in lower levels than control subjects (176). The same trend was observed in this work (Table 7).

Hydrocarbons are metabolites of great diagnostic interest since they are closely related to oxidative stress (206). Alkanes and other methylated hydrocarbons typically result from the lipid peroxidation of polyunsaturated fatty acids found mainly in cell membranes (206). Significant changes in the levels of alkanes and methyl alkanes in cancer patients might be related to the activity of CYP 450 enzymes (176). On the other hand, unsaturated hydrocarbons, typically alkenes, are often related to the mevalonic acid pathway of cholesterol synthesis (206). In this work, only one alkene was detected, 2,7-Dimethylocta-1,7-diene, which was found in 9 % of RP subjects (Table S.1). PAHs are carcinogenic substances to which humans are exposed from the environment, at certain industrial workplaces, and from tobacco smoke (206). Naphthalene is one of the PAHs often associated with cancer development and it is released by industrial, domestic, and natural burning processes, leading to general population exposure (206, 249). However, so far, no metabolic pathway can clearly explain the origin of naphthalene derivatives in urine (183). Some researchers indicate a potential relationship with steroid metabolism, while others suggest that these compounds may come from the environment to which the individual is exposed (206, 249).

Furanic compounds and benzene derivatives can be found in both exogenous and endogenous sources, as metabolic products of food and different processes in the human organism (236-238). The primary source of furan and furanic compounds is the thermal degradation and rearrangement of carbohydrates, such as glucose, lactose, and fructose, in natural food components and processed commercial foods (233, 237, 238). Furan can also be produced endogenously, but it results most probably from the ingestion of food (238). Furan was found in samples belonging to the three studied groups (Table S.1) and has been proposed

as a PCa biomarker by Jiménez-Pacheco *et al.* (222) (Table 5). 2-Pentylfuran was only found in the RP subjects and this metabolite has been associated with bacterial infections (176). Benzene derivates are often related to environmental sources, such as air and environmental pollution from industrial (pesticides, dyes) or natural processes (fires). The major sources of benzene exposure are automobile service stations and tobacco smoke (239).

3.2. Chemometric analysis of urine samples

The statistical analysis was performed using MetaboAnalyst 5.0 (232). The variables were initially normalized (as described in Chapter II) to obtain a homogeneous distribution and generate reliable and interpretable models. The normalized matrix was subjected to univariate analysis through one-way ANOVA with a post-hoc Tukey test ($p < 0.05$), in which the p values obtained proved that 22 of the 60 VOMs identified presented statistically significant differences between the three groups analysed, the healthy subjects - control group, oncological patients - PCa group, and RP individuals - RP group (Table 8). Some of these 22 metabolites have been previously related to oncological pathologies in different studies, according to the Human Metabolome Database (176). TDN and *p*-cresol have been detected in urine samples of colorectal, leukaemia, and lymphoma cancers, where they were found increasingly expressed in the samples of oncological patients (176). 2-Ethyl-1-hexanol has been detected in five types of cancer, namely lung, laryngeal, thyroid, colorectal and breast cancers (176). This metabolite was found in 62 and 59 % of PCa and RP samples, respectively (Table S.1), and in both cases, its expression was higher than in the control group (Table 7). 1,2,4-Trimethylbenzene has been detected in urine samples of four types of cancer, namely colorectal, lymphoma, leukaemia, and lung cancers, where it was increasingly expressed in oncological patients (176). 4-Heptatone is considered a cancer biomarker, but its origin is still unknown. Previous studies have shown that 4-heptanone is most probably produced from an exogenous origin, as a product of *in vivo* β-oxidation of 2-ethylhexanoic acid from plasticizers (183, 250). Although the origin of some of these metabolites is known (Table 7), most of them still need a more detailed evaluation to establish a relationship with PCa.

Table 8 - Important features identified by one-way ANOVA with post-hoc Turkey test ($p < 0.05$) for the three groups under study, control (0), PCa (1), and RP (2).

Metabolite	*f* value	*p* value	-log10(p)	FDR	Post-hoc tests
2-Bromophenol	18.31	2.85E-08	7.55	1.53E-06	1-0; 2-0
o-Methoxyphenol	17.67	5.10E-08	7.29	1.53E-06	2-0; 2-1
3,4-Dehydro-β-ionone	16.55	1.41E-07	6.85	2.17E-06	2-0; 2-1
2-Ethyl-5-methylfuran	16.52	1.45E-07	6.84	2.17E-06	2-0; 2-1
p-Cresol	12.53	5.65E-06	5.25	6.78E-05	2-0; 2-1
Theaspirane	8.38	2.81E-04	3.55	2.68E-03	2-0; 2-1
TDN	8.11	3.63E-04	3.44	2.68E-03	2-0
Methanethiol	7.99	4.08E-04	3.39	2.68E-03	2-0; 2-1
α-Calacorene	7.84	4.72E-04	3.33	2.68E-03	1-0; 2-0
2-Ethyl-1-hexanol	7.75	5.11E-04	3.29	2.68E-03	1-0; 2-0
2-(1-Cyclopentenyl)furan	7.71	5.32E-04	3.27	2.68E-03	1-0; 2-0
β-Damascenone	7.71	5.35E-04	3.27	2.68E-03	1-0; 2-1
2-Pentanone	5.57	4.19E-03	2.38	1.85E-02	1-0; 2-0
2-Methyl-5-(methylthio)furan	5.49	4.50E-03	2.35	1.85E-02	2-0
Carvone	5.46	4.62E-03	2.34	1.85E-02	1-0
4-Heptanone	4.93	7.79E-03	2.11	2.92E-02	1-0; 2-1
p-Cymenene	4.43	1.27E-02	1.90	4.32E-02	2-1
2-Acetylfuran	4.39	1.31E-02	1.88	4.32E-02	1-0; 2-1
Cyclohexanone	4.35	1.37E-02	1.86	4.32E-02	1-0; 2-1
α-Corocalene	4.28	1.47E-02	1.83	4.40E-02	1-0; 2-1
1,2, 4-Trimethylbenzene	4.15	1.65E-02	1.78	4.54E-02	2-0
3-Methylfuran	4.15	1.67E-02	1.78	4.54E-02	1-0

Legend: FDR - false discovery rate.

The PCA and PLS-DA multivariate pattern recognition procedures use the information contained in the VOMs fingerprint as several variables to visualize group trends and clustering patterns, respectively, according to the separations among sample sets. Both PCA (Figure 11A) and PLS-DA (Figure 11B) were not able to provide complete discrimination of the three groups under study. The variances of the PLS-DA component 1 and component 2 were 24.7 and 11.1 %, respectively, representing 35.8 % of the total variability of the data (Figure 11B). The heatmap created using Pearson's correlation with the 22 statistically significant VOMs (Table 8) illustrated the correlations between these VOMs and the sample groups (Figure 11C). This hierarchical cluster analysis showed that each cluster of the studied groups is well defined by a

distinct panel of metabolites. For instance, 2-(1-cyclopentyl)furan, 2-pentanone, and 2-bromophenol were the metabolites more associated with the control group, while carvone, α-corocalene, 2-acetylfuran, and cyclohexanone showed high correlations with the PCa group. In turn, 2-ethyl-5-methylfuran, methanethiol, *o*-methoxyphenol, *p*-cymenene, and 3,4-dehydro-β-ionene, were the more associated VOMs with the RP group. The panel of VOMs with the highest correlations in the RP group was mostly composed of terpenic compounds, which might be related to dietary factors.

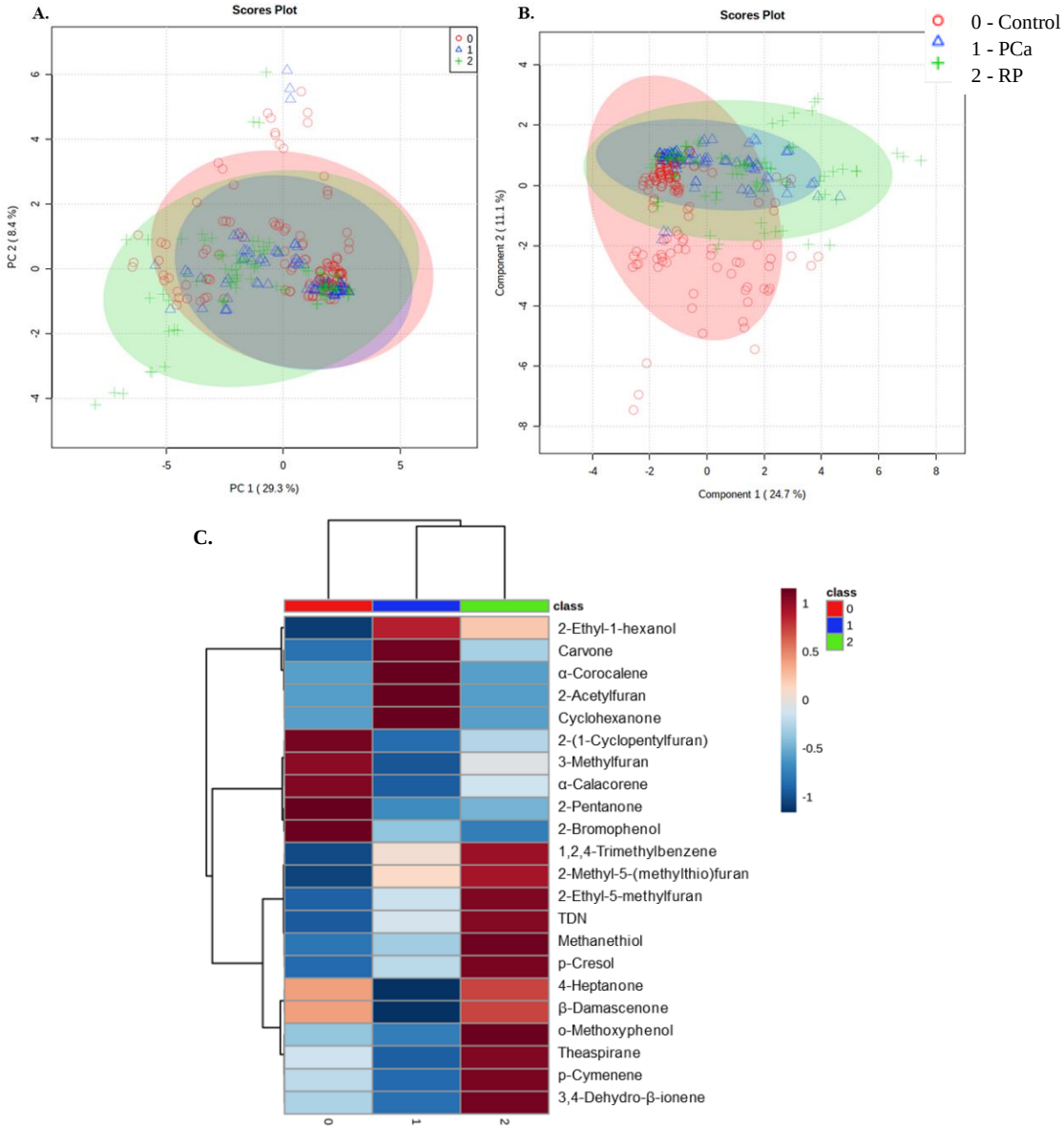


Figure 11 – Score scatter plots of the volatilomic fingerprint of urine samples from the three groups under study, control (0) *vs* PCa (1) *vs* RP (2). **A.** PCA PC1 × PC2 score scatter plot. **B.** PLS-DA score scatter plot. **C**. Heatmap illustrating the mean levels (normalized peak areas) of the 22 VOMs statistically significant by Pearson's distance analysis. Rows correspond to the mean normalized peak area of each metabolite with the sample groups in the columns.

The VIP scores computed through the PLS-DA-based algorithm were used to select the metabolites that best discriminated the groups' control *vs* PCa *vs* RP (Figures 12A and 12B), but the PLS-DA model obtained only partially discriminated the groups under study. Concerning the PLS-DA model validation, the R^2 and Q^2 values were lower than 50 % (Figure 12C). Any R^2 value lower than 1.0 indicates that at least some variability in the data cannot be explained by the model (251). In this case, an R^2 of 0.237 (Table 9) means that 76 % of the variability in the outcome data cannot be explained by the PLS-DA model obtained. Moreover, the model was not overfitted, as proved by the difference between R^2 and Q^2 (R^2 - Q^2 < 0.03). A random permutation test with 1000 permutations was used to evaluate whether the class assignment was good or bad. Despite having returned a *p*-value < 0.001 and further guaranteeing that there was no overfitting in the model obtained, it failed to show predictive ability to classify the samples, as seen in the distribution of the permuted samples (Figure 12D).

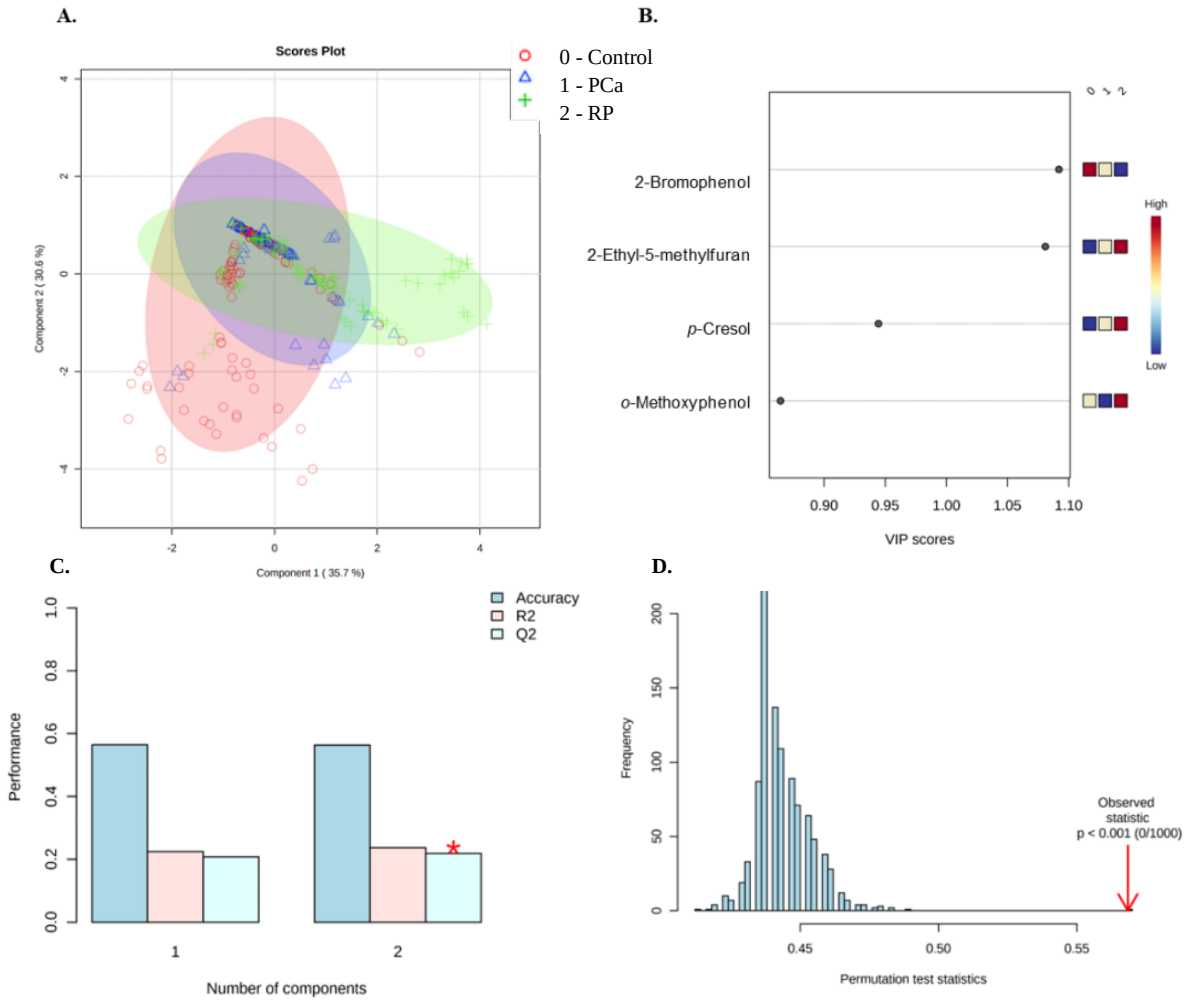


Figure 12 – **A:** PLS-DA score scatter plots of the volatilomic fingerprint of urine samples from the three groups under study, control (0) *vs* PCa (1) *vs* RP (2). **B.** VIP scores **C:** 10-fold CV performance of control *vs* PCa *vs* RP (*means best Q^2 value). **D:** Model validation by permutation test based on 1000 permutations of VOMs obtained by GC-MS of urine samples from the groups under study.

Table 9 – PLS-DA CV details of the comparison of control *vs* PCa *vs* RP.

Measure	Control *vs* PCa *vs* RP (2 components)
Accuracy	0.563
R^2	0.237
Q^2	0.218

The 22 statistically significant VOMs (Table 8) were then used to compare the urinary volatilomic profiles of the control *vs* PCa, control *vs* RP, and PCa *vs* RP (Figure S.2). The metabolites with the highest contribution to group discrimination in the three comparisons (control *vs* PCa, control *vs* RP, and PCa *vs* RP) were selected according to the VIP results and considered for further analysis (Figure 13). The PLS-DA model obtained for the volatilomic fingerprint of control *vs* PCa showed a total variability of data of 54.4 % (Figure 13A) but did not result in complete discrimination between the two groups. This can be due to the fact that most samples analysed corresponded to patients that were already under treatment when the urines were collected, as described in Table 6. To clarify this hypothesis the volatilomic profile of these cases should be monitored more closely during the treatment. Regarding the control group, the urinary profiles revealed some interindividual variability (Figure 13A). External and uncontrolled variables might be related to this observation, including genetic and epigenetic factors, such as diet, medication, and lifestyle. However, it is challenging to assess which VOMs are directly related to these factors since some metabolites show both endogenous and exogenous origins and the metabolic origin of most of them is not fully understood yet, as well as their role in PCa. The PLS-DA of the control *vs* RP (Figure 13B) and PCa *vs* RP (Figure 13C) both achieved good discrimination with a total variability of data of 74.3 % and 77.8 %, respectively. The individuals who were subjected to surgery (RP group) had their prostate removed in different years, with the average happening in 2017, and all had been diagnosed with PCa in stages I-II (Table 6). Nevertheless, it was not possible to establish a correlation between the time after the prostatectomy treatment and the closeness to the control group. The same was valid for the comparison between these patients and the PCa group. Other characteristics related to the morphology of the tumour and the clinical characteristics of each individual could help explain the results obtained in the PLS-DA. It is also worth noticing that PCa is very heterogeneous in terms of grade and genetics (1), meaning that the comparison of different cases could be a challenging task and would require more subjects to be recruited to mitigate such interference.

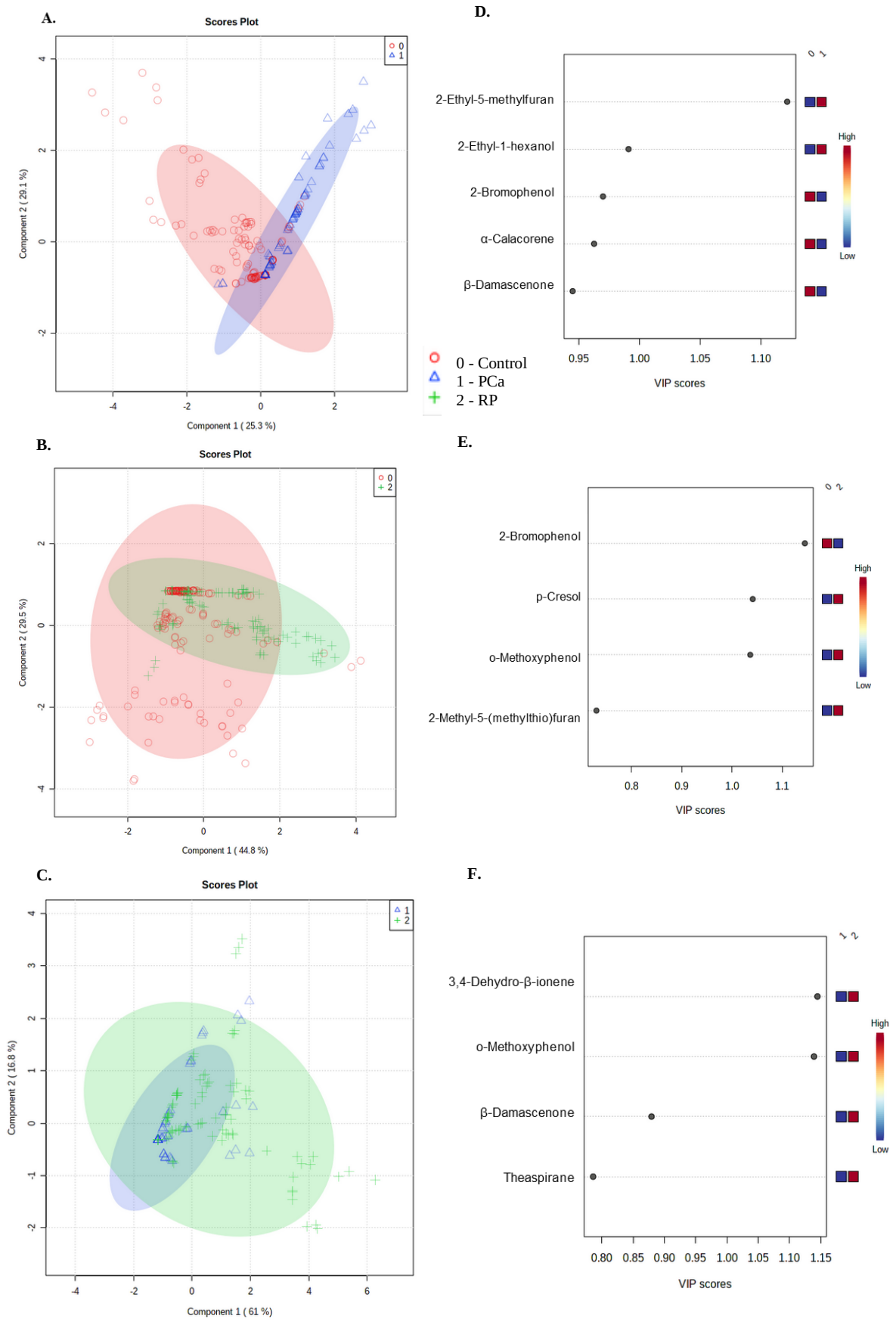


Figure 13 – **A-C**: PLS-DA score scatter plots of the volatilomic fingerprint of urine samples from **A.** control (0) *vs* PCa (1), **B.** control (0) *vs* RP (2), and **C.** PCa (1) *vs* RP (2). **D-F:** Corresponding VIP scores.

The quality and reliability of the PLS-DA models generated were evaluated (Table 9). Despite the models not being overfitted, as proved by the difference between R^2 and Q^2 lower than 0.03 in all comparisons (Table S.2), the goodness of fit and predictability values were lower than 50 % (Figure 14A-C). The R^2 value measures the percentage of the variation in the dependent variable that is explained by variation in the independent variable. Consequently, the variance in the study population can strongly influence the R^2 magnitude, which in turn is deeply influenced by the variation in the independent variable (251). The histograms obtained in the permutation tests (Figure 14D-F) show the distribution of the permuted samples, which failed to show predictive ability to classify the samples.

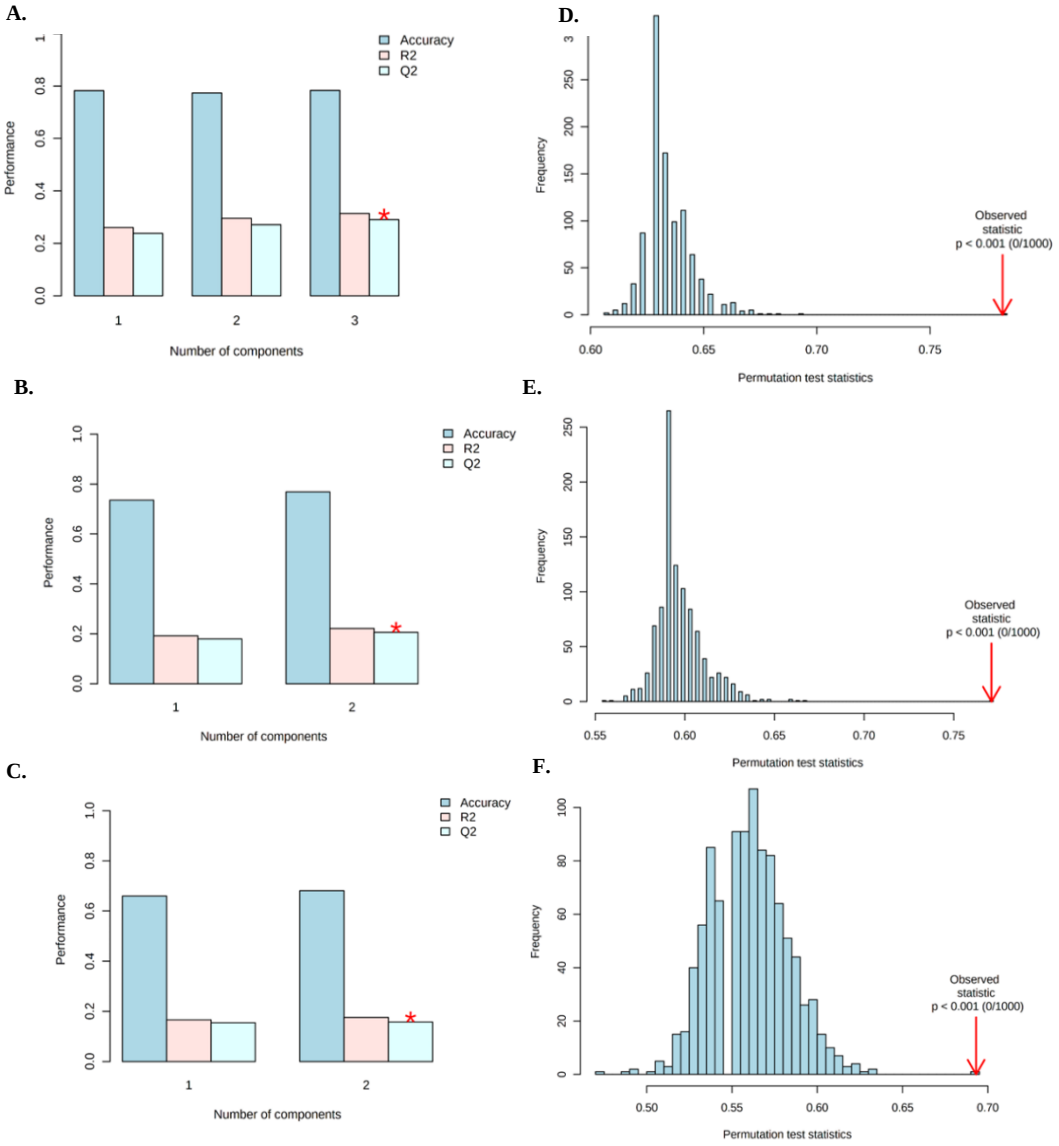


Figure 14 – **A-C:** 10-fold CV performance of **A.** control *vs* PCa, **B.** control *vs* RP, and **C.** PCa *vs* RP (*means best Q^2 value). **D-F:** Model validation by permutation test based on 1000 permutations of VOMs obtained by GC-MS of urine samples from the groups under study.

Considering the results obtained for the comparisons of the three groups under study (Figures 13 and 14) and the differences among the PCa group for the type of treatment, PLS-DA were applied using the kind of therapy as a discriminant factor. A total of 29 PCa patients enrolled in this study, 8 of which donated their urine samples before the confirmatory prostate biopsy has been done. Thus, these patients were sampled before starting PCa treatment, while the remaining 21 corresponded to individuals who were already under different treatments (Table 6). The treatment options for patients with low-grade PCa include RP, radiotherapy, hormonotherapy (generally androgen deprivation therapy), and chemotherapy (252, 253). Hormonotherapy is used to reduce the androgens' levels in the body or to stop them from getting into PCa cells and stimulating their growth (252, 253). Radiotherapy comprises external beam radiotherapy and brachytherapy and is mostly applied to patients who are in a reasonably good physiological state that can tolerate possible serious comorbidities (253). This treatment is based on the use of ionizing radiation to kill tumour cells and is mostly applied to PCa patients with lesions confined to the pelvis. Brachytherapy comprises the implantation of radioactive particles into the prostate gland. This form of radiation therapy causes less damage to the healthy tissue nearby and includes temporary seed implantation (high-dose rate) or permanent seed implantation (low-dose rate) (252). Radiotherapy can also be used for the corrective treatment of local recurrence after RP (252). This treatment is mostly used in combination with hormonotherapy, as the combination of these therapies seems to improve overall survival (253). For instance, of the 21 PCa patients under treatment, only 8 were submitted to combined therapies, namely brachytherapy combined with hormonotherapy (2 subjects) and radiotherapy combined with hormonotherapy (6 subjects) (Table 6). Hence, to understand if the type of treatment could discriminate the PCa samples, these were classified in combined, one or no therapy (Table 6). The 22 statistically significant VOMs in Table 8 were used for further analysis. Figure 15 shows the PCA and PLS-DA score scatter plots obtained for the volatilomic fingerprint of the studied groups. The results showed that the PCa individuals who received different types of treatment can be grouped accordingly. However, the total variability of the data obtained in the PLS-DA was 35.5 % in the comparison of all groups (Figure 15B).

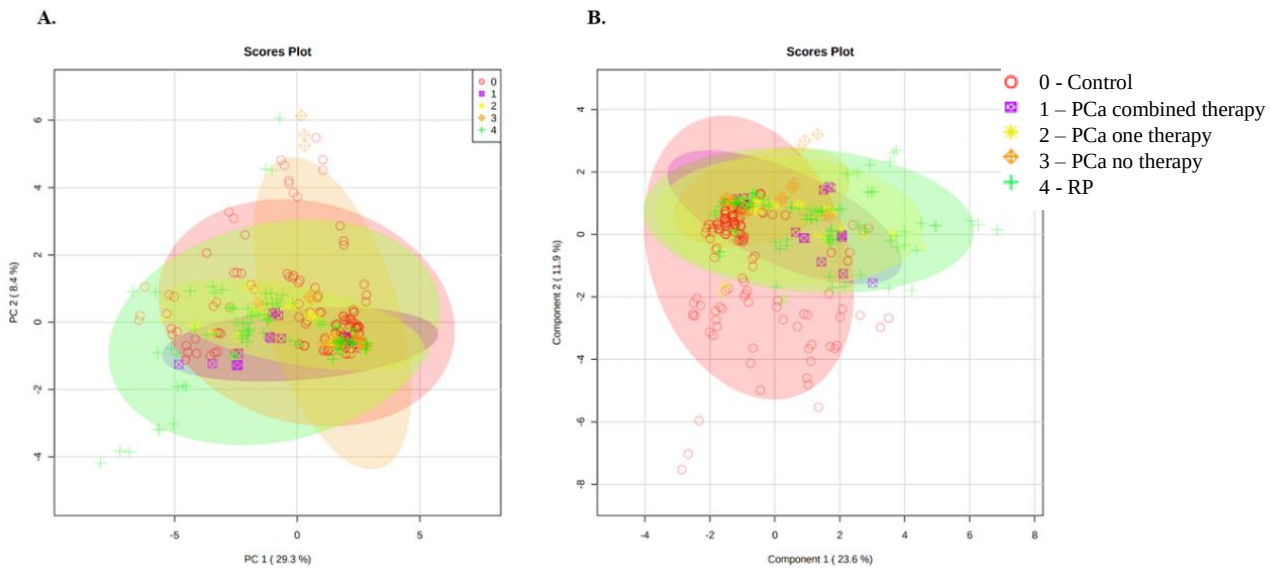


Figure 15 - Score scatter plots of the volatilomic profile of urine samples from the control (0) *vs* PCa with combined therapy (1) *vs* PCa with one therapy (2) *vs* PCa with no therapy (3) *vs* RP (4). **A.** PCA PC1 × PC2 score scatter plot. **B.** PLS-DA score scatter plot.

The PCa samples classified according to the kind of treatment and the control group were compared (Figure S.3A). The top metabolites with the highest contribution to group discrimination were selected according to the VIP results of the PLS-DA, in which the total variability of the data was 44.3 % (Figure 16A). The discrimination between the groups showed that the oncological patients under treatment could be more related to the control group, but genetic and epigenetic factors such as age, health status, and duration of therapy influence the treatment's outcome. Moreover, the samples collected before a prostate biopsy are still under analysis through imaging tests, to characterize and classify the type and stage of the tumour in more detail. The comparison of the PCa groups showed good discrimination with a total variability of the data of 80.9 % (Figure 16B). The RP subjects were also compared to the PCa patients classified according to their treatment (Figure S.3C). The top metabolites with the highest contribution to group discrimination were also selected according to the VIP results of the PLS-DA, in which the total variability of the data was 33.8 % (Figure 16C). Nevertheless, the results obtained in these comparisons failed to show the predictive ability to classify the samples according to the different types of treatment that the patients recruited were subjected to. Additionally, it would be important to know how long ago these patients initiated their therapy. Another key feature would be the PSA value when the urine sample was collected since the PSA indicated in Table 6 is relative to the date of diagnosis. The clinical stage was also not known in detail, since most subjects that agreed to participate in this study began their

treatment in private health facilities and it was not possible to collect this crucial data so far. The data about other external factors that could contribute to the high variance obtained, such as BMI and smoking habits, were also not available for all the recruited subjects. Obtaining all clinical and demographic characteristics would be crucial to understanding how they may have contributed to the results obtained, as well as testing them as discriminant factors. VOMs may vary according to the patient's stage of cancer but also differ between individuals due to uncontrolled variables, including genetic and epigenetic factors, such as the environment, diet, lifestyle, BMI, medication, hydration status, and smoking habits. All these confounding factors strongly impact qualitatively and quantitatively the individual urine volatilomic composition, even when strong disease associations are found. Furthermore, the metabolic origin of the majority of the VOMs identified and their relationship with PCa are not completely elucidated. Many of the metabolites already described in the literature have a mixed origin, both endogenous and exogenous, which may limit their potential use as biomarkers. Ideally, a potential biomarker candidate should be an endogenous VOM, linked to disease-related changes in the metabolism. In this thesis, it was not possible to accomplish a comprehensive evaluation of the origin and relationships of the metabolites identified, which will be addressed in future work. Overall, to improve the predictive power and reliability of the statistical models developed concerning the identification of putative volatile urinary biomarkers for PCa, it would be necessary to recruit a much larger population and fully characterized in terms of confounding factors.

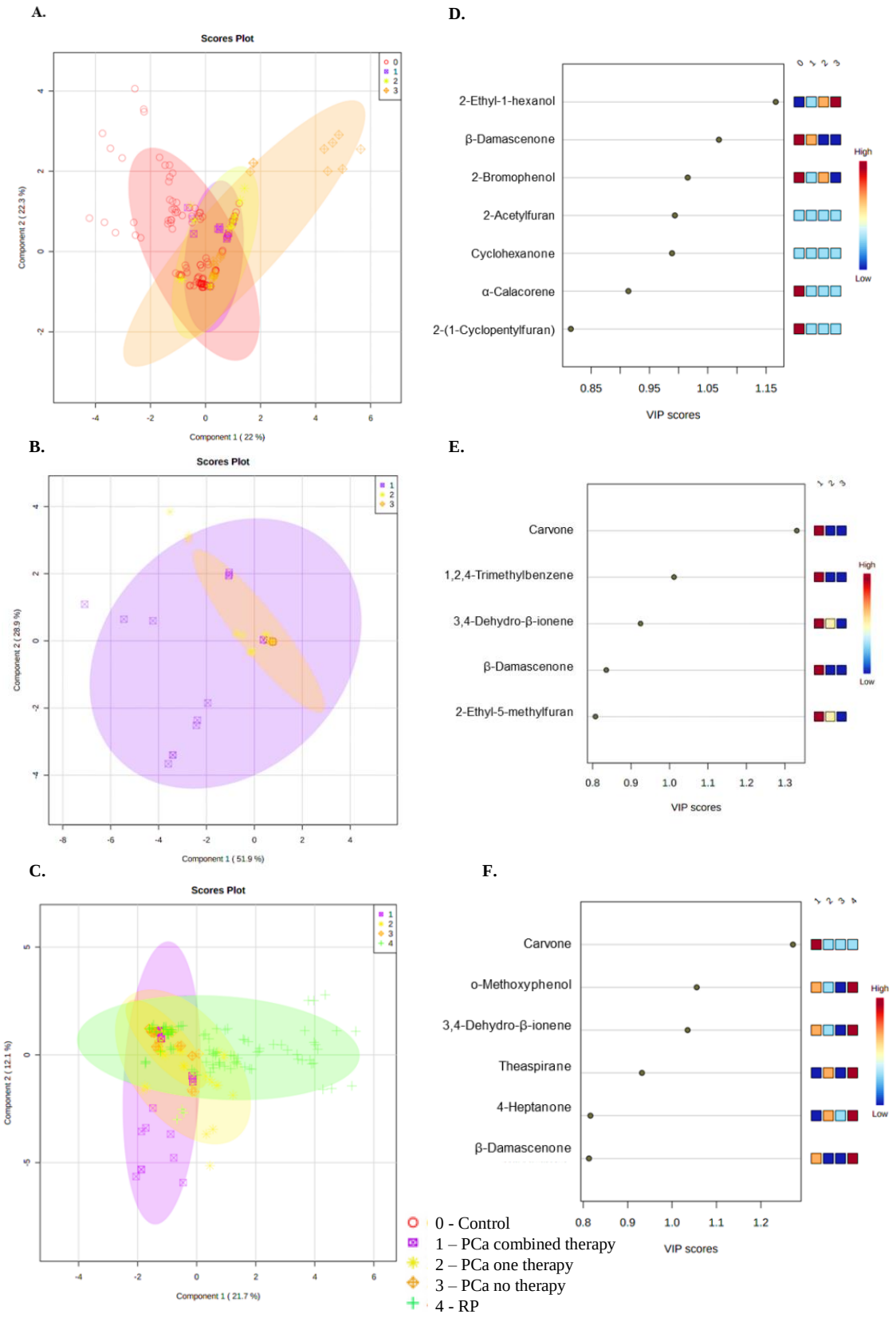


Figure 16 – A-C: PLS-DA score scatter plots of the volatilomic fingerprint of urine samples from **A.** control (0) *vs* PCa with combined therapy (1) *vs* PCa with one therapy (2) *vs* PCa with no therapy (3) **B.** PCa with combined therapy (1) *vs* PCa with one therapy (2) *vs* PCa with no therapy (3), and **C.** PCa with combined therapy (1) *vs* PCa with one therapy (2) *vs* PCa with no therapy (3) *vs* RP (4). **D-F:** Corresponding VIP scores.

CHAPTER IV.

CONCLUSIONS AND FUTURE PERSPECTIVES

Considering the main goal of this thesis, a total of 60 VOMs were identified belonging to different chemical families and retrieving different chromatographic profiles for the three groups of subjects recruited. Terpenes, volatile phenols, and norisoprenoids were the chemical families that contributed the most to the volatilomic profile of the three studied groups, control, PCa and RP. The statistical analysis revealed that 22 of the 60 VOMs identified presented statistically significant differences between the recruited groups, according to the one-way ANOVA with post-hoc Turkey test ($p < 0.05$). The multivariate analysis, through PLS-DA, however, did not show clear discrimination between these groups (control *vs* PCa *vs* RP). The robustness of the PLS-DA model generated was evaluated by 10-fold CV and permutation tests, which indicated that more than 50 % of the variability in the outcome data could not be explained by the model. The hierarchical cluster analysis by Pearson's correlation showed that each cluster of the studied groups is well defined by a distinct panel of metabolites. The metabolites 2-(1-cyclopentylfuran), 2-pentanone and 2-bromophenol were more associated with the control group, while carvone, α-corocalene, 2-acetylfuran, and cyclohexanone showed high correlations with the PCa group. In turn, 2-ethyl-5-methylfuran, methanethiol, *o*-methoxyphenol, *p*-cymenene, and 3,4-dehydro-β-ionene were the more associated VOMs with the RP group. The comparison of the groups' control *vs* PCa, control *vs* RP, and PCa *vs* RP also did not result in clear discrimination in the models obtained. These results can be related to the influence of both endogenous and exogenous factors on the volatilomic profile of individuals. However, it was not possible to do an exhaustive analysis of the study population to understand how their demographic and clinical characteristics could help explain these results by testing confounding factors as discriminant factors. Further validation of the findings in this thesis is required using a much larger sample cohort to improve the predictive power and reliability of the statistical models developed. Likewise, additional research is required to determine which of the metabolites are of tumour origin and which originate from normal metabolic processes and external contaminations (environment or diet), as well as which are down (or up) regulated by tumour growth. These open questions will be assessed in future work.

REFERENCES

1. Rawla P. Epidemiology of prostate cancer. World J Oncol. 2019;10(2):63-89.
2. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin. 2021;71(3):209-49.
3. Beltran H, Demichelis F. Intrapatient heterogeneity in prostate cancer. Nat Rev Urol. 2015;12(8):430-1.
4. Dudka I, Thysell E, Lundquist K, Antti H, Iglesias-Gato D, Flores-Morales A, et al. Comprehensive metabolomics analysis of prostate cancer tissue in relation to tumor aggressiveness and TMPRSS2-ERG fusion status. BMC Cancer. 2020;20(1):437.
5. Ostrander EA, Johannesson B. Prostate cancer susceptibility loci: finding the genes. Adv Exp Med Biol. 2008;617:179-90.
6. Kgatle MM, Kalla AA, Islam MM, Sathekge M, Moorad R. Prostate Cancer: Epigenetic alterations, risk factors, and therapy. Prostate Cancer. 2016;2016:1-11.
7. Culp MBB, Soerjomataram I, Efstathiou JA, Bray F, Jemal A. Recent global patterns in prostate cancer incidence and mortality rates. Eur Urol. 2020;77(1):38-52.
8. Tonon L, Fromont G, Boyault S, Thomas E, Ferrari A, Sertier AS, et al. Mutational profile of aggressive, localised prostate cancer from African Caribbean men Versus European ancestry men. Eur Urol. 2019;75(1):11-5.
9. Tikkinen KAO, Dahm P, Lytvyn L, Heen AF, Vernooij RWM, Siemieniuk RAC, et al. Prostate cancer screening with prostate-specific antigen (PSA) test: A clinical practice guideline. BMJ. 2018;362:k3581.
10. Braga R, Costa AR, Pina F, Moura-Ferreira P, Lunet N. Prostate cancer screening in Portugal: Prevalence and perception of potential benefits and adverse effects. Eur J Cancer Prev. 2020;29(3):248-51.
11. Salciccia S, Capriotti AL, Lagana A, Fais S, Logozzi M, De Berardinis E, et al. Biomarkers in prostate cancer diagnosis: From current knowledge to the role of metabolomics and exosomes. Int J Mol Sci. 2021;22(9):4367.
12. Lee S, Ku JY, Kang BJ, Kim KH, Ha HK, Kim S. A unique urinary metabolic feature for the determination of bladder cancer, prostate cancer, and renal cell carcinoma. Metabolites. 2021;11(9):591.

13. Spur EM, Decelle EA, Cheng LL. Metabolomic imaging of prostate cancer with magnetic resonance spectroscopy and mass spectrometry. Eur J Nucl Med Mol Imaging. 2013;40 Suppl 1(0 1):S60-71.

14. Center MM, Jemal A, Lortet-Tieulent J, Ward E, Ferlay J, Brawley O, et al. International variation in prostate cancer incidence and mortality rates. Eur Urol. 2012;61(6):1079-92.

15. Fowke JH, McLerran DF, Gupta PC, He J, Shu X-O, Ramadas K, et al. Associations of body mass index, smoking, and alcohol consumption with prostate cancer mortality in the Asia cohort consortium. Am J Epidemiol. 2015;182(5):381-9.

16. Markozannes G, Tzoulaki I, Karli D, Evangelou E, Ntzani E, Gunter MJ, et al. Diet, body size, physical activity and risk of prostate cancer: An umbrella review of the evidence. Eur J Cancer. 2016;69:61-9.

17. Marhold M, Kramer G, Krainer M, Le Magnen C. The prostate cancer landscape in Europe: Current challenges, future opportunities. Cancer Lett. 2022;526:304-10.

18. Gandaglia G, Leni R, Bray F, Fleshner N, Freedland SJ, Kibel A, et al. Epidemiology and prevention of prostate cancer. Eur Urol Oncol. 2021;4(6):877-92.

19. Miranda A, Mayer-da-Silva A, Glória L, Brito C. Registo oncológico nacional de todos os tumores na população residente em Portugal, em 2018. Lisbon; 2021.

20. Ferlay J, Colombet M, Soerjomataram I, Mathers C, Parkin DM, Piñeros M, et al. Estimating the global cancer incidence and mortality in 2018: GLOBOCAN sources and methods. Int J Cancer. 2019;144(8):1941-53.

21. Baade PD, Youlden DR, Krnjacki LJ. International epidemiology of prostate cancer: Geographical distribution and secular trends. Mol Nutr Food Res. 2009;53(2):171-84.

22. Camacho C, Nunes C, Fraga C, Santos F, Camacho J, Barradas N, et al. Registo oncológico na Região Autónoma da Madeira, 2018. Funchal; 2020.

23. Instituto Nacional de Estatística, IP - Portugal. 2019 Jan 8 [cited 2022 May 12]. Available from: http://www.ine.pt/.

24. Rebbeck TR. Prostate cancer disparities by race and ethnicity: from nucleotide to neighborhood. Cold Spring Harb Perspect Med. 2018;8(9):a030387.

25. Taitt HE. Global Trends and Prostate Cancer: A review of incidence, detection, and mortality as influenced by race, ethnicity, and geographic location. Am J Mens Health. 2018;12(6):1807-23.

26. McAllister BJ. The association between ethnic background and prostate cancer. Br J Nurs. 2019;28(18):S4-s10.

27. Brown CR, Hambleton I, Hercules SM, Unwin N, Murphy MM, Nigel Harris E, et al. Social determinants of prostate cancer in the Caribbean: a systematic review and meta-analysis. BMC Public Health. 2018;18(1):900.

28. Grossman DC, Curry SJ, Owens DK, Bibbins-Domingo K, Caughey AB, Davidson KW, et al. Screening for prostate cancer US Preventive services task force recommendation statement. JAMA. 2018;319(18):1901-13.

29. Vietri MT, D'Elia G, Caliendo G, Resse M, Casamassimi A, Passariello L, et al. Hereditary prostate cancer: Genes related, target therapy and prevention. Int J Mol Sci. 2021;22(7):3753.

30. Brandão A, Paulo P, Teixeira MR. Hereditary predisposition to prostate cancer: from genetics to clinical implications. Int J Mol Sci. 2020;21(14):5036.

31. Bhanji Y, Isaacs WB, Xu J, Cooney KA. Prostate Cancer Predisposition. Urol Clin North Am. 2021;48(3):283-96.

32. Bree KK, Hensley PJ, Pettaway CA. Germline predisposition to prostate cancer in diverse populations. Urol Clin North Am. 2021;48(3):411-23.

33. Coughlin SS, Vernon M, Klaassen Z, Tingen MS, Cortes JE. Knowledge of prostate cancer among African American men: A systematic review. Prostate. 2021;81(3):202-13.

34. Ventimiglia E, Salonia A, Briganti A, Montorsi F. Re: Family history and probability of prostate cancer, differentiated by risk category — A nationwide population-based study. Eur Urol. 2017;71(1):143-4.

35. Vidal AC, Freedland SJ. Obesity and Prostate Cancer: a focused update on active surveillance, race, and molecular subtyping. Eur Urol. 2017;72(1):78-83.

36. Wilson RL, Taaffe DR, Newton RU, Hart NH, Lyons-Wall P, Galvão DA. Obesity and prostate cancer: A narrative review. Crit Rev Oncol Hematol. 2022;169:103543.

37. Adesunloye BA. Mechanistic insights into the link between obesity and prostate cancer. Int J Mol Sci. 2021;22(8):3935.

38. Bandini M, Gandaglia G, Briganti A. Obesity and prostate cancer. Curr Opin Urol. 2017;27(5):415-21.

39. Fujita K, Hayashi T, Matsushita M, Uemura M, Nonomura N. Obesity, Inflammation, and prostate cancer. J Clin Med. 2019;8(2):201.

40. Wilson KM, Mucci LA. Diet and lifestyle in prostate cancer. Adv Exp Med Biol. 2019;1210:1-27.

41. Pernar CH, Ebot EM, Wilson KM, Mucci LA. The epidemiology of prostate cancer. Cold Spring Harb Perspect Med. 2018;8(12):a030361.

42. Kaiser A, Haskins C, Siddiqui MM, Hussain A, D'Adamo C. The evolving role of diet in prostate cancer risk and progression. Curr Opin Oncol. 2019;31(3):222-9.

43. Darcey E, Boyle T. Tobacco smoking and survival after a prostate cancer diagnosis: A systematic review and meta-analysis. Cancer Treat Rev. 2018;70:30-40.

44. Sato N, Shiota M, Shiga KI, Kashiwagi E, Takeuchi A, Inokuchi J, et al. Effect of smoking on oncological outcome among prostate cancer patients after radical prostatectomy with neoadjuvant hormonal therapy. Cancer Invest. 2020;38(10):559-64.

45. Khan S, Thakkar S, Drake B. Smoking history, intensity, and duration and risk of prostate cancer recurrence among men with prostate cancer who received definitive treatment. Ann Epidemiol. 2019;38:4-10.

46. Jochems SHJ, Fritz J, Häggström C, Järvholm B, Stattin P, Stocks T. Smoking and risk of prostate cancer and prostate cancer death: a pooled study. Eur Urol. 2022;S0302-2838(22)01804-8.

47. Foerster B, Pozo C, Abufaraj M, Mari A, Kimura S, D'Andrea D, et al. Association of smoking status with recurrence, metastasis, and mortality among patients with localized prostate cancer undergoing prostatectomy or radiotherapy: a systematic review and meta-analysis. JAMA Oncol. 2018;4(7):953-61.

48. Fraser GE, Jacobsen BK, Knutsen SF, Mashchak A, Lloren JI. Tomato consumption and intake of lycopene as predictors of the incidence of prostate cancer: the Adventist Health Study-2. Cancer Causes Control. 2020;31(4):341-51.

49. Li N, Wu X, Zhuang W, Xia L, Chen Y, Wu C, et al. Tomato and lycopene and multiple health outcomes: Umbrella review. Food Chem. 2021;343:128396.

50. Soares N, Elias MB, Lima Machado C, Trindade BB, Borojevic R, Teodoro AJ. Comparative analysis of lycopene content from different tomato-based food products on the cellular activity of prostate cancer cell lines. Foods. 2019;8(6):201.

51. Puah BP, Jalil J, Attiq A, Kamisah Y. New insights into molecular mechanism behind anti-cancer activities of lycopene. Molecules. 2021;26(13):3888.

52. Oczkowski M, Dziendzikowska K, Pasternak-Winiarska A, Włodarek D, Gromadzka-Ostrowska J. Dietary factors and prostate cancer development, progression, and reduction. nutrients. 2021;13(2):496.

53. Beynon RA, Richmond RC, Santos Ferreira DL, Ness AR, May M, Smith GD, et al. Investigating the effects of lycopene and green tea on the metabolome of men at risk of prostate cancer: The ProDiet randomised controlled trial. Int J Cancer. 2019;144(8):1918-28.

54. Maksymchuk OV, Kashuba VI. Altered expression of cytochrome P450 enzymes involved in metabolism of androgens and vitamin D in the prostate as a risk factor for prostate cancer. Pharmacol Rep. 2020;72(5):1161-72.

55. Capiod T, Barry Delongchamps N, Pigat N, Souberbielle JC, Goffin V. Do dietary calcium and vitamin D matter in men with prostate cancer? Nat Rev Urol. 2018;15(7):453-61.

56. Grant WB. Review of Recent Advances in understanding the role of vitamin D in reducing cancer risk: breast, colorectal, prostate, and overall cancer. Anticancer Res. 2020;40(1):491-9.

57. Ardura JA, Álvarez-Carrión L, Gutiérrez-Rojas I, Alonso V. Role of calcium signaling in prostate cancer progression: Effects on cancer hallmarks and bone metastatic mechanisms. Cancers (Basel). 2020;12(5):1071.

58. Applegate CC, Rowles JL, Ranard KM, Jeon S, Erdman JW. Soy consumption and the risk of prostate cancer: an updated systematic review and meta-analysis. Nutrients. 2018;10(1):40.

59. Tsugane S. Why has Japan become the world's most long-lived country: insights from a food and nutrition perspective. Eur J Clin Nutr. 2021;75(6):921-8.

60. Rogovskii VS, Popov SV, Sturov NV, Shimanovskii NL. The possibility of preventive and therapeutic use of green tea catechins in prostate cancer. Anticancer Agents Med Chem. 2019;19(10):1223-31.

61. Musial C, Kuban-Jankowska A, Gorska-Ponikowska M. Beneficial properties of green tea catechins. Int J Mol Sci. 2020;21(5):1744.

62. Miyata Y, Shida Y, Hakariya T, Sakai H. Anti-cancer effects of green tea polyphenols against prostate cancer. Molecules. 2019;24(1):193.

63. Ferreira PMP, Rodrigues L, de Alencar Carnib LP, de Lima Sousa PV, Nolasco Lugo LM, Nunes NMF, et al. Cruciferous vegetables as antioxidative, chemopreventive and antineoplasic functional foods: Preclinical and clinical evidences of sulforaphane against prostate cancers. Curr Pharm Des. 2018;24(40):4779-93.

64. Ahmad F, Cherukuri MK, Choyke PL. Metabolic reprogramming in prostate cancer. Bri J Cancer. 2021;125(9):1185-96.

65. Kheirandish P, Chinegwundoh F. Ethnic differences in prostate cancer. Br J Cancer. 2011;105(4):481-5.

66. Haiman CA, Chen GK, Blot WJ, Strom SS, Berndt SI, Kittles RA, et al. Characterizing genetic risk at known prostate cancer susceptibility loci in African Americans. PLoS Genetics. 2011;7(5):e1001387.

67. Van Poppel H, Hogenhout R, Albers P, van den Bergh RCN, Barentsz JO, Roobol MJ. A European model for an organised risk-stratified early detection programme for prostate cancer. Eur Urol Oncol. 2021;4(5):731-9.

68. Rebbeck TR, Devesa SS, Chang B-L, Bunker CH, Cheng I, Cooney K, et al. Global patterns of prostate cancer incidence, aggressiveness, and mortality in men of African descent. Prostate Cancer. 2013;2013:1-12.

69. Panigrahi GK, Praharaj PP, Kittaka H, Mridha AR, Black OM, Singh R, et al. Exosome proteomic analyses identify inflammatory phenotype and novel biomarkers in African American prostate cancer patients. Cancer Med. 2019;8:1110-23.

70. Bratt O, Drevin L, Akre O, Garmo H, Stattin P. Family history and probability of prostate cancer, differentiated by risk category: a nationwide population-based study. J Natl Cancer Inst. 2016;108(10):djw110.

71. Gathirua-Mwangi WG, Zhang J. Dietary factors and risk for advanced prostate cancer. Eur J Cancer Prev. 2014;23(2):96-109.

72. Liu X, Allen JD, Arnold JT, Blackman MR. Lycopene inhibits IGF-I signal transduction and growth in normal prostate epithelial cells by decreasing DHT-modulated IGF-I production in co-cultured reactive stromal cells. Carcinogenesis. 2008;29(4):816-23.

73. Peters U, Leitzmann MF, Chatterjee N, Wang Y, Albanes D, Gelmann EP, et al. Serum lycopene, other carotenoids, and prostate cancer risk: a nested case-control study in the prostate, lung, colorectal, and ovarian cancer screening trial. Cancer Epidemiol Biomarkers Prev. 2007;16(5):962-8.

74. Aronson WJ, Barnard RJ, Freedland SJ, Henning S, Elashoff D, Jardack PM, et al. Growth inhibitory effect of low fat diet on prostate cancer cells: results of a prospective, randomized dietary intervention trial in men with prostate cancer. J Urol. 2010;183(1):345-50.

75. Rohrmann S, Platz EA, Kavanaugh CJ, Thuita L, Hoffman SC, Helzlsouer KJ. Meat and dairy consumption and subsequent risk of prostate cancer in a US cohort study. Cancer Causes Control. 2007;18(1):41-50.

76. Gibson TM, Ferrucci LM, Tangrea JA, Schatzkin A. Epidemiological and clinical studies of nutrition. Semin Oncol. 2010;37(3):282-96.

77. Ubago-Guisado E, Rodríguez-Barranco M, Ching-López A, Petrova D, Molina-Montes E, Amiano P, et al. Evidence update on the relationship between diet and the most common cancers from the European prospective investigation into cancer and nutrition (epic) study: a systematic review. Nutrients. 2021;13(10):3582.

78. Sinha R, Park Y, Graubard BI, Leitzmann MF, Hollenbeck A, Schatzkin A, et al. Meat and meat-related compounds and risk of prostate cancer in a large prospective cohort study in the United States. Am J Epidemiol. 2009;170(9):1165-77.

79. Berquin IM, Min Y, Wu R, Wu J, Perry D, Cline JM, et al. Modulation of prostate cancer genetic risk by omega-3 and omega-6 fatty acids. J Clin Invest. 2007;117(7):1866-75.

80. Daniyal M, Siddiqui ZA, Akram M, Asif HM, Sultana S, Khan A. Epidemiology, etiology, diagnosis and treatment of prostate cancer. Asian Pac J Cancer Prev. 2014;15(22):9575-8.

81. Rodriguez C, McCullough ML, Mondul AM, Jacobs EJ, Fakhrabadi-Shokoohi D, Giovannucci EL, et al. Calcium, dairy products, and risk of prostate cancer in a prospective cohort of United States men. Cancer Epidemiol Biomarkers Prev. 2003;12(7):597-603.

82. Tomaszewski JJ, Cummings JL, Parwani AV, Dhir R, Mason JB, Nelson JB, et al. Increased cancer cell proliferation in prostate cancer patients with high levels of serum folate. The Prostate. 2011;71(12):1287-93.

83. Petersen LF, Brockton NT, Bakkar A, Liu S, Wen J, Weljie AM, et al. Elevated physiological levels of folic acid can increase in vitro growth and invasiveness of prostate cancer cells. BJU Int. 2012;109(5):788-95.

84. Bistulfi G, Foster BA, Karasik E, Gillard B, Miecznikowski J, Dhiman VK, et al. Dietary folate deficiency blocks prostate cancer progression in the TRAMP model. Cancer Prev Res (Phila). 2011;4(11):1825-34.

85. de Vogel S, Meyer K, Fredriksen Å, Ulvik A, Ueland PM, Nygård O, et al. Serum folate and vitamin B12 concentrations in relation to prostate cancer risk--a Norwegian population-based nested case-control study of 3000 cases and 3000 controls within the JANUS cohort. Int J Epidemiol. 2013;42(1):201-10.

86. Kasperzyk JL, Fall K, Mucci LA, Håkansson N, Wolk A, Johansson JE, et al. One-carbon metabolism-related nutrients and prostate cancer survival. Am J Clin Nutr. 2009;90:561-9.

87. Tomaszewski JJ, Richman EL, Sadetsky N, O'Keefe DS, Carroll PR, Davies BJ, et al. Impact of folate intake on prostate cancer recurrence following definitive therapy: Data from CaPSURE™. J Urol. 2014;191:971-6.

88. Banez LL, Hamilton RJ, Partin AW, Vollmer RT, Sun L, Rodriguez C, et al. Obesity-related plasma hemodilution and PSA concentration among men with prostate cancer. JAMA. 2007;298(19):2275-80.

89. Parekh N, Lin Y, Dipaola RS, Marcella S, Lu-Yao G. Obesity and prostate cancer detection: Insights from three national surveys. Am J Med. 2010;123:829-35.

90. Kaaks R, Stattin P. Obesity, Endogenous hormone metabolism, and prostate cancer risk: a conundrum of "highs" and "lows". Cancer Prev Res. 2010;3(3):259-62.

91. Huncharek M, Sue Haddock K, Reid R, Kupelnick B. Smoking as a risk factor for prostate cancer: a meta-analysis of 24 prospective cohort studies. Am J Public Health. 2010;100:693-701.

92. Rohrmann S, Genkinger JM, Burke A, Helzlsouer KJ, Comstock GW, Alberg AJ, et al. Smoking and risk of fatal prostate cancer in a prospective U.S. study. Urology. 2007;69:721-5.

93. Parent M-E, Désy M, Siemiatycki J. Does exposure to agricultural chemicals increase the risk of prostate cancer among farmers? McGill J Med. 2009;12(1):70-7.

94. Ragin C, Davis-Reyes B, Tadesse H, Daniels D, Bunker CH, Jackson M, et al. Farming, reported pesticide use, and prostate cancer. Am J Mens Health. 2013;7(2):102-9.

95. Sauvé JF, Lavoué J, Parent MÉ. Occupation, industry, and the risk of prostate cancer: a case-control study in Montréal, Canada. Environ Health. 2016;15(1):100.

96. Parker C, Castro E, Fizazi K, Heidenreich A, Ost P, Procopio G, et al. Prostate cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. Ann Oncol. 2020;31(9):1119-34.

97. Kim C-J, Dong L, Amend SR, Cho Y-K, Pienta KJ. The role of liquid biopsies in prostate cancer management. Lab on a Chip. 2021;21(17):3263-88.

98. Lima AR, Pinto J, Amaro F, Bastos MdL, Carvalho M, Guedes de Pinho P. Advances and perspectives in prostate cancer biomarker discovery in the last 5 years through tissue and urine metabolomics. Metabolites. 2021;11(3):181.

99. Louie KS, Seigneurin A, Cathcart P, Sasieni P. Do prostate cancer risk models improve the predictive accuracy of PSA screening? A meta-analysis. Ann Oncol. 2015;26(5):848-64.

100. Das CJ, Razik A, Sharma S, Verma S. Prostate biopsy: when and how to perform. Clin Radiol. 2019;74(11):853-64.

101. McDunn JE, Li Z, Adam KP, Neri BP, Wolfert RL, Milburn MV, et al. Metabolomic signatures of aggressive prostate cancer. Prostate. 2013;73(14):1547-60.

102. Dimakakos A, Armakolas A, Koutsilieris M. Novel Tools for prostate cancer prognosis, diagnosis, and follow-up. BioMed Res Int. 2014;2014:890697.

103. Barry MJ. Prevention of prostate cancer morbidity and mortality primary prevention and early detection. Medical Clinics of NA. 2017;101(4):787-806.

104. Pal RP, Maitra NU, Mellon JK, Khan MA. Defining prostate cancer risk before prostate biopsy. Urol Oncol. 2013;31(8):1408-18.

105. Rigau M, Olivan M, Garcia M, Sequeiros T, Montes M, Colás E, et al. The present and future of prostate cancer urine biomarkers. Int J Mol Sci. 2013;14(6):12620-49.

106. Lima AR, Pinto J, Azevedo AI, Barros-Silva D, Jerónimo C, Henrique R, et al. Identification of a biomarker panel for improvement of prostate cancer diagnosis by volatile metabolic profiling of urine. Br J Cancer. 2019;121(10):857-68.

107. Gaglani S, Gonzalez-Kozlova E, Lundon DJ, Tewari AK, Dogra N, Kyprianou N. Exosomes as a next-generation diagnostic and therapeutic tool in prostate cancer. Int J Mol Sci. 2021;22(18):10131.

108. Lima AR, Bastos MdL, Carvalho M, Guedes de Pinho P. Biomarker discovery in human prostate cancer: an update in metabolomics studies. Trans Oncol. 2016;9(4):357-70.

109. Campos-Fernández E, Barcelos LS, de Souza AG, Goulart LR, Alonso-Goulart V. Research landscape of liquid biopsies in prostate cancer. Am J Cancer Res. 2019;9(7):1309-28.

110. Andriole GL, Crawford ED, Grubb RL, 3rd, Buys SS, Chia D, Church TR, et al. Prostate cancer screening in the randomized prostate, lung, colorectal, and ovarian cancer screening trial: Mortality results after 13 years of follow-up. J Natl Cancer Inst. 2012;104(2):125-32.

111. Parsi M, Desai MH, Desai D, Singhal S, Khandwala PM, Potdar RR. PSMA: a game changer in the diagnosis and treatment of advanced prostate cancer. Med Oncol. 2021;38(8):89.

112. Uijen MJM, Derks YHW, Merkx RIJ, Schilham MGM, Roosen J, Privé BM, et al. PSMA radioligand therapy for solid tumors other than prostate cancer: background, opportunities, challenges, and first clinical reports. Eur J Nucl Med Mol Imaging. 2021;48(13):4350-68.

113. Seifert R, Alberts IL, Afshar-Oromieh A, Rahbar K. Prostate Cancer Theranostics: PSMA targeted therapy. PET Clin. 2021;16(3):391-6.

114. Hofman MS, Violet J, Hicks RJ, Ferdinandus J, Thang SP, Akhurst T, et al. [(177)Lu]-PSMA-617 radionuclide treatment in patients with metastatic castration-resistant prostate cancer (LuPSMA trial): a single-centre, single-arm, phase 2 study. Lancet Oncol. 2018;19(6):825-33.

115. Hofman M, Violet JA, Hicks RJ, Ferdinandus J, Thang SP, Iravani A, et al. Results of a 50 patient single-center phase II prospective trial of Lutetium-177 PSMA-617 theranostics in metastatic castrate-resistant prostate cancer. J Clin Oncol. 2019;37(7_suppl):228.

116. Calais J, Fendler WP, Eiber M, Lassmann M, Dahlbom M, Esfandiari R, et al. RESIST-PC phase 2 trial: 177Lu-PSMA-617 radionuclide therapy for metastatic castrate-resistant prostate cancer. J Clin Oncol. 2019;37(15_suppl):5028.

117. Moradi F, Farolfi A, Fanti S, Iagaru A. Prostate cancer: Molecular imaging and MRI. Eur J Radiol. 2021;143:109893.

118. Mena E, Black PC, Rais-Bahrami S, Gorin M, Allaf M, Choyke P. Novel PET imaging methods for prostate cancer. World J Urol. 2021;39(3):687-99.

119. Thomas L, Balmus C, Ahmadzadehfar H, Essler M, Strunk H, Bundschuh RA. Assessment of bone metastases in patients with prostate cancer-A comparison between (99m)Tc-Bone-Scintigraphy and [(68)Ga]Ga-PSMA PET/CT. Pharmaceuticals (Basel). 2017;10(3):68.

120. Oey O, Ghaffari M, Li JJ, Hosseini-Beheshti E. Application of extracellular vesicles in the diagnosis and treatment of prostate cancer: Implications for clinical practice. Crit Rev Oncol Hematol. 2021;167:103495.

121. Ludwig M, Rajvansh R, Drake JM. Emerging role of extracellular vesicles in prostate cancer. Endocrinology. 2021;162(9):bqab139.

122. Lorenc T, Klimczyk K, Michalczewska I, Słomka M, Kubiak-Tomaszewska G, Olejarz W. Exosomes in prostate cancer diagnosis, prognosis and therapy. Int J Mol Sci. 2020;21(6):2118.

123. McKiernan J, Donovan MJ, O'Neill V, Bentink S, Noerholm M, Belzer S, et al. A novel urine exosome gene expression assay to predict high-grade prostate cancer at initial biopsy. JAMA Oncol. 2016;2(7):882-9.

124. Ji J, Chen R, Zhao L, Xu Y, Cao Z, Xu H, et al. Circulating exosomal mRNA profiling identifies novel signatures for the detection of prostate cancer. Mol Cancer. 2021;20(1):58.

125. Xu Y-H, Deng J-L, Wang G, Zhu Y-S. Long non-coding RNAs in prostate cancer: Functional roles and clinical implications. Cancer Letters. 2019;464:37-55.

126. Li W, Dou Z, We S, Zhu Z, Pan D, Jia Z, et al. Long noncoding RNA BDNF-AS is associated with clinical outcomes and has functional role in human prostate cancer. Biomed Pharmacother. 2018;102:1105-10.

127. Zhao R, Sun F, Bei X, Wang X, Zhu Y, Jiang C, et al. Upregulation of the long non-coding RNA FALEC promotes proliferation and migration of prostate cancer cell lines and predicts prognosis of PCa patients. The Prostate. 2017;77(10):1107-17.

128. Hu JC, Wang SS, Chou YE, Chiu KY, Li JR, Chen CS, et al. Associations between lncRNA MALAT1 polymorphisms and lymph node metastasis in prostate cancer. Diagnostics (Basel). 2021;11(9):1692.

129. Li Y, Ji J, Lyu J, Jin X, He X, Mo S, et al. A novel urine exosomal lncRNA assay to improve the detection of prostate cancer at initial biopsy: a retrospective multicenter diagnostic feasibility study. Cancers (Basel). 2021;13(16):4075.

130. Kidd SG, Carm KT, Bogaard M, Olsen LG, Bakken AC, Løvf M, et al. High expression of SCHLAP1 in primary prostate cancer is an independent predictor of biochemical recurrence, despite substantial heterogeneity. Neoplasia. 2021;23(6):634-41.

131. Huang W, Su X, Yan W, Kong Z, Wang D, Huang Y, et al. Overexpression of AR-regulated lncRNA TMPO-AS1 correlates with tumor progression and poor prognosis in prostate cancer. Prostate. 2018;78(16):1248-61.

132. Coleman RE, Lipton A, Roodman GD, Guise TA, Boyce BF, Brufsky AM, et al. Metastasis and bone loss: advancing treatment and prevention. Cancer Treat Rev. 2010;36(8):615-20.

133. Parker C, Nilsson S, Heinrich D, Helle SI, O'Sullivan JM, Fosså SD, et al. Alpha emitter radium-223 and survival in metastatic prostate cancer. N Engl J Med. 2013;369(3):213-23.

134. Maffioli L, Florimonte L, Costa D, Castanheira Correira J, Grana C, Luster M, et al. New radiopharmaceutical agents for the treatment of castration-resistant prostate cancer. Q J Nucl Med Mol Imaging. 2015;59(4):420-38.

135. Du Y, Dizdarevic S. Molecular radiotheragnostics in prostate cancer. Clin Med (Lond). 2017;17(5):458-61.

136. Retter A, Gong F, Syer T, Singh S, Adeleke S, Punwani S. Emerging methods for prostate cancer imaging: evaluating cancer structure and metabolic alterations more clearly. Mol Oncol. 2021;15(10):2565-79.

137. Maurer T, Gschwend JE, Rauscher I, Souvatzoglou M, Haller B, Weirich G, et al. Diagnostic efficacy of (68)gallium-PSMA positron emission tomography compared to conventional imaging for lymph node staging of 130 consecutive patients with intermediate to high risk prostate cancer. J Urol. 2016;195(5):1436-43.

138. Cursano MC, Iuliani M, Casadei C, Stellato M, Tonini G, Paganelli G, et al. Combination radium-223 therapies in patients with bone metastases from castration-resistant prostate cancer: A review. Crit Rev Oncol Hematol. 2020;146:102864.

139. Afshar-Oromieh A, Babich JW, Kratochwil C, Giesel FL, Eisenhut M, Kopka K, et al. The rise of PSMA ligands for diagnosis and therapy of prostate cancer. J Nucl Med. 2016;57(Suppl 3):79S-89S.

140. Rahbar K, Ahmadzadehfar H, Kratochwil C, Haberkorn U, Schäfers M, Essler M, et al. German multicenter study investigating 177Lu-PSMA-617 radioligand therapy in advanced prostate cancer patients. J Nucl Med. 2017;58(1):85-90.

141. Scarpa L, Buxbaum S, Kendler D, Fink K, Bektic J, Gruber L, et al. The (68)Ga/(177)Lu theragnostic concept in PSMA targeting of castration-resistant prostate cancer: correlation of SUV(max) values and absorbed dose estimates. Eur J Nucl Med Mol Imaging. 2017;44(5):788-800.

142. Morrison GJ, Goldkorn A. Development and application of liquid biopsies in metastatic prostate cancer. Curr Oncol Rep. 2018;20(4):35.

143. Matuszczak M, Schalken JA, Salagierski M. Prostate cancer liquid biopsy biomarkers' clinical utility in diagnosis and prognosis. Cancers. 2021;13(13):3373.

144. Oto J, Fernández-Pardo Á, Royo M, Hervás D, Martos L, Vera-Donoso CD, et al. A predictive model for prostate cancer incorporating PSA molecular forms and age. Sci Rep. 2020;10(1):2463.

145. Auprich M, Bjartell A, Chun FK, de la Taille A, Freedland SJ, Haese A, et al. Contemporary role of prostate cancer antigen 3 in the management of prostate cancer. Eur Urol. 2011;60(5):1045-54.

146. Parekh DJ, Punnen S, Sjoberg DD, Asroff SW, Bailen JL, Cochran JS, et al. A multi-institutional prospective trial in the USA confirms that the 4Kscore accurately identifies men with high-grade prostate cancer. Eur Urol. 2015;68(3):464-70.

147. Nordström T, Vickers A, Assel M, Lilja H, Grönberg H, Eklund M. Comparison between the Four-Kallikrein panel and Prostate Health Index for predicting prostate cancer. Eur Urol. 2015;68(1):139-46.

148. Zappala SM, Scardino PT, Okrongly D, Linder V, Dong Y. Clinical performance of the 4Kscore test to predict high-grade prostate cancer at biopsy: a meta-analysis of US and European clinical validation study results. Rev Urol. 2017;19(3):149-55.

149. White J, Shenoy BV, Tutrone RF, Karsh LI, Saltzstein DR, Harmon WJ, et al. Clinical utility of the Prostate Health Index (phi) for biopsy decision management in a large group urology practice setting. Prostate Cancer Prostatic Dis. 2018;21(1):78-84.

150. Saidi S, Al Riyami N, Marhoon M, Saraf M, Busaidi S, Mula-Abed W-A, et al. Validity of prostate health index and percentage of [-2] pro-prostate-specific antigen as novel biomarkers

in the diagnosis of prostate cancer: Omani tertiary hospitals experience. Oman Med J. 2017;32(4):275-83.

151. Rodríguez SVM, García-Perdomo HA. Diagnostic accuracy of prostate cancer antigen 3 (PCA3) prior to first prostate biopsy: a systematic review and meta-analysis. Can Urol Assoc J. 2020;14(5):E214-e9.

152. Nicholson A, Mahon J, Boland A, Beale S, Dwan K, Fleeman N, et al. The clinical effectiveness and cost-effectiveness of the PROGENSA® prostate cancer antigen 3 assay and the Prostate Health Index in the diagnosis of prostate cancer: a systematic review and economic evaluation. Health Technol Assess. 2015;19(87):i-xxxi, 1-191.

153. Van Neste L, Hendriks RJ, Dijkstra S, Trooskens G, Cornel EB, Jannink SA, et al. Detection of high-grade prostate cancer using a urinary molecular biomarker-based risk score. Eur Urol. 2016;70(5):740-8.

154. Tomlins SA, Day JR, Lonigro RJ, Hovelson DH, Siddiqui J, Kunju LP, et al. Urine TMPRSS2:ERG Plus PCA3 for individualized prostate cancer risk assessment. Eur Urol. 2016;70(1):45-53.

155. BridgeSpan. Gene-Based Tests for Screening, Detection, and/or Management of Prostate Cancer. In Genetic Testing, Policy No.17; BridgeSpan: Boston, MA, USA, 2020.

156. Ross AE, Feng FY, Ghadessi M, Erho N, Crisan A, Buerki C, et al. A genomic classifier predicting metastatic disease progression in men with biochemical recurrence after prostatectomy. Prostate Cancer Prostatic Dis. 2014;17(1):64-9.

157. Narayan VM, Konety BR, Warlick C. Novel biomarkers for prostate cancer: An evidence-based review for use in clinical practice. Int J Urol. 2017;24(5):352-60.

158. Chistiakov DA, Myasoedova VA, Grechko AV, Melnichenko AA, Orekhov AN. New biomarkers for diagnosis and prognosis of localized prostate cancer. Semin Cancer Biol. 2018;52(Pt 1):9-16.

159. Porzycki P, Ciszkowicz E. Modern biomarkers in prostate cancer diagnosis. Cent Eur J Urol. 2020;73(3):300-6.

160. Bae J, Yang S-H, Kim A, Kim HG. RNA-based biomarkers for the diagnosis, prognosis, and therapeutic response monitoring of prostate cancer. Urol Oncol. 2022;40(3):105.e1-.e10.

161. Kan Y, Li B, Yang D, Liu Y, Liu J, Yang C, et al. Emerging roles of long non-coding RNAs as novel biomarkers in the diagnosis and prognosis of prostate cancer. Discov Med. 2021;32(165):29-37.

162. Crawford ED, Ventii K, Shore ND. New biomarkers in prostate cancer. Oncology (Williston Park). 2014;28(2):135-42.

163. Clinton TN, Bagrodia A, Lotan Y, Margulis V, Raj GV, Woldu SL. Tissue-based biomarkers in prostate cancer. Expert Rev Precis Med Drug Dev. 2017;2(5):249-60.

164. Kretschmer A, Tilki D. Biomarkers in prostate cancer – Current clinical utility and future perspectives. Crit Rev Oncol Hematol. 2017;120:180-93.

165. Hanjani NA, Esmaelizad N, Zanganeh S, Gharavi AT, Heidarizadeh P, Radfar M, et al. Emerging role of exosomes as biomarkers in cancer treatment and diagnosis. Crit Rev Oncol Hematol. 2022;169:103565.

166. Wang J, Ni J, Beretov J, Thompson J, Graham P, Li Y. Exosomal microRNAs as liquid biopsy biomarkers in prostate cancer. Crit Rev Oncol Hematol. 2020;145:102860.

167. Van Poppel H, Roobol MJ, Chapple CR, Catto JWF, N'Dow J, Sønksen J, et al. Prostate-specific antigen testing as part of a risk-adapted early detection strategy for prostate cancer: European association of urology position and recommendations for 2021. Eur Urol. 2021;80(6):703-11.

168. Litwin MS, Tan HJ. The diagnosis and treatment of prostate cancer: A review. JAMA. 2017;317:2532-42.

169. Ossoliński K, Nizioł J, Arendowski A, Ossolińska A, Ossoliński T, Kucharz J, et al. Mass spectrometry-based metabolomic profiling of prostate cancer - a pilot study. J Cancer Metastasis Treat. 2019;5:1.

170. de Cobelli O, Terracciano D, Tagliabue E, Raimondi S, Galasso G, Cioffi A, et al. Body mass index was associated with upstaging and upgrading in patients with low-risk prostate cancer who met the inclusion criteria for active surveillance. Urol Oncol. 2015;33(5):201.e1-8.

171. Ferro M, Lucarelli G, Bruzzese D, Di Lorenzo G, Perdonà S, Autorino R, et al. Low serum total testosterone level as a predictor of upstaging and upgrading in low-risk prostate cancer patients meeting the inclusion criteria for active surveillance. Oncotarget. 2017;8(11):18424-34.

172. Monteiro M, Carvalho M, Bastos M, Guedes de Pinho P. Biomarkers in renal cell carcinoma: a metabolomics approach. Metabolomics. 2014;10:1210-22.

173. Trock BJ. Application of metabolomics to prostate cancer. Urol Oncol. 2011;29(5):572-81.

174. Rudaz S. Identification and Data-Processing Methods in Metabolomics.vIn Identification and Data Processing Methods in Metabolomics. Future Science Book Series: Future Science Ltd; London, UK, 2015. p. 2-5.

175. Silva C, Perestrelo R, Silva P, Tomás H, Câmara JS. Breast cancer metabolomics: from analytical platforms to multivariate data analysis. A review. Metabolites. 2019;9(5):102.

176. Janfaza S, Khorsand B, Nikkhah M, Zahiri J. Digging deeper into volatile organic compounds associated with cancer. Biol Methods Protoc. 2019;4(1):bpz014.

177. Spratlin JL, Serkova NJ, Gail Eckhardt S. Clinical applications of metabolomics in oncology: a review. Clin Cancer Res. 2009;15(2):431-40.

178. Aggarwal P, Baker J, Boyd MT, Coyle S, Probert C, Chapman EA. Optimisation of urine sample preparation for headspace-solid phase microextraction gas chromatography-mass spectrometry: Altering sample ph, sulphuric acid concentration and phase ratio. Metabolites. 2020;10(12):482.

179. Živković Semren T, Brčić Karačonji I, Safner T, Brajenović N, Tariba Lovaković B, Pizent A. Gas chromatographic-mass spectrometric analysis of urinary volatile organic metabolites: Optimization of the HS-SPME procedure and sample storage conditions. Talanta. 2018;176:537-43.

180. Lima AR, Araújo AM, Pinto J, Jerónimo C, Henrique R, Bastos MdL, et al. Discrimination between the human prostate normal and cancer cell exometabolome by GC-MS. Sci Rep. 2018;8(1):5539.

181. Câmara JS, Perestrelo R, Berenguer CV, Andrade CFP, Gomes TM, Olayanju B, et al. Green extraction techniques as advanced sample preparation approaches in biological, food, and environmental matrices: a review. Molecules. 2022;27(9):2953.

182. Arthur CL, Pawliszyn J. Solid phase microextraction with thermal desorption using fused silica optical fibers. Anal Chem. 1990;62(19):2145-8.

183. Silva CL, Passos M, Câmara JS. Solid phase microextraction, mass spectrometry and metabolomic approaches for detection of potential urinary cancer biomarkers - A powerful strategy for breast cancer diagnosis. Talanta. 2012;89:360-8.

184. Porto-Figueira P, Pereira J, Miekisch W, Câmara JS. Exploring the potential of NTME/GC-MS, in the establishment of urinary volatomic profiles. Lung cancer patients as case study. Sci Rep. 2018;8(1):13113.

185. Capelli L, Taverna G, Bellini A, Eusebio L, Buffi N, Lazzeri M, et al. Application and uses of electronic noses for clinical diagnosis on urine samples: a review. Sensors (Basel). 2016;16(10):1708.

186. Filianoti A, Costantini M, Bove AM, Anceschi U, Brassetti A, Ferriero M, et al. Volatilome analysis in prostate cancer by electronic nose: a pilot monocentric study. Cancers (Basel). 2022;14(12):2927.

187. Tyagi H, Daulton E, Bannaga AS, Arasaradnam RP, Covington JA. Urinary volatiles and chemical characterisation for the non-invasive detection of prostate and bladder cancers. Biosensors (Basel). 2021;11(11):437.

188. Bax C, Taverna G, Eusebio L, Sironi S, Grizzi F, Guazzoni G, et al. Innovative diagnostic methods for early prostate cancer detection through urine analysis: a review. Cancers (Basel). 2018;10(4):123.

189. Gao Q, Lee WY. Urinary metabolites for urological cancer detection: a review on the application of volatile organic compounds for cancers. Am J Clin Exp Urol. 2019;7(4):232-48.

190. Opitz P, Herbarth O. The volatilome – investigation of volatile organic metabolites (VOM) as potential tumor markers in patients with head and neck squamous cell carcinoma (HNSCC). J Otolaryngol - Head Neck Surg. 2018;47(1):42.

191. Taware R, Taunk K, Pereira JAM, Dhakne R, Kannan N, Soneji D, et al. Investigation of urinary volatomic alterations in head and neck cancer: a non-invasive approach towards diagnosis and prognosis. Metabolomics. 2017;13(10):111.

192. Silva CL, Passos M, Camara JS. Investigation of urinary volatile organic metabolites as potential cancer biomarkers by solid-phase microextraction in combination with gas chromatography-mass spectrometry. Br J Cancer. 2011;105(12):1894-904.

193. Wang D, Wang C, Pi X, Guo L, Wang Y, Li M, et al. Urinary volatile organic compounds as potential biomarkers for renal cell carcinoma. Biomed Rep. 2016;5(1):68-72.

194. Daulton E, Wicaksono AN, Tiele A, Kocher HM, Debernardi S, Crnogorac-Jurcevic T, et al. Volatile organic compounds (VOCs) for the non-invasive detection of pancreatic cancer from urine. Talanta. 2021;221:121604.

195. Rudnicka J, Kowalkowski T, Buszewski B. Searching for selected VOCs in human breath samples as potential markers of lung cancer. Lung Cancer. 2019;135:123-9.

196. Kischkel S, Miekisch W, Fuchs P, Schubert JK. Breath analysis during one-lung ventilation in cancer patients. Eur Respir J. 2012;40(3):706-13.

197. Markar SR, Brodie B, Chin ST, Romano A, Spalding D, Hanna GB. Profile of exhaled-breath volatile organic compounds to diagnose pancreatic cancer. Br J Surg. 2018;105(11):1493-500.

198. Altomare DF, Di Lena M, Porcelli F, Trizio L, Travaglio E, Tutino M, et al. Exhaled volatile organic compounds identify patients with colorectal cancer. Br J Surg. 2013;100(1):144-50.

199. Jung YJ, Seo HS, Kim JH, Song KY, Park CH, Lee HH. Advanced diagnostic technology of volatile organic compounds real time analysis analysis from exhaled breath of

gastric cancer patients using proton-transfer-reaction time-of-flight mass spectrometry. Front Oncol. 2021;11:56059.

200. Taware R, Taunk K, Pereira JAM, Shirolkar A, Soneji D, Câmara JS, et al. Volatilomic insight of head and neck cancer via the effects observed on saliva metabolites. Sci Rep. 2018;8(1):17725.

201. Bel'skaya LV, Sarf EA, Shalygin SP, Postnova TV, Kosenok VK. Identification of salivary volatile organic compounds as potential markers of stomach and colorectal cancer: A pilot study. J Oral Biosci. 2020;62(2):212-21.

202. Cavaco C, Pereira JAM, Taunk K, Taware R, Rapole S, Nagarajaram H, et al. Screening of salivary volatiles for putative breast cancer discrimination: an exploratory study involving geographically distant populations. Anal Bioanal Chem. 2018;410(18):4459-68.

203. Cavaco C, Perestrelo R, Silva C, Aveiro F, and PJ, JS C. Establishment of the saliva volatomic profile as an exploratory and non-invasive strategy to find potential breast cancer biomarkers. Int Labmate. 2014;39(4):4-5.

204. Monedeiro F, Monedeiro-Milanowski M, Zmysłowski H, De Martinis BS, Buszewski B. Evaluation of salivary VOC profile composition directed towards oral cancer and oral lesion assessment. Clin Oral Investig. 2021;25(7):4415-30.

205. Shirasu M, Touhara K. The scent of disease: volatile organic compounds of the human body related to disease and disorder. J Biochem (Tokyo). 2011;150(3):257-66.

206. Pereira J, Porto-Figueira P, Cavaco C, Taunk K, Rapole S, Dhakne R, et al. Breath analysis as a potential and non-invasive frontier in disease diagnosis: an overview. Metabolites. 2015;5(1):3-55.

207. Tienpont B, David F, Bicchi C, Sandra P. High capacity headspace sorptive extraction. J Microcolumn Sep. 2000;12(11):577-84.

208. Pugliese G, Trefz P, Brock B, Schubert JK, Miekisch W. Extending PTR based breath analysis to real-time monitoring of reactive volatile organic compounds. Analyst. 2019;144(24):7359-67.

209. Streckfus CF, Brown RE, Bull JM. Proteomics, morphoproteomics, saliva and breast cancer: an emerging approach to guide the delivery of individualised thermal therapy, thermochemotherapy and monitor therapy response. Int J Hyperthermia. 2010;26(7):649-61.

210. Amann A, Costello BdL, Miekisch W, Schubert J, Buszewski B, Pleil J, et al. The human volatilome: volatile organic compounds (VOCs) in exhaled breath, skin emanations, urine, feces and saliva. J Breath Res. 2014;8(3):034001.

211. Malathi N, Mythili S, Vasanthi HR. Salivary diagnostics: a brief review. ISRN Dentistry. 2014;2014:1-8.

212. Pereira JAM, Taware R, Porto-Figueira P, Rapole S, Câmara JS. The salivary volatome in breast cancer. Elsevier; Amsterdam, The Netherlands, 2020. p. 301-7.

213. Lima AR, Pinto J, Carvalho-Maia C, Jerónimo C, Henrique R, Bastos MdL, et al. A panel of urinary volatile biomarkers for differential diagnosis of prostate cancer from other urological cancers. Cancers. 2020;12(8):2017.

214. da Costa BRB, De Martinis BS. Analysis of urinary VOCs using mass spectrometric methods to diagnose cancer: A review. Clin Mass Spectrom. 2020;18:27-37.

215. Dai C, Heemers H, Sharifi N. Androgen signaling in prostate cancer. Cold Spring Harbor Perspect Med. 2017;7(9):a030452.

216. Koppenol WH, Bounds PL, Dang CV. Otto Warburg's contributions to current concepts of cancer metabolism. Nat Rev Cancer. 2011;11(5):325-37.

217. Lima AR, Pinto J, Barros-Silva D, Jerónimo C, Henrique R, Bastos MdL, et al. New findings on urinary prostate cancer metabolome through combined GC–MS and 1H NMR analytical platforms. Metabolomics. 2020;16:1-9.

218. Chen TC, Sakaki T, Yamamoto K, Kittaka A. The roles of cytochrome P450 enzymes in prostate cancer development and treatment. Anticancer Res. 2012;32(1):291.

219. Khalid T, Aggio R, White P, De Lacy Costello B, Persad R, Al-Kateb H, et al. Urinary volatile organic compounds for the detection of prostate cancer. PLOS ONE. 2015;10(11):e0143283.

220. Struck-Lewicka W, Kordalewska M, Bujak R, Yumba Mpanga A, Markuszewski M, Jacyna J, et al. Urine metabolic fingerprinting using LC-MS and GC-MS reveals metabolite changes in prostate cancer: A pilot study. J Pharm Biomed Anal. 2015;111:351-61.

221. Gao Q, Su X, Annabi MH, Schreiter BR, Prince T, Ackerman A, et al. Application of urinary volatile organic compounds (VOCs) for the diagnosis of prostate cancer. Clin Genitourin Cancer. 2019;17(3):183-90.

222. Jiménez-Pacheco A, Salinero-Bachiller M, Iribar MC, López-Luque A, Miján-Ortiz JL, Peinado JM. Furan and p-xylene as candidate biomarkers for prostate cancer. Urol Oncol. 2018;36(5):243.e21-.e27.

223. Peng G, Hakim M, Broza YY, Billan S, Abdah-Bortnyak R, Kuten A, et al. Detection of lung, breast, colorectal, and prostate cancers from exhaled breath using a single array of nanosensors. Br J Cancer. 2010;103(4):542-51.

224. Taverna G, Grizzi F, Tidu L, Bax C, Zanoni M, Vota P, et al. Accuracy of a new electronic nose for prostate cancer diagnosis in urine samples. Int J Urol. 2022;29:890-6.

225. Capelli L, Bax C, Grizzi F, Taverna G. Optimization of training and measurement protocol for eNose analysis of urine headspace aimed at prostate cancer diagnosis. Sci Rep. 2021;11(1):20898.

226. Bannaga AS, Kvasnik F, Persaud K, Arasaradnam RP. Differentiating cancer types using a urine test for volatile organic compounds. J Breath Res. 2020;15(1):017102.

227. Waltman CG, Marcelissen TAT, van Roermund JGH. Exhaled-breath testing for prostate cancer based on volatile organic compound profiling using an electronic nose device (Aeonose™): a preliminary report. Eur Urol Focus. 2020;6(6):1220-5.

228. Worley B, Powers R. Multivariate analysis in metabolomics. Curr Metabolomics. 2013;1(1):92-107.

229. Trygg J, Holmes E, Lundstedt T. Chemometrics in metabonomics. J Proteome Res. 2007;6(2):469-79.

230. Monteiro M. Metabolomics Analysis for Renal Cell Carcinoma Detection and Biomarker Discovery [dissertation]. Porto: Porto University; 2016.

231. Smolinska A, Hauschild AC, Fijten RR, Dallinga JW, Baumbach J, van Schooten FJ. Current breathomics - a review on data pre-processing techniques and machine learning in metabolomics breath analysis. J Breath Res. 2014;8(2):027105.

232. Pang Z, Chong J, Zhou G, de Lima Morais DA, Chang L, Barrette M, et al. MetaboAnalyst 5.0: narrowing the gap between raw spectra and functional insights. Nucleic Acids Res. 2021;49(W1):W388-W96.

233. Taunk K, Porto-Figueira P, Pereira JAM, Taware R, da Costa NL, Barbosa R, et al. Urinary volatomic expression pattern: paving the way for identification of potential candidate biosignatures for lung cancer. Metabolites. 2022;12(1):36.

234. Schmidt K, Podmore I. Current challenges in volatile organic compounds analysis as potential biomarkers of cancer. J Biomark. 2015;2015:981458.

235. Pleil JD, Stiegel MA, Risby TH. Clinical breath analysis: discriminating between human endogenous compounds and exogenous (environmental) chemical confounders. J Breath Res. 2013;7(1):017107.

236. Hasnip S, Crews C, Castle L. Some factors affecting the formation of furan in heated foods. Food Addit Contam. 2006;23(3):219-27.

237. Wegener JW, López-Sánchez P. Furan levels in fruit and vegetables juices, nutrition drinks and bakery products. Anal Chim Acta. 2010;672(1-2):55-60.

238. Wishart DS, Guo A, Oler E, Wang F, Anjum A, Peters H, et al. HMDB 5.0: the Human Metabolome Database for 2022. Nucleic Acids Res. 2022;50(D1):D622-D31.

239. de Lacy Costello B, Amann A, Al-Kateb H, Flynn C, Filipiak W, Khalid T, et al. A review of the volatiles from the healthy human body. J Breath Res. 2014;8(1):014001.

240. Buszewski B, Ulanowska A, Ligor T, Jackowski M, Kłodzińska E, Szeliga J. Identification of volatile organic compounds secreted from cancer tissues and bacterial cultures. J Chromatogr B Analyt Technol Biomed Life Sci. 2008;868(1-2):88-94.

241. Mills GA, Walker V. Headspace solid-phase microextraction profiling of volatile compounds in urine: Application to metabolic investigations. J Chromatogr B Biomed Appl. 2001;753(2):259-68.

242. Smith S, Burden H, Persad R, Whittington K, de Lacy Costello B, Ratcliffe NM, et al. A comparative study of the analysis of human urine headspace using gas chromatography–mass spectrometry. J Breath Res. 2008;2(3):037022.

243. Miekisch W, Schubert JK, Vagts DA, Geiger K. Analysis of volatile disease markers in blood. Clin Chem. 2001;47(6):1053-60.

244. Tangerman A. Measurement and biological significance of the volatile sulfur compounds hydrogen sulfide, methanethiol and dimethyl sulfide in various biological matrices. J Chromatogr B Analyt Technol Biomed Life Sci. 2009;877(28):3366-77.

245. Blom HJ, Boers GH, van den Elzen JP, Gahl WA, Tangerman A. Transamination of methionine in humans. Clin Sci (Lond). 1989;76(1):43-9.

246. Schöller C, Molin S, Wilkins K. Volatile metabolites from some gram-negative bacteria. Chemosphere. 1997;35(7):1487-95.

247. Garner CE, Smith S, de Lacy Costello B, White P, Spencer R, Probert CS, et al. Volatile organic compounds from feces and their potential for diagnosis of gastrointestinal disease. FASEB J. 2007;21(8):1675-88.

248. Filipiak W, Sponring A, Baur MM, Filipiak A, Ager C, Wiesenhofer H, et al. Molecular analysis of volatile metabolites released specifically by Staphylococcus aureus and Pseudomonas aeruginosa. BMC Microbiol. 2012;12:113.

249. Preuss R, Angerer J, Drexler H. Naphthalene - an environmental and occupational toxicant. Int Arch Occup Environ Health. 2003;76(8):556-76.

250. Wahl HG, Hoffmann A, Luft D, Liebich HM. Analysis of volatile organic compounds in human urine by headspace gas chromatography-mass spectrometry with a multipurpose sampler. J Chromatogr A. 1999;847(1-2):117-25.

251. Hamilton DF, Ghert M, Simpson AH. Interpreting regression models in clinical outcome studies. Bone Joint Res. 2015;4(9):152-3.

252. Xie X, Zhang Y, Ge C, Liang P. Effect of Brachytherapy vs. External beam radiotherapy on sexual function in patients with clinically localized prostate cancer: a meta-analysis. Front Cell Dev Biol. 2021;9:792597.

253. Wallis CJD, Glaser A, Hu JC, Huland H, Lawrentschuk N, Moon D, et al. Survival and complications following surgery and radiation for localized prostate cancer: an international collaborative review. Eur Urol. 2018;73(1):11-20.

APPENDICES.

SUPPLEMENTARY MATERIAL

Additional figures and tables

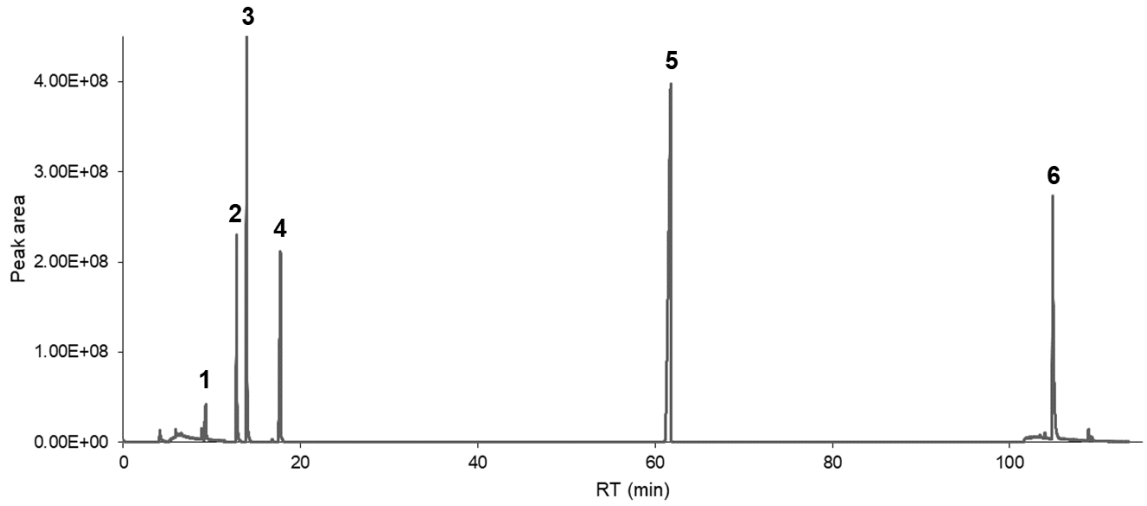


Figure S.1 – Representative chromatogram of the pure standards applied to confirm the VOMs identified using the NIST library. Legend: 1: 2-pentanone (9.31 min); 2: 3-hexanone (12.80 min); 3: dimethyl disulfide (13.94 min); 4: 4-heptanone (17.75 min); 5: 2-ethyl-1-hexanol (61.76 min); 6: octanoic acid (104.84 min).

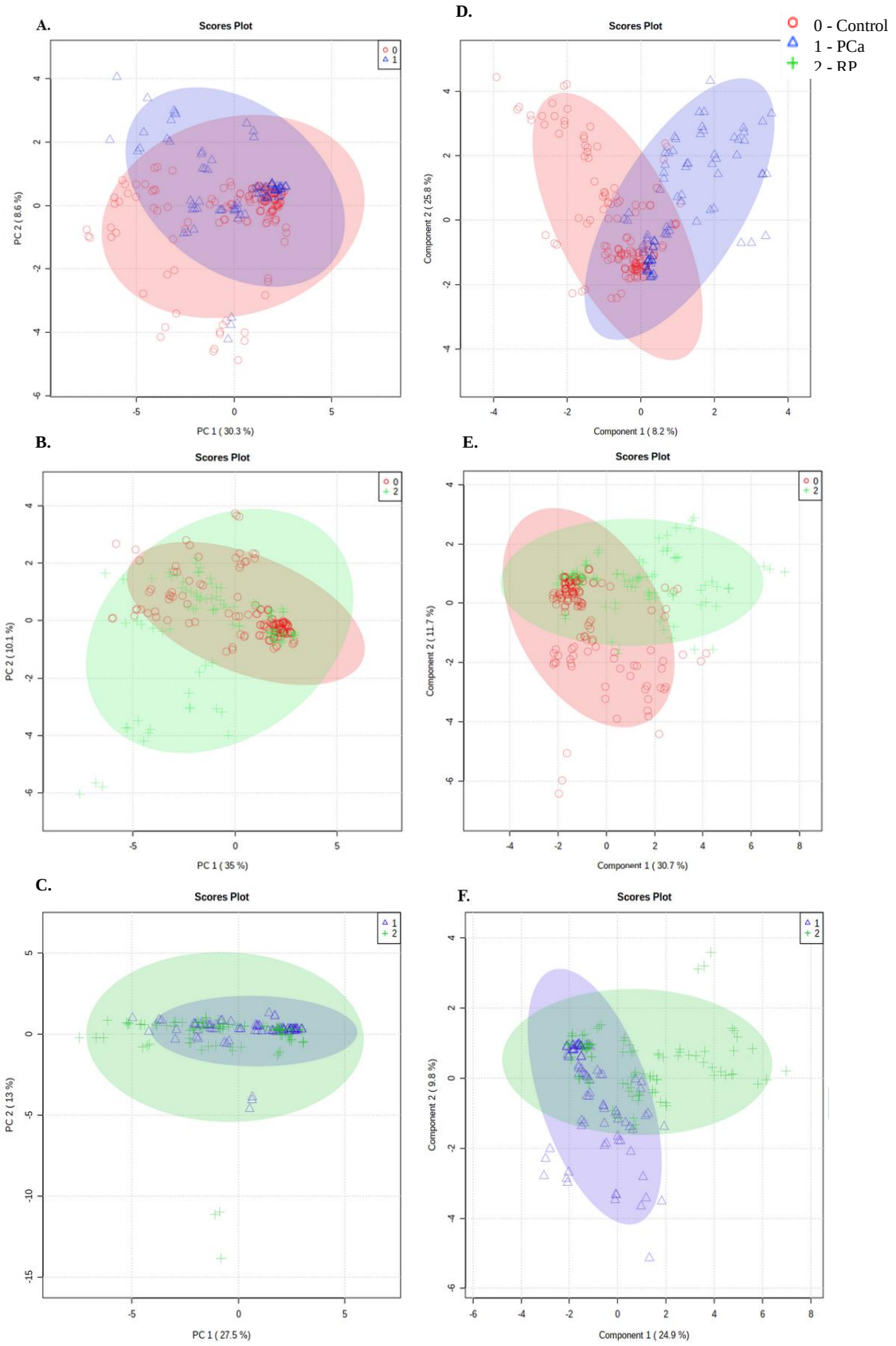


Figure S.2 – Score scatter plots of the volatile profile of urine samples from the groups under study. **A-C:** PCA PC1 x PC2 score scatter plot of **A.** control (0) *vs* PCa (1), **B.** control (0) *vs* RP (3), and **C.** PCa (1) *vs* RP (3). **C.** PCA PC1 × PC2 score scatter plot. **D-F:** Corresponding PLS-DA score scatter plots.

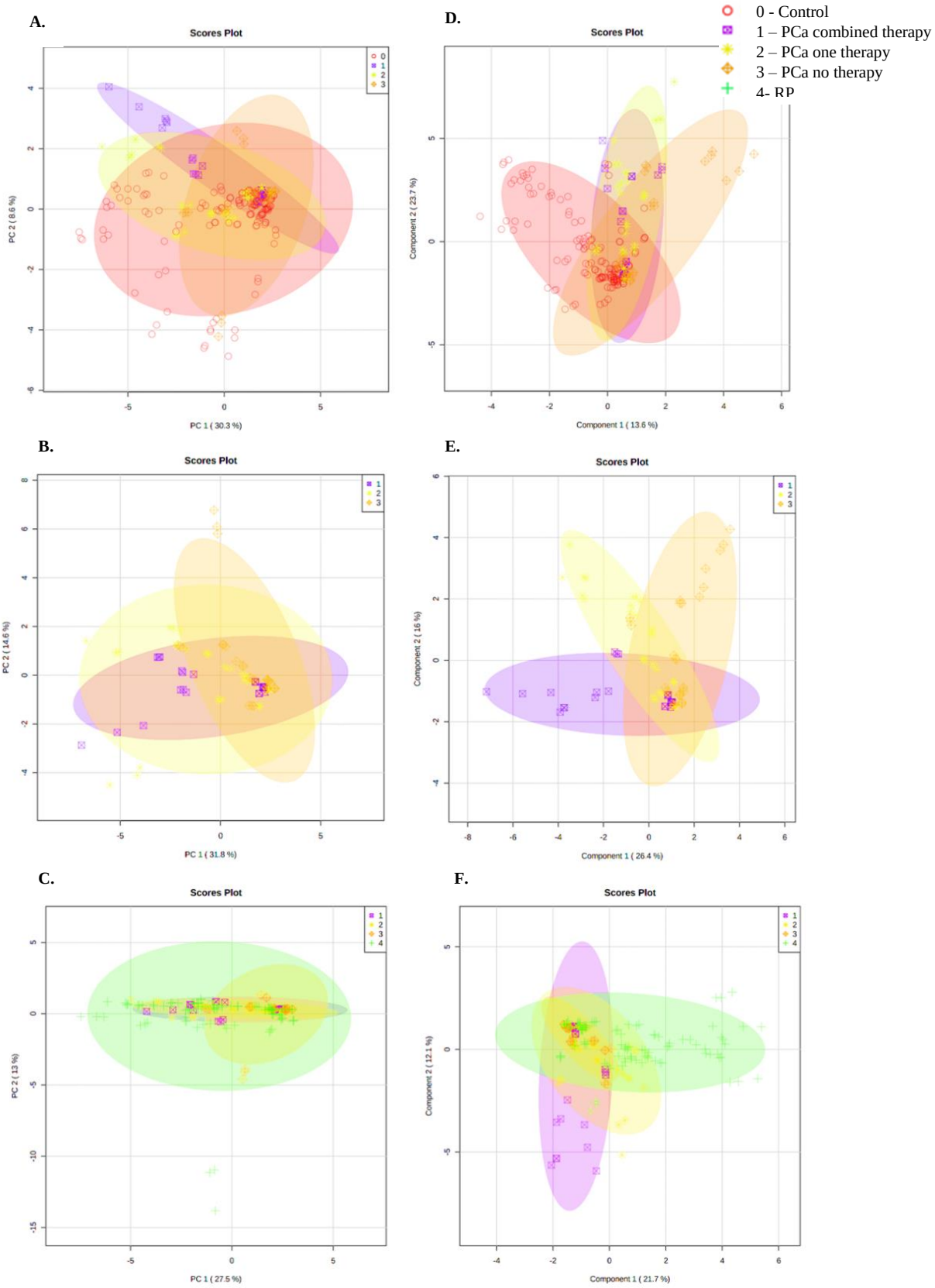


Figure S.3 – Score scatter plots of the volatile profile of urine samples from the groups under study. **A-C:** PCA PC1 x PC2 score scatter plot of **A.** control (0) *vs* PCa with combined therapy (1) *vs* PCa with one therapy (2) *vs* PCa with no therapy (3), **B.** PCa with combined therapy (1) *vs* PCa with one therapy (2) *vs* PCa with no therapy (3), and **C.** PCa with combined therapy (1) *vs* PCa with one therapy (2) *vs* PCa with no therapy (3) *vs* RP (4). **D-F:** Corresponding PLS-DA score scatter plots.

Table S.1 – Identification, chemical family, possible origin, and frequency of occurrence of the identified VOMs in the control group, PCa patients and RP subjects (n = 3; RSD < 20 %).

RT (min)	Metabolites	Formula	CAS nr	Possible origin	Frequency of occurrence (%)		
					CTRL	PCa	RP
Alcohols							
23.65	2-Hexanol[a]	$C_6H_{14}O$	26549-24-6	Endo/Exo (diet)	6	17	24
59.31	2,6-Dimethyl-7-octen-2-ol[a]	$C_{10}H_{20}O$	18479-58-8	Exo (diet)	49	28	38
61.84	2-Ethyl-1-hexanol[a,b]	$C_8H_{18}O$	104-76-7	Endo/Exo (diet)	37	62	59
80.89	2,6-Dimethyl-5,7-octadien-2-ol[a]	$C_{10}H_{18}O$	5986-38-9	Unk	0	0	3
Furanic compounds							
5.48	Furan[a]	C_4H_4O	110-00-9	Endo/Exo (diet, env)	59	52	56
6.51	3-Methylfuran[a]	C_5H_6O	930-27-8	Exo (diet, env)	18	0	3
8.60	2,5-Dimethylfuran[a]	C_6H_8O	625-86-5	Exo (diet, env)	6	0	3
11.78	2-Acetylfuran[a]	$C_6H_6O_2$	1192-62-7	Exo (diet)	0	3	0
11.89	2-Ethyl-5-methylfuran[a]	$C_7H_{10}O$	1703-52-2	End/Exo (diet)	0	14	26
29.02	2-Pentylfuran[a]	$C_9H_{14}O$	3777-69-3	End/Exo (diet)	0	0	3
30.19	2,2-dimethyl-5-(1 -methyl-1-propenyl)-tetrahydrofuran[a]	$C_{10}H_{18}O$	7416-35-5	Unk	43	45	35
85.15	4,7-Dimethylbenzofuran[a]	$C_{10}H_{10}O$	28715-26-6	Endo (syst)	6	3	0
107.68	2-(1-Cyclopentenyl)furan[a]	$C_9H_{10}O$	115754-78-4	Unk	20	0	9
Ketones							
9.31	2-Pentanone[a,b]	$C_5H_{10}O$	107-87-9	Exo (diet)	20	7	9
12.78	3-Hexanone[a]	$C_6H_{12}O$	589-38-8	Endo (syst)/ Exo (diet)	2	3	0
17.88	4-Heptanone[a,b]	$C_7H_{14}O$	123-19-3	Unk	98	97	100
23.28	Acetone[a]	C_3H_6O	67-64-1	Endo (syst, bact)	4	0	3
34.59	Cyclohexanone[a]	$C_6H_{10}O$	108-94-1	Endo (syst)/ Exo (diet)	0	3	0
Norisoprenoids							
60.49	Theaspirane[a]	$C_{13}H_{22}O$	36431-72-8	Exo (diet)	24	10	30
63.70	3,4,4a,5,6,7-Hexahydro-1,1,4a-trimethyl-2(1H)-naphthalenone[a]	$C_{13}H_{20}O$	4668-61-5	Unk	18	17	18
64.10	γ-Ionone[a]	$C_{13}H_{20}O$	79-76-5	Exo (diet)	2	17	9
89.38	1,1,6-Trimethyl-1,2-dihydronapththalene (TDN) [a]	$C_{13}H_{16}$	30364-38-6	Exo (diet)	76	93	88
98.08	β-Damascenone[a]	$C_{13}H_{18}O$	23726-93-4	Exo (diet)	33	10	30
101.60	α-Ionene[a]	$C_{13}H_{18}$	475-03-6	Exo (diet)	4	7	9
103.86	3,4-Dehydro-β-ionone[a]	$C_{13}H_{18}O$	1203-08-3	Unk	29	21	53

Table S.1 – *Continuation.*

RT (min)	Metabolites	Formula	CAS nr	Possible origin	Frequency of occurrence (%)		
					CTRL	PCa	RP
Sulphur-containing compounds							
4.66	Methanethiol[a]	CH_4S	74-93-1	Endo (syst, bact)	96	97	94
14.05	Dimethyl disulfide[a,b]	$C_2H_6S_2$	624-92-0	Endo (bact)/ Exo (diet)	100	97	97
45.81	Dimethyl trisulfide[a]	$C_2H_6S_3$	3658-80-8	Endo (bact)/ Exo	67	52	73
47.28	2-Methyl-5-(methylthio)furan[a]	C_6H_8OS	13678-59-6	Exo (diet)	29	48	79
Terpenes							
12.24	β-Citronellene[a]	$C_{10}H_{18}$	10281-56-8	Exo (diet)	45	69	50
27.12	α-Phellandrene[a]	$C_{10}H_{16}$	99-83-2	Endo/Exo (diet)	0	0	3
27.56	γ-Terpinene[a]	$C_{10}H_{16}$	99-85-4	Exo (diet)	4	7	0
33.06	*o*-Cymene[a]	$C_{10}H_{14}$	13877-91-3	Exo (diet)	88	90	82
35.47	Isoterpinolene[a]	$C_{10}H_{16}$	586-63-0	Endo (syst)/ Exo (diet)	8	0	6
54.39	*p*-Cymenene[a]	$C_{10}H_{12}$	1195-32-0	Endo/Exo (diet)	84	72	82
58.37	Linaloloxide[a]	$C_{10}H_{18}O_2$	5989-33-3	Exo (diet)	29	14	15
67.65	Pinocarvone[a]	$C_{10}H_{14}O$	30460-92-5	Exo (diet)	0	12	0
78.86	di-Menthol[a]	$C_{10}H_{20}O$	89-78-1	Exo (diet)	41	34	45
86.15	Phellandral[a]	$C_{10}H_{16}O$	21391-98-0	Exo (diet)	2	0	0
87.89	Carvone[a]	$C_{10}H_{14}O$	6485-40-1	Exo (diet)	2	10	12
102.05	α-Calacorene[a]	$C_{15}H_{20}$	21391-99-1	Endo/Exo (diet)	22	0	6
102.78	Dehydro-Ar-α-himachalene[a]	$C_{15}H_{20}$	78204-62-3	Exo (env)	2	0	0
105.01	α-Corocalene[a]	$C_{15}H_{20}$	20129-39-9	Endo/Exo (diet)	0	3	0
105.65	p-Cymen-7-ol[a]	$C_{10}H_{14}O$	536-60-7	Exo (diet)	2	0	0
107.94	1,6-Dimethyl-4-isopropylnaphthalene[a]	$C_{15}H_{18}$	483-78-3	Exo (diet)	29	3	26
Volatile phenols							
100.27	*o*-Methoxyphenol[a]	$C_7H_8O_2$	90-05-1	Unk	49	28	76
103.37	2-Bromophenol[a]	C_6H_5BrO	95-56-7	Endo/Exo (diet)	43	14	12
103.94	Phenol[a]	C_6H_6O	108-95-2	Endo/Exo (env)	100	97	97
105.28	*p*-Cresol[a]	C_7H_8O	106-44-5	Endo (bact)/ Exo	98	93	94
106.75	Eugenol[a]	$C_{10}H_{12}O_2$	97-53-0	Exo (diet, env)	0	0	3
108.83	m-tert-Butylphenol[a]	$C_{10}H_{14}O$	585-34-2	Endo/Exo (diet)	98	100	100

Table S.1 – *Continuation.*

RT (min)	Metabolites	Formula	CAS nr	Possible origin	Frequency of occurrence (%)		
					CTRL	PCa	RP
Others							
11.65	2,7-Dimethylocta-1,7-diene[a]	$C_{10}H_{18}$	59840-10-7	Unk	0	0	9
12.25	Toluene[a]	C_7H_8	108-88-3	Exo (env)	4	0	3
41.00	1,2,4-Trimethylbenzene[a]	C_9H_{12}	95-63-6	Endo/Exo (diet)	4	10	12
53.65	Acetic acid[a]	$C_2H_4O_2$	64-19-7	Endo (syst)/ Exo (env)	4	3	6
59.86	1,2,3,5-Tetramethylbenzene[a]	$C_{10}H_{14}$	527-53-7	Exo (env)	10	0	6
80.57	4-Methyl-4-vinylbutyrolactone[a]	$C_7H_{10}O_2$	1073-11-6	Exo (diet)	0	2	0
92.59	Methyl salicylate[a]	$C_8H_8O_3$	119-36-8	Endo/Exo (diet)	2	0	3
104.86	Octanoic acid[a,b]	$C_8H_{16}O_2$	124-07-2	Endo (bact)/ Exo	12	10	18

[a]Metabolite identification using mass spectra of the NIST library search.
[b]Metabolite identification using a chemical reference standard, with all analyses performed under identical analytical conditions.
Legend: Bact – bacterial; Diet – dietary; Endo – endogenous; Env – environmental; Exo - exogenous; Unk – unknown; Syst – systemic. CTRL – control group; PCa – prostate cancer; RP – radical prostatectomy; RT – retention time.

Table S.2 – PLS-DA CV details of the models generated in the three comparisons, control *vs* PCa, control *vs* RP, and PCa *vs* RP.

Measure	Control *vs* PCa (3 components)	Control *vs* RP (2 components)	PCa *vs* RP (2 components)
Accuracy	0.783	0.768	0.681
R^2	0.313	0.222	0.175
Q^2	0.289	0.206	0.157

Supplementary information

Appendice 1.

Folheto informativo para o participante do Estudo de Investigação Clínica/Translacional no SESARAM, EPERAM

Convidámos o Sr. (a) a participar em um estudo de Investigação Clínica/Translacional que está a ser realizado no SESARAM, EPERAM em parceria com o Laboratório de Bioquímica da Universidade da Madeira (UMa).

Nome do Estudo de Investigação:

«**ESTABELECIMENTO DE PADRÕES METABOLÓMICOS COMO PLATAFORMAS INOVADORAS E INTEGRADAS PARA O DIAGNÓSTICO DE CANCRO DA PRÓSTATA»**

Este estudo tem como objetivo detetar, na urina dos doentes com cancro da próstata, metabolitos voláteis endógenos (eVOMs) que poderão ser marcadores deste tipo de cancro, sendo assim úteis para o diagnóstico, monitorização da eficiência terapêutica nos doentes com este tipo de cancro.

Serão introduzidos neste estudo doentes com cancro da próstata, assim como participantes que não tenham essa patologia.

Para participar no estudo, os participantes necessitam de colher sangue (para exames analíticos como o PSA) e urina (para deteção dos metabolitos voláteis endógenos), que serão enviados para o Laboratório de Bioquímica da Universidade da Madeira de forma anónima e não identificável.

Os dados de saúde serão integrados numa «base de dados» de forma anónima e não identificável, tendo como fim o estudo de Investigação Clínica/Translacional acima referido. Esta «base de dados» sairá do SESARAM, EPERAM para o Laboratório de Bioquímica da Universidade da Madeira.

Informámos também que o participante neste estudo, pode retirar o seu consentimento e os seus dados serão de imediato extintos da «base de dados» do estudo.

Termo de Consentimento Informado para Estudo de Investigação Clínica/Translacional no SESARAM, EPERAM

(modelo 2)*

Nome do Estudo de Investigação:

«ESTABELECIMENTO DE PADRÕES METABOLÓMICOS COMO PLATAFORMAS INOVADORAS E INTEGRADAS PARA O DIAGNÓSTICO DE CANCRO DA PRÓSTATA»

Eu, …… CC nº…….

(nome do Participante no estudo)

Confirmo ter lido e compreendido toda a informação deste consentimento. Todas as minhas questões foram esclarecidas.

Concordo voluntariamente participar no estudo acima referido.

…… ……..

(data) (Assinatura do Participante)

Autorizo que os meus dados de Saúde, sejam tratados de forma anónima e não identificável pelos Investigadores do estudo acima referido e publicados em artigos científicos.

Concordo que os meus dados de Saúde, sejam integrados numa base de dados de forma não identificável, tendo como fim o estudo de investigação clínica/translacional acima referido e saiam do SESARAM, EPERAM.

Não autorizo que os meus dados de Saúde, saiam do SESARAM, EPERAM para serem integrados de forma anonimizada e tratados na plataforma digital do projeto FAIR4Health (base de dados de acesso aberto).

Concordo com a colheita de sangue e de urina para a realização de exames laboratoriais e autorizo saiam do SESARAM, EPERAM, de forma anónima e não identificável.

Sei que posso retirar o meu consentimento e que os meus dados de saúde serão extintos da base de dados do estudo.

Sei que receberei uma fotocópia deste formulário de consentimento informado.

…… ……

(data) (Assinatura do Participante)

…… ……

(data) (Assinatura do Profissional de Saúde que explicou o estudo)

…… ……..

(data) (Assinatura da Testemunha)

(modelo 2)* - Aplica-se aos Estudos de Investigação em que os dados de Saúde saem do SESARAM, EPERAM