NON-INTERVENTIONAL (NI) STUDY PROTOCOL

Study information

Title	Registry-based non-interventional analysis of advanced/metastatic renal carcinoma treatment patterns and outcomes in the Hospital district of Southwest Finland.
Protocol number	A4061098
Protocol version identifier	Version 1
Date	28 March 2022
Active substance	axitinib, ATC-code L01EK01 (previously L01XE17) sunitinib, ATC-code L01EX01 (previously L01XE04)
Medicinal product	Inlyta; Sutent
Research question and objectives	The aim of the study is to describe the characteristics, treatment pattern and outcomes of patients with advanced/metastatic renal cell carcinoma (RCC) with medical records in the HDSF data lake. The treatment sequence of mRCC patients and treatment outcomes (overall survival, time-to-next treatment) will be analysed with regard to individual risk score of the patient and over time. In addition, co-morbidities, procedures and healthcare usage will be analyzed. As an exploratory objective, the coding practices especially with regard to risk score calculation are assessed.
Author	PPD , PhD, PPD email: PPD tel. PPD

TABLE OF CONTENTS

1. TABLE OF CONTENTS....2
2. LIST OF ABBREVIATIONS....4
3. RESPONSIBLE PARTIES....5
4. AMENDMENTS AND UPDATES....6
5. MILESTONES....7
6. RATIONALE AND BACKGROUND....7
7. RESEARCH QUESTION AND OBJECTIVES....9
7.1. Primary objectives....9
7.2. Secondary objectives....9
7.3. Exploratory objective....10
8. RESEARCH METHODS....11
8.1. Study design....11
8.2. Setting....11
8.2.1. Inclusion criteria....12
8.2.2. Exclusion criteria....13
8.2.3. Patient stratification....13
8.3. Variables....14
8.4. Data sources....16
8.5. Study size....17
8.6. Data management....17
8.7. Data analysis....18
8.7.1. Patient characteristics....18
8.7.2. Treatment patterns and treatment lines....19
8.7.3. Treatment outcomes....20
8.7.4. Morbidity....21
8.7.5. Healthcare resource utilization....21
8.7.6. Changes in coding practices over time....22
8.8. Quality control....22
8.9. Limitations of the research methods....22
8.10. Other aspects....23

9. PROTECTION OF HUMAN SUBJECTS 23
9.1. Patient information 23
9.2. Patient consent 23
9.3. Institutional review board (IRB)/Independent ethics committee (IEC) 23
9.4. Ethical conduct of the study 23
10. MANAGEMENT AND REPORTING OF ADVERSE EVENTS/ADVERSE REACTIONS 23
11. PLANS FOR DISSEMINATING AND COMMUNICATING STUDY RESULTS.....24
12. REFERENCES 24
13. LIST OF TABLES 25
14. LIST OF FIGURES 25
ANNEX 1. LIST OF STAND ALONE DOCUMENTS 25
ANNEX 2. ENCEPP CHECKLIST FOR STUDY PROTOCOLS 25
ANNEX 3. ADDITIONAL INFORMATION.....25

1. LIST OF ABBREVIATIONS

Abbreviation	Definition
AJCC	American Joint Committee on Cancer
ATC	Anatomic therapeutic chemical classification system
BMI	Body mass index
HCRU	Healthcare resource utilization
HDSF	Hospital District of Southwest Finland
HTA	Health Technology Assessment
ICD-10	International Classification of Diseases 10th revision
ICMJE	International Committee of Medical Journal Editors
IMDC	International Metastatic RCC Database Consortium
LOT	Line of treatment
mRCC	(Advanced/)metastatic RCC
OS	Overall survival
PFS	Progression free survival
PO	Primary objective
RCC	Renal cell carcinoma
RWE	Real-world evidence
RWD	Real-world data
SAP	Statistical analysis plan
SO	Secundary objective
SoC	Standard-of-Care
TKI	Tyrosine kinase inhibitor
TNM	Tumor – Node – Metastasis (classification of malignant tumors)
TTNT	Time to next treatment
VEGFR	Vascular endothelial growth factor receptor

2. RESPONSIBLE PARTIES

Principal Investigator(s) of the Protocol

Name, degree(s)	Job Title	Affiliation	Address
PPD , PhD	PPD	PPD	PPD
PPD , PhD	PPD	PPD	PPD
PPD	PPD	PPD	PPD
PPD , PhD	PPD	PPD	PPD
PPD , DI	PPD	PPD	PPD
PPD , TkK	PPD	PPD	PPD
PPD , PhD	PPD	PPD	PPD

3. AMENDMENTS AND UPDATES

None.

4. MILESTONES

Milestone	Planned date
Auria permission and start of data collection	01-May-2022
End of data collection	31-May-2022
The following milestones are dependent on data set completion date. The planned dates have been estimated based on data completion by the end of May 2022:	
Data analyses	01 September 2022
Final study report	01-Nov-2022
Publication submission	To be determined

5. RATIONALE AND BACKGROUND

Renal cell carcinoma (RCC) is among the 10 most frequently occurring cancers and the most common type of kidney cancer occurring in Western communities [1]. RCCs comprise a heterogeneous group of malignant neoplasms arising from the nephron. Risk factors for RCC include obesity, hypertension, cigarette smoking, diet, diabetes, and male gender [1].

After decades of increasing trends in RCC incidence and mortality rates, it seems that rates are stabilising or starting to decline in many Western countries [1]. The worldwide number of kidney cancer cases was 431,288 (62.9% male) and deaths 179,368 (64.4% male) in 2020 [2]. Kidney cancer is more common in males compared to females. The incidence was 7.8 and mortality 3.0 per 100,000 males in high human development index countries in 2020 [2]. In Finland, the number of new kidney cancer cases was 635 in males and 353 in females in 2019 [3], and the number of prevalent patients was 4,943 males and 3,867 females in 2019 according to the Finnish Cancer Registry [4].

Localized RCC can be curatively treated with nephrectomy. However, about 20–30% of patients present with metastatic RCC (mRCC) at the time of diagnosis [5]. In addition, another 20% of patients with localized RCC undergoing nephrectomy will have a relapse and develop mRCC during follow-up. Despite the recent advances in mRCC treatment, metastatic disease is often fatal [6].

Over the last decade, the treatment of advanced/metastatic RCC has been constantly evolving, especially in the last 2-3 years with the emergence of several novel treatment options in rapid succession. Current recommendations for first-line therapy of mRCC include several combination treatments, superseding the previously recommended antiangiogenic agent monotherapies. The novel options combine an anti-PD-1-antibody with either an antiangiogenic drug (vascular endothelial growth factor receptor (VEGFR)) tyrosine kinase inhibitor (TKI)), or an anti-CTLA-4-antibody. The combination treatments are recommended

variably according to the patient's IMDC (International mRCC Database Consortium) risk score. IMDC represents a large collection of real-world data (RWD) on patients with advanced kidney cancer treated with targeted therapies [7,8]. IMDC combinations available for favourable risk patients include pembrolizumab + axitinib, nivolumab + cabozantinib and pembrolizumab + lenvatinib; for intermediate or poor risk patients the combination of ipilimumab and nivolumab is recommended in addition to the aforementioned three combinations. None of the recommended first-line treatment options have been compared head-to-head but all of the pivotal clinical studies have sunitinib monotherapy as a control, a former standard-of-care (SoC). The current second and later line treatment recommendations include VEGFR-TKI monotherapies, such as pazopanib and sunitinib, but will likely be reshaped in the coming years as clinical evidence on the optimal treatment sequence is accumulated.

The prognostic model constructed by the IMDC has been used to stratify patients in contemporary clinical trials and to provide risk-directed treatment selection in everyday clinical practice. The IMDC prognostic model contains six independent predictors of poor survival, which are: Karnofsky performance status of less than 80%, less than 1 year from diagnosis to treatment, anaemia (haemoglobin concentration <lower limit of normal), hypercalcaemia (corrected calcium concentration >upper limit of normal), neutrophilia (neutrophil count >upper limit of normal), and thrombocytosis (platelet count >upper limit of normal). An IMDC risk score is determined for a patient according to the number of poor prognostic factors the patient has; favourable risk patients have 0 factors; intermediate risk patients have one or two factors; and poor risk patients have three or more factors.

In Finland, the mRCC treatment landscape has been shaped by the access and reimbursement status of medications. Here, publicly funded medicines are processed via two different access and reimbursement processes based on the route of administration: outpatient or oral medicines (patient acquires medicine with a prescription from the pharmacy) and hospital medicines (medicines administered at hospital, mainly intravenous medicines).For oral medicines, the Finnish Social Insurance Institution (Kela) reimburses the patient for the costs of outpatient medicines if the Pharmaceuticals Pricing Board (Hila) has confirmed the reimbursement of the medicine. For oral drugs, the marketing authorization holder is responsible for applying for reimbursement. The Finnish Medicines Agency (Fimea) is the national competent authority for regulating pharmaceuticals. Fimea is responsible for initiating the evaluation of hospital medicines to be used in Finland, and makes the initial health technology assessment (HTA). This HTA-evaluation is not made automatically for all new pharmacotherapies or indication extensions, but at the discretion of Fimea. Hospital medicines assessed by Fimea are further assessed by the Council for Choices in Health Care in Finland (COHERE Finland) to issue recommendations on services that should be included in the selection of public health services. Hospital medicines are paid by the municipalities that fund the hospital districts. The use of hospital medicines and prices may also vary greatly from region to region. Use of combination treatments is often problematic due to the Finnish reimbursement system because the combination partners may fall under different processes (eg. a combination of an oral drug and an intravenously administered drug).

Finland has equal healthcare which is primarily funded by taxation. Thus, all permanent residents in Finland are entitled to public healthcare at a uniform level regardless of their financial situation. Finnish registries cover all individuals living in Finland and no patients are excluded from our study based on their social status or financial capability. In addition, health record data available via data lake technology enable extraction and analysis of large data sets including RWD on disease-related clinical and molecular characteristics and a number of medical procedures. Data for this study will be retrieved from the data lake of the Hospital District of Southwest Finland (HDSF). With a total of 470,000 residents and about 200,000 patients each year, HDSF is one of the biggest hospital districts in Finland.

The aim of this retrospective study is to investigate treatment patterns and outcomes of patients with advanced/metastatic RCC treated at the HDSF. In this protocol, mRCC refers to advanced and/or metastatic RCC. In addition, coding practices of related data and characteristics of all RCC patients will be reported. To our knowledge, this will be the first study to utilize regional Finnish data to characterize RCC patients, and investigate the treatment of the mRCC patient population. The resulting real-world evidence (RWE) is likely to be of interest for the national and global RCC scientific communities.

6. RESEARCH QUESTION AND OBJECTIVES

The objective of this study is to describe the characteristics and treatment outcomes of patients with advanced/metastatic RCC with medical records in the HDSF data lake. In the first primary objective, all RCC patients (with local or metastatic disease) are analyzed (treated and non-treated mRCC patients separately). In all other objectives, only treated mRCC patients will be analyzed. The objectives of the study are to:

6.1. Primary objectives

1. Characterize patients with RCC and with advanced/metastatic RCC (treated and non-treated separately) including e.g. demographics and IMDC risk score.

2. Define treatment patterns and treatment sequences in mRCC.

3. Assess outcomes by treatment line and from diagnosis
    a. Progression free survival (via proxies such as time-to-next treatment)
    b. Overall survival
        i. time to event analyses
        ii. Cox models

6.2. Secondary objectives

1. Assess morbidity of metastatic RCC including
    a. co-diagnoses and procedures
    b. infections and antibiotic use
    c. antihypertensive use before and after mRCC diagnosis, as applicable

2. Assess speciality care healthcare resource use (absolute, per patient, per patient year)
    a. cumulatively over time from diagnosis
    b. by treatment line

6.3. Exploratory objective

1. Assess changes in coding practices over time, specifically IMDC risk score
    a. including but not limited to changes in structured IMDC score recordings and the use of metastatic ICD-10 diagnosis codes (C77-C79)

7. RESEARCH METHODS

7.1. Study design

This is a retrospective registry-based cohort investigation, utilizing existing data generated during routine clinical practice and available in the Hospital District of Southwest Finland HDSF data lake. All adult patients (age ≥18 years) with RCC diagnosis (C64*; '*' indicates any number) will be identified from the data lake of HDSF.

The main purpose of this study is to characterize the patient population of advanced/metastatic RCC, and their treatment patterns and treatment outcomes using medical records retrieved from HDSF data lake. The main patient stratification will be the IMDC risk category (see 8.2.3 Patient stratification). Patients with unknown risk score will be analyzed as an own patient group.

7.2. Setting

The target population of this study is adult patients with RCC and medical records available at HDSF data lake. Estimated number of patients in this study is 220-550. The inclusion criteria is a diagnosis of renal cell carcinoma (ICD-10: C64) between January 2010 and December 2021 (Figure 1). Advanced/metastatic disease stage (mRCC) will be assessed based on ICD-10 diagnosis code for metastasis (C77*-C79*), American Joint Committee on Cancer (AJCC) stage 4, visit to oncologist (specialty code 65), or initiation of treatment for mRCC.

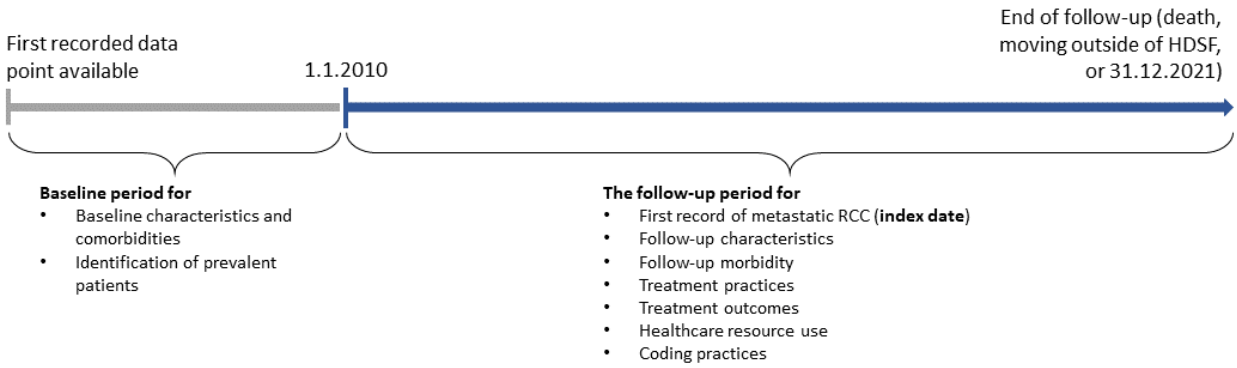

Figure 1. Graphical illustration of study time windows and corresponding analyses.

The data collected has been generated as part of standard clinical care, treatment, and follow-up of patients with RCC. In accordance with the Finnish Act on Secondary Use of Health and Social Data 552/2019 (in Finnish: Laki sosiaali- ja terveystietojen toissijaisesta käytöstä), health registry data can be processed in scientific studies without patient consent. The purpose of this act is to establish conditions for the effective and secure processing of, and access to, personal health and social data for certain secondary purposes, such as research and statistics, innovation and development, knowledge management, teaching and authority planning. Thus, the Secondary Act creates a clear legal basis for the use of such registered data for research and innovation related to, for example, the health and well-being of citizens, the prevention of

disease and the development of new treatment methods. This relatively new legislation is also a welcome unification of the fragmented Finnish national rules regarding the use of healthcare and social welfare data. In addition, it takes into consideration current data protection requirements.

In the primary objective 1 (characteristics), all RCC patients will be analysed (local and metastatic diseases separately), while in the other objectives only treated mRCC patients will be analysed. Therefore, there are two cohorts to be analysed: 1) characteristics cohort (including all RCC patients), 2) mRCC cohort (including only treated mRCC patients).

Definitions for mRCC cohort

Index: Date of the first mRCC diagnosis; date of the metastatic disease identified as first of the following after the initial RCC diagnosis

- ICD-10 diagnosis code (C77*-C79*) for metastatic disease
- AJCC stage 4
- Visit to oncologist (specialty code 65) with RCC as main diagnosis
- Treatment initiation for mRCC (sunitinib, pazopanib, cabozantinib, ipilimumab, nivolumab)

When analyzing the treatment outcomes, index is defined as the date of the treatment initiation.

End of follow-up (whichever of the following occurs first):

- Death
- Moving outside of HDSF
    - Approximated using data of healthcare contacts; no contacts to specialty care of HDSF within 2 years (patient censored at the last recorded visit)
- End of study (31.12.2021)

Length of follow-up (in years): (‘end of follow-up’ – ‘index’ + 1)/365.25

Definitions for characteristics cohort

Index:

- For mRCC patients, same as for the mRCC cohort (i.e. approximated metastasis date)
- For RCC patients, the date of the first record of RCC diagnosis (ICD-10: C64)

7.2.1. Inclusion criteria

For the **characteristics cohort**, patients must meet all of the following inclusion criteria to be eligible for inclusion:

- Diagnosis of renal cell carcinoma (ICD-10: C64) during 1.1.2010-31.12.2021
- Age at least 18 years at index.
- Resident of HDSF at index.

For the **mRCC cohort**, patients must meet the following inclusion criteria in addition to the above criteria to be eligible for inclusion:

- metastatic RCC during 1.1.2010-31.12.2021; defined as
    - ICD-10 diagnosis code for metastasis (C77*-C79*), or
    - AJCC stage 4, or
        - AJCC stage data potentially not available for all patients
    - a visit to oncologist (specialty code 65) with RCC as main diagnosis, or
        - In Finland, the ICD-10 codes for metastatic disease are rarely used. In contrast, RCC patients visit oncologist, when and only when disease metastasises and therefore, the visit can be used as a proxy to a metastatic disease.
    - initiation of treatment for mRCC

7.2.2. Exclusion criteria

For the **characteristics cohort**, patients meeting any of the following criteria will be excluded from the study:

- Prevalent mRCC patients (i.e. diagnosis of metastatic RCC before 1.1.2010)

- Prevalent RCC patients (i.e. diagnosis of RCC before 1.1.2010) if there is no records of metastatic disease during 2010-2021

For the **mRCC cohort**, patients meeting any of the following criteria will be excluded from the study:

- Prevalent patients with mRCC (i.e. diagnosis of metastatic RCC before 1.1.2010)
    - Note that patients with local RCC before 2010 are included in this study if the disease has metastasised during 2010-2021
- Patients without treatment for mRCC (see Table 3 for details)

7.2.3. Patient stratification

Analyses will be stratified by:

- IMDC risk category (defined at the initiation of the first treatment for the metastatic RCC; see Table 1 and listing below for definitions)
    - Applied to all primary and secondary objectives
- Treatment line
    - Applied to primary objective 2 and 3

IMDC risk factors (each criterion accounts for +1 for the risk score):

1. Less than a year from time of initial RCC diagnosis to initiation of treatment for the metastatic RCC

2. Performance status* <80% (Karnofsky)
3. Hemoglobin* below normal (women < 117 g/l; men < 134 g/l)
4. Serum calcium* above normal (> 2.51 mmol/l)
5. Neutrophiles* above normal (> 6.3 x 10^9/l)
6. Thrombocytes* above normal (> 350 x 10^9/l)

** for performance status and laboratory values data of ± 3 months around the treatment initiation will be utilized due to possibly sprase data. If multiple records are available, the record closest to the treatment initiation will be utilized.*

Table 1. IMDC risk scores and categories based on two different approaches depending on the data coverage

	IMDC risk category	**Total score of IMDC risk factors**	**Explanations/missing values**
Option 1*	Favorable risk	0	no missing values allowed
	Intermediate risk	1-2	1) total score = 1; max 1 risk factor allowed to have missing value 2) total score = 2; no missing values allowed
	Poor risk	3 or more	max 3 risk factors allowed to have missing values
	Unknown risk**	-	when none of the above risk categories can be assigned
Option 2*	Appear favorable	0	all non-missing values for the risk factors indicate favorable risk max 3 risk factors allowed to have missing values
	At least intermediate	≥1	at least one risk factor criterion is met max 3 risk factors allowed to have missing values
	Poor	≥3	at least three risk factor criteria is met max 3 risk factors allowed to have missing values
	Unknown**	-	when none of the above risk categories can be assigned, i.e. more than 3 risk factors have missing value

** If small number of missing values is observed, the option 1 will be selected. If large number of missing values is observed and thus majority of patients were to belong to the "unknown" group of the option 1, the option 2 will be selected. The selection will be done after the data release and monitoring of data coverage.*
*** Patients with unknown risk category, will be analysed as own patient group.*

7.3. Variables

Source for all variables is the data lake of HDSF.

Table 2. List of study variables. Abbreviations: PO=primary objective, SO=secondary objective.

Variable	Role/Relevant objective	Operational definition
ICD-10 diagnoses; RCC diagnoses	Patient inclusion/exclusion IMDC risk category	yes/no
ICD-10 diagnoses: co-diagnoses (e.g. renal insufficiency and hypertension)	PO1, SO1	yes/no; codiagnoses before and after index
Date of birth	Patient inclusion PO1	Date; age defined at index based on birth date
Date of death	End of follow-up PO3	Date
Gender	PO1	Female/male
Home municipality	Patient inclusion End of follow-up	Home municipality at each healtcare contact
Weight	PO1	kg
Height	PO1	m
Body mass index (BMI)	PO1	derived from weight and height
Smoking status (if available in structured format)	PO1	yes/no
Outpatient visits (including emergency room visits)	Patient inclusion and index definition (visits to oncologist) SO2	Date, physician's specialty, type of arrival
Hospitalizations	SO2	Start and end dates
Procedures and operations (including nephrectomy)	PO1 SO1 SO2	Procedure code, date
Medical imaging	SO1 SO2	Imaging code, date
Medications; including - hospital medications - prescriptions	Patient inclusion/exclusion and index definition PO2 PO3 SO1 SO2	Dates, ATC-codes, active substance
Laboratory testing	IMDC risk category PO1	Dates, test id, result
Zubrod/ECOG performance status	IMDC risk category PO1	scale 0-5; converted to Karnofsky
PAD histology (if available in structured format)	PO1	
Staging of RCC, the tumor-node-metastasis (TNM) classification	Patient inclusion and index definition PO1	scale 1-4 (stages)

Table 3. List of medications for RCC treatment. ATC= Anatomical Therapeutic Chemical.

	Active substance	Brand name	ATC-code at the moment in use	ATC-code previously in use
Oral medications	axitinib	Inlyta	L01EK01	L01XE17
	everolimus	Afinitor	L01EG02	L01XE10
	cabozantinib	Cabometyx	L01EX07	L01XE26
	pazopanib	Votrient	L01EX03	L01XE11
	sorafenib	Nexavar	L01EX02	L01XE05
	sunitinib	Sutent	L01EX01	L01XE04
Intravenous medications	ipilimumab	Yervoy	L01XC11	
	nivolumab	Opdivo	L01XC17	

7.4. Data sources

Cohort and data set formation overview is presented in Figure 2. The patient population will be selected from the health records at the data lake of HDSF based on diagnosis data and includes a personal identity code that is unique for every individual in Finland. For the identified RCC cohort, data of patient characteristics, diagnoses, healthcare contacts, treatments, laboratory and pathology tests, and performed procedures, operations and imagings will be collected from the same data lake of HDSF. The identification of the RCC cohort and collection of the additional data with personal identification numbers will be handled solely by the personnel of Auria Clinical Informatics. The identification codes are removed after linkage so that only pseudonymized data or group specific summary data is provided for analyses.

This study is conducted using existing and available clinical electronic records accessible from Auria Clinical Informatics. The study protocol and associated data request for all data will therefore be submitted to Auria Clinical Informatics for approval.

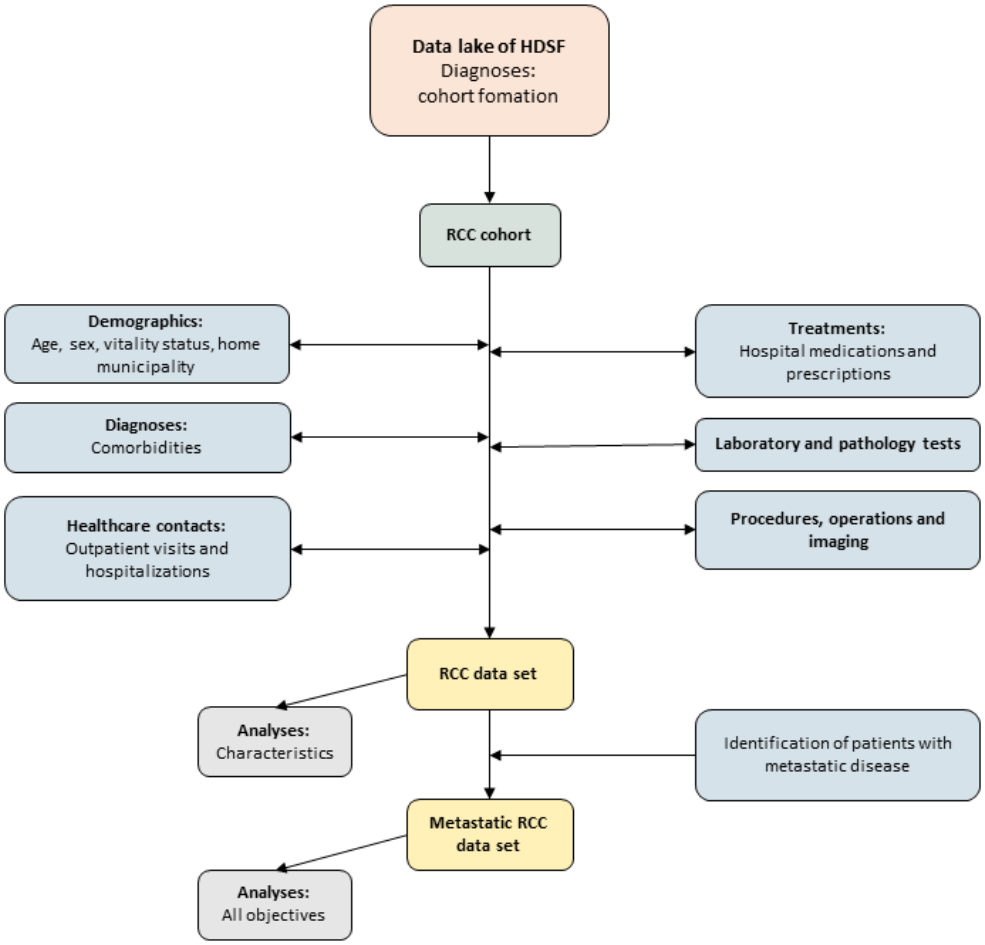

Figure 2. Overview of cohorts and data set formation.

7.5. Study size

The expected number of patients eligible for this study is approximated based on publicly available data from the Finnish Cancer registry (https://syoparekisteri.fi/tilastot/tautitilastot/):

- In HDSF, there have been approximately 1,100 RCC patients during 2010-2021.
- 20-50% of RCC patients develop advanced/metastatic disease [9] and therefore, the size of mRCC-patient cohort is approximately 220-550 patients.

7.6. Data management

All data collection, storage and handling will be coordinated by Auria Clinical Informatics. Auria Clinical Informatics collects the required data, pseudonymizes the IDs and releases the row level pseudonymized data to a secure analysis environment maintained by Auria Clinical

Informatics. All data-analyses will be performed by the Medaffcon Oy analytics team in the secure analysis environment using the statistical software R (version 3.6.1. or higher). No patient level data will be transferred outside of this system. Only aggregate level and/or fully anonymous data and/or summary statistics may be transferred outside this system for reporting and publication purposes. Thus, individual patients cannot be identified from this data set.

7.7. Data analysis

Data will be analyzed using primarily descriptive measures, which include mean/median values for continuous variables and proportions for categorical variables. Costs will be estimated using data of visits, hospitalizations, procedures, etc. and utilizing publicly available price listings. For treatment outcomes, Kaplan-Meier estimates and/or corresponding time to event competing risk models will be utilized.

All analyses will be reported for the whole advanced/metastatic RCC patient population and stratified by the IMDC risk score. Additionally, treatment patterns and treatment outcomes will be reported by treatment lines.

Analyses are based on structured data, with the exception for potential text-mining of IMDC risk score or its components.

Due to the retrospective nature of the study, missing and/or incomplete data is expected in some of the medical records. Patients' records will not be excluded because of missing values, and missing data will not be imputed, i.e. the data will be analyzed as they are recorded in the electronic medical records. The proportion of missing values per variable will be reported.

Analyses will be performed using R, a language and environment for statistical computing, in Rstudio-server environment.

7.7.1. Patient characteristics

Primary objective 1:

Demographical and clinical variables of RCC and mRCC patients will be summarized at index. Treated and non-treated mRCC patients will be analysed separately. For laboratory measures and performance status, the closest available measure +/- 90 days from index will be utilized. In case of a high rate of missing values, the time window can be extended to 6 months.

The demographical and clinical variables include but are not limited to:

- age at index
- sex
- history of renal insufficiency (any records until index)
    - ICD-10 code: N17-N19
- history of dialysis (any records until index)
    - procedure codes: TK800, TK829, TK810,
    - ICD-10 code: Z49

- history of kidney transplant (any records until index)
    - procedure codes: KAS10, KAS20
    - ICD-10 code: Z94.0
- history of nephroctomy (any records until index)
    - procedure codes: KAC*, KAD*, KAS40, KAS41
    - ICD-10 code: Z90.5
- performance status at index
    - Karnofsky (converted from Zubrod/ECOG)
- time from initial diagnosis of RCC to initiation of treatment for metastatic RCC
- various laboratory measures (e.g. hemoglobin, serum calcium, neutrophiles, thrombocytes, creatinine, albumin, eGFR)
- smoking status (if available in structured format)
- histology (if available in structured format)
- BMI (if available in structured format)

For categorical variables, the number (N) and the proportion (%) of patients in each class will be reported. For continuous normally distributed variables, mean and standard deviation (SD) will be reported. For continuous, non-normally distributed variables, median, the 1st quartile and the 3rd quartile will be reported.

The difference between strata will be tested using chi-squared/Fisher's exact test (categorical variables), t-test (continuous normally distributed variables) and Kruskal-Wallis test (continuous non-normally distributed variables).

7.7.2. Treatment patterns and treatment lines

Primary objective 2:

The treatment lines will be defined using drug administration and prescription data. The treatment options and lines will be visualized using Sankey plots and the proportion of patients per treatment type will be reported.

In general in mRCC, line of treatment (LOT) changes when the active substance changes, as the treatments are mainly monotherapies and only a few options consist of two substances (e.g. combination of ipilimumab and nivolumab in 1L).

The treatment lines will be defined *post hoc* using following process:

- For each active substance, single administrations/prescriptions are merged to treatment continuums.

- To identify, whether the treatment is monotherapy or combination of two active substances, the first records of the active substances will be considered

- if the first records of the two active substances are within 28 days, the treatment is considered as combination therapy of these two medications
- otherwise the two active substances are considered as two separate treatment lines

- The initiation of the treatment line is defined as the date of first record (dosing or prescription) of the active substance.

7.7.3. Treatment outcomes

Primary objective 3:

The treatment outcomes will be assessed using time-to-event analysis, namely Kaplan-Meier fit or other relevant competing risk models. For each outcome, the time will be defined as follows (Figure 3):

- overall survival:
  - time from index until death (event) or end of study (censoring event)
    - when analysing by treatment lines, start the time from initiation of each treatment line
- progression free survival (PFS, defined using time to next treatment (TTNT) proxy):
  - time from initiation of the current treatment line until the initiation of the next treatment line (event), death (competing risk/event), or end of study (censoring event)

The Kaplan-Meier fits will be visualized, and the median survival/PFS of the patients will be reported as well, if reached.

Additionally, corresponding Cox proportional-hazards model will be fitted, including clinical variables as covariates (including but not limited to age, sex, IMDC risk score, diagnosis year, and treatment line number). The hazard ratios, 95% confidence intervals and p-values corresponding to each covariate will be reported.

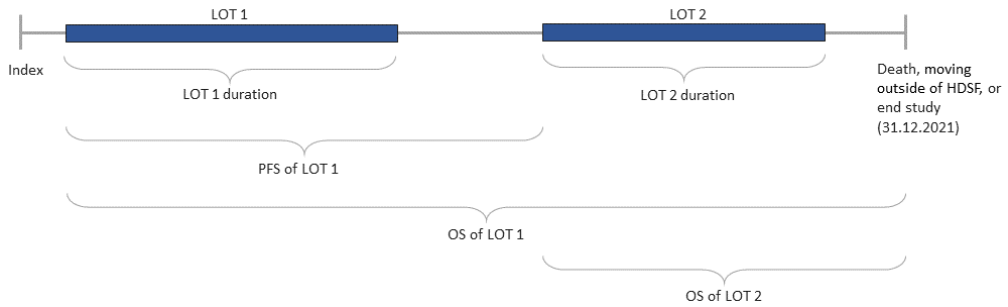

Figure 3. Graphical illustration of definitions of treatment outcomes. Abbreviations: LOT = line of treatment.

7.7.4. Morbidity

Secondary objective 1:

Morbidity will be assessed using data of co-diagnoses, procedures, and medications before and after diagnosis of advanced/metastatic RCC using all available data (ICD-10, procedure codes, and ATC-codes) in baseline and during the follow-up, respectively. The number (N) and proportion (%) of patients with each morbidity will be reported.

The list of morbidities of interest are:

- co-diagnoses using ICD-10 codes
    - especially interest of infections (ICD-10: A00*-B99*) and hypertension (ICD-10: I10)
- procedures related to kidneys
    - all KA* (kidneys and pelvis of kidneys) and KA_* (diagnostic radiology of kidney and pelvis of kidneys) procedures codes
- use of antibiotics (ATC-code: J01*, J02* and J04*) and antihypertensive medication (ATC-code: C02*)

If data coverage seems low, use of antibiotics and antihypertensives can be used as a proxy for infections and hypertension, respectively, along with the diagnosis (ICD-10) data. Therefore, diagnoses and medications will be analyzed as a combination instead of separately.

7.7.5. Healthcare resource utilization

Secondary objective 2:

Healthcare resource utilization (HCRU) will be defined as number of outpatient clinic visits, ER visits, hospitalizations, hospital inpatient days, performed procedures, operations, imaging, and laboratory tests. The price will be evaluated for each entry using publicly available price listings.

The absolute number of each HCRU type and the absolute costs associated will be reported. Additionally, the estimates will be scaled to “per-patient” estimates (by dividing the absolute estimates with the number of contributing patients) and to “per patient year” estimates (by dividing the absolute estimates by the number of contributing patient years).

Additionally, assess the HCRU cumulatively over time, i.e. annual costs per patient during the first 5 years from index, each year separately (i.e., the costs during the first year after the diagnosis of advanced/metastatic RCC, during the second year, etc).

No statistical testing for difference between groups will be performed and estimates will be descriptive in nature. However 95% confidence intervals (CI) for the estimates will be derived using bootstrapping over patients.

7.7.6. Changes in coding practices over time

Exploratory objective 1:

Assess changes in coding practices as a proportion of patients with data available respect to the diagnosis year of advanced/metastatic RCC. Variables of interest include but are not limited to

- structured IMDC risk score
- ICD-10 diagnosis for metastatic disease (C77*-C79*)

Other aspects of changes in coding practices may also be analyzed, if reasonable.

7.8. Quality control

Internal quality will be assured by consulting a clinical expert on data integrity, clinical relevance, and plausibility of the results.

A quality control will be performed, whereby all data will be sanity checked by the allocated data scientist(s). This control includes e.g. data coverages, number of individuals and data rows, changes as a function of time and checks for systematic gaps on data coverage and outliers in the data. Additionally, all the results will be sanity checked with a clinical expert and study team for plausibility and clinical relevance, and especially in case of unexpected results, both analytical methods and data will be discussed and validated with the whole Medaffcon analytics team.

All R scripts to process and analyze the data will be saved, and there is version control and external back-ups for the scripts. At the end of the study, all scripts will be archived to assure analysis reproducibility, and plausible later audits (by client and/or from scientific publication side).

7.9. Limitations of the research methods

All the data is recorded by the hospital during everyday practice and most of the data is stored in structured format. Therefore, rather high quality of the data is expected. However, as with all RWD, it is plausible to have erroneous entries. Also due to the retrospective nature of the study, missing and/or incomplete data is expected in some of the medical records. Patients' records will not be excluded because of missing values, and missing data will not be imputed, i.e. the data will be analyzed as they are recorded in the electronic medical records. The proportion of missing values per variable will be reported.

The treatment lines will be defined based on the medication data and therefore, misclassification of treatment lines is possible. However, the start and end dates of each main drug are expected to be recorded rather precisely, and the clinician PPD will be consulted when constructing the treatment lines. Thus, the majority of the treatment lines are expected to be defined correctly.

7.10. Other aspects

Not applicable.

8. PROTECTION OF HUMAN SUBJECTS

8.1. Patient information

This study involves data that exist in anonymized structural or non-structural format and contain no patient personal information.

8.2. Patient consent

As this study involves anonymized structural data, which according to applicable legal requirements do not contain data subject to privacy laws, obtaining informed consent from patients by Pfizer is not required.

8.3. Institutional review board (IRB)/Independent ethics committee (IEC)

Per the Act on Secondary Use of Health and Social Data 552/2019 (in Finnish: Laki sosiaali- ja terveystietojen toissijaisesta käytöstä), no institutional review board or independent ethics committee for this retrospective registry study is required.

8.4. Ethical conduct of the study

Per the Act on Secondary Use of Health and Social Data 552/2019 (in Finnish: Laki sosiaali- ja terveystietojen toissijaisesta käytöstä), no ethical approval for this retrospective registry study is required. This study protocol will be evaluated and approved by the Auria Clinical Informatics before any activities.

The study will be conducted in accordance with legal and regulatory requirements, as well as with scientific purpose, value and rigor and will follow generally accepted research practices described in the Guidelines for Good Pharmacoepidemiology Practices (GPP) [10].

9. MANAGEMENT AND REPORTING OF ADVERSE EVENTS/ADVERSE REACTIONS

This study involves data that exist as structured data by the time of study start or a combination of existing structured data and unstructured data, which will be converted to structured form during the implementation of the protocol solely by a computer using automated/algorithmic methods, such as natural language processing.

In these data sources, individual patient data are not retrieved or validated, and it is not possible to link (ie, identify a potential association between) a particular product and medical event for any individual. Thus, the minimum criteria for reporting an adverse event (AE) (i.e., identifiable patient, identifiable reporter, a suspect product, and event) cannot be met.

10. PLANS FOR DISSEMINATING AND COMMUNICATING STUDY RESULTS

The detailed results obtained from this study will be available upon publication. Research results from this study will be published in peer reviewed scientific journals and/or international scientific congresses. The Authorship of any publications resulting from this study will be determined based on the International Committee of Medical Journal Editors (ICMJE) Recommendations for the Conduct, Reporting, Editing and Publication of Scholarly Work in Medical Journals.

In the event of any prohibition or restriction imposed (e.g., clinical hold) by an applicable competent authority in any area of the world, or if the party responsible for collecting data from the participant is aware of any new information which might influence the evaluation of the benefits and risks of a Pfizer product, Pfizer should be informed immediately.

11. REFERENCES

[1] Ljungberg B, Campbell SC, Cho HY, Jacqmin D, Lee JE, Weikert S, et al. The Epidemiology of Renal Cell Carcinoma. European Urology 2011;60:615–21. https://doi.org/10.1016/j.eururo.2011.06.049.
[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 A Cancer J Clin 2021;71:209–49. https://doi.org/10.3322/caac.21660.
[3] Pitkäniemi J, Malila N, Tanskanen T, Degerlund H, Heikkinen S, Seppä K. Syöpä 2019. Tilastoraportti Suomen syöpätilanteesta. Helsinki: Suomen Syöpäyhdistys; 2021.
[4] Suomen Syöpärekisteri. Suomen Syöpärekisteri datan. Https://TilastotSyoparekisteriFi/Syovat 2021.
[5] Gupta K, Miller JD, Li JZ, Russell MW, Charbonneau C. Epidemiologic and socioeconomic burden of metastatic renal cell carcinoma (mRCC): A literature review. Cancer Treatment Reviews 2008;34:193–205. https://doi.org/10.1016/j.ctrv.2007.12.001.
[6] Capitanio U, Montorsi F. Renal cancer. The Lancet 2016;387:894–906. https://doi.org/10.1016/S0140-6736(15)00046-X.
[7] Heng DYC, Xie W, Regan MM, Warren MA, Golshayan AR, Sahi C, et al. Prognostic Factors for Overall Survival in Patients With Metastatic Renal Cell Carcinoma Treated With Vascular Endothelial Growth Factor–Targeted Agents: Results From a Large, Multicenter Study. JCO 2009;27:5794–9. https://doi.org/10.1200/JCO.2008.21.4809.
[8] Heng DY, Xie W, Bjarnason GA, Vaishampayan UN, Donskov F, Wood L, et al. A unified prognostic model for first- and second-line targeted therapy in metastatic renal cell carcinoma (mRCC): Results from a large international study. JCO 2010;28:4523–4523. https://doi.org/10.1200/jco.2010.28.15_suppl.4523.

[9] Flanigan RC, Campbell SC, Clark JI, Picken MM. Metastatic renal cell carcinoma. Curr Treat Options in Oncol 2003;4:385–90. https://doi.org/10.1007/s11864-003-0039-2.
[10] International society for Pharmacoepidemiology. Guidelines for good pharmacoepidemiology practice (GPP). Pharmacoepidemiology and Drug Safety 2016;25:2–10. https://doi.org/10.1002/pds.3891.

12. LIST OF TABLES

Table 1. IMDC risk scores and categories based on two different approaches depending on the data coverage 14
Table 2. List of study variables. Abbreviations: PO=primary objective, SO=secondary objective....15
Table 3. List of medications for RCC treatment. ATC= Anatomical Therapeutic Chemical. 16

13. LIST OF FIGURES

Figure 1. Graphical illustration of study time windows and corresponding analyses. 11
Figure 2. Overview of cohorts and data set formation....17
Figure 3. Graphical illustration of definitions of treatment outcomes. Abbreviations: LOT = line of treatment. 20

ANNEX 1. LIST OF STAND ALONE DOCUMENTS

None.

ANNEX 2. ENCEPP CHECKLIST FOR STUDY PROTOCOLS

Not required.

ANNEX 3. ADDITIONAL INFORMATION

Not applicable.