Research Article

Clinical Characteristics and Management of Patients with Secondary Hyperparathyroidism Undergoing Hemodialysis: A Feasibility Analysis of Electronic Health Records Using Natural Language Processing

Alberto Ortiz[a, b]   Jose Portoles[b, c]   Maria Dolores Pino-Pino[d]   Jesús Barea[e]

María López[e]   Patricia de Sequera[f]   Borja Quiroga[g]   Rocio Echarri[h]

Mario Prieto Velasco[i]   Rafael Díaz[j]   Gonzalo Gómez Marqués[k]

Pilar Sanchez Perez[l]   Vicens Torregrosa[m]   Mariano Rodriguez[b, n]

on behalf of the SAVANA Research Group, SENEFRO-BD-SHPT working group

[a]Department of Nephrology and Hypertension, IIS-Fundación Jimenez Diaz UAM, Madrid, Spain; [b]RICORS2040, Madrid, Spain; [c]Department of Nephrology and Transplantation, University Hospital of Puerta de Hierro, Big Data Working Group of the Spanish Society of Nephrology (BIGSEN), Madrid Spain; [d]Spanish Society of Nephrology, Madrid, Spain; [e]Savana, Madrid, Spain; [f]Department of Nephrology, H. U. Infanta Leonor, Madrid, Spain; [g]Department of Nephrology, H. La Princesa, Madrid, Spain; [h]Department of Nephrology, H. Infanta Sofía, Madrid, Spain; [i]Department of Nephrology, C.A.U. Leon, Leon, Spain; [j]Department of Nephrology, H. Virgen de La Salud, Toledo, Spain; [k]Department of Nephrology, H. Son Espases, Baleares, Spain; [l]Department of Nephrology, H.U. I P. La Fe, Valencia, Spain; [m]Department of Nephrology, Clínic Hospital, Barcelona, Spain; [n]Department of Nephrology, University Hospital Reina Sofía, Córdoba, Spain

Abstract

Introduction: This study aimed to assess the feasibility of applying natural language processing (NLP) to analyze real-world data (RWD) and resolve clinical problems in patients with secondary hyperparathyroidism and chronic kidney disease undergoing hemodialysis (SHPT/CKD-HD). The primary objective was to evaluate how well the guideline-recommended analytical goals are achieved in a Spanish cohort of SHPT/CKD-HD patients based on RWD. **Methods:** Unstructured data in the electronic health records (EHRs) from 8 hospitals were retrospectively analyzed using the *EHRead®* technology, based on NLP and machine learning. Variables extracted from EHRs included demographics, CKD-related clinical characteristics, comorbidities and complications, mineral and bone disorder parameter levels, and treatments at baseline, 6-month, and 12-month follow-up. **Results:** A total of 623 prevalent SHPT/CKD-HD patients were identified; of those, 282 fulfilled the inclusion criteria. They

Alberto Ortiz and Jose Portoles contributed equally to this work.

Correspondence to:
Alberto Ortiz, aortiz@fjd.es

were predominantly elderly males with cardiovascular comorbidities, and the first cause of CKD was diabetic nephropathy. Diagnosis of SHPT was associated with an improvement in median values for PTH, calcium, and phosphate. However, the percentage of patients with normal PTH ranges remained stable during the study period (52.8–60.4%), while the percentage of patients with within-target range serum calcium or phosphate values showed an increasing trend (43.2–60% and 38.8–50%). At baseline, 74.1% of patients were using SHPT-related medication, including at least one vitamin D or analog (63.1%), phosphate binders (46.8%), and/or calcimimetics (9.6%). **Conclusions:** This study represents the first attempt to use clinical NLP to analyze SHPT/CKD-HD patients based on unstructured clinical data. This methodology is useful to address clinical problems based on RWD and identified a high rate of out-of-range mineral-bone analytical values in patients with HPT/CKD-HD and an increasing trend of out-of-range values for serum calcium and phosphate.

Introduction

Secondary hyperparathyroidism (SHPT) is a common complication in patients suffering from chronic kidney disease (CKD) and CKD mineral and bone disorder. It is characterized by the enlargement and hyperactivity of the parathyroid glands, high parathyroid hormone (PTH) levels in the blood, and altered bone and mineral metabolism, most notably altered serum phosphate and calcium levels, which may lead to skeletal fractures, bone disease, and cardiovascular complications with related mortality due to progressive vascular calcification [1–5]. Importantly, SHPT worsens over time in CKD patients undergoing hemodialysis (CKD-HD) [6].

Because of the multifactorial nature of SHPT and the existing knowledge gaps in currently available clinical guidelines, the treatment of SHPT in CKD-HD patients (SHPT/CKD-HD) is particularly challenging [7, 8]. Current treatment options for SHPT include dietary phosphate restriction and use of phosphate binders to reduce serum phosphate levels, administration of vitamin D or vitamin D receptor activators (such as paricalcitol), administration of calcimimetics to inhibit PTH secretion, and surgical parathyroidectomy if medical treatment fails [9, 10]. International *Kidney Disease: Improving Global Outcomes* (KDIGO) guidelines specifically recommend some mineral and bone disorder parameter goals since 2009 for SHPT/CKD-HD patients which include lowering elevated phosphorus levels toward the normal range, maintaining serum calcium in the normal range, and maintaining PTH levels within 2–9 times the upper normal limit since lower PTH levels are associated with increased mortality [11–14]. However, as recently revealed in COSMOS, a prospective study aimed at describing the European dialysis population [15, 16], clinical standards for the prevention, diagnosis, and treatment of SHPT markedly vary across European countries [16], and there are no data regarding the ability to achieve those goals in the routine, day-by-day practice.

In light of the above and the clinical value of real-life information on the management and outcomes of SHPT [6, 14], the analysis of vast amounts of real-world data (RWD) holds great potential to improve our understanding of SHPT/CKD-HD patients from an epidemiological and clinical standpoint. The clinical information in patients' electronic health records (EHRs) represents a paramount source of RWD; particularly, the extraction and analysis of the unstructured clinical information in EHRs using machine learning tools (most notably natural language processing [NLP]) has yielded novel insights into patients' characteristics, management, epidemiological data, and disease prognosis in a variety of clinical areas [17–24].

The SENEFRO-BD-SHPT study aimed to assess the feasibility of applying NLP to analyze the unstructured clinical information in EHRs to answer clinical questions in SHPT/CKD-HD patients using the *EHRead*® deep NLP technology [21–25]. Then, our primary study objective was to evaluate how well the guideline-recommended analytical goals are achieved in a Spanish cohort of SHPT/CKD-HD patients based on RWD.

Materials and Methods

Data Source

This was a multicenter, retrospective study based on the secondary use of the unstructured free-text in the EHRs of 8 hospitals from the Spanish National Healthcare Network (online suppl. Table S1; for all online suppl. material, see www.karger.com/doi/10.1159/000528784). The study period was from January 1, 2014, to December 31, 2018. Data were collected from all available departments in each participating site (including inpatient hospital, outpatient hospital, and emergency service). Of note, access to primary and specialized healthcare (including kidney replacement therapy) is universal and free of charge at the point of delivery in the Spanish National Healthcare System. Online supplementary Figure S1 displays the total number of screened records by the main data sources and hospital departments. The number of records available in structured and unstructured formats in each participating site is shown in online supplementary Figure S2.
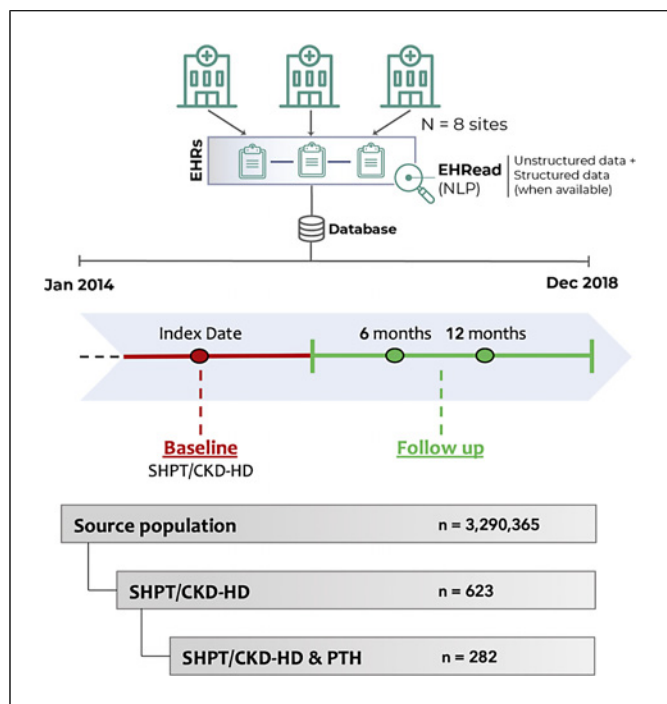
Fig. 1. Study design. Using *EHRead* technology (see Methods section for further details), the unstructured clinical data from patients' EHRs were extracted and analyzed at two different time windows, namely, Baseline and Follow-Up. For all CKD-HD patients, the index date was defined as the timepoint when diagnostic criteria for either HD or SHPT were first met; the baseline period considered the time window between −6 months and +1 month from the index date. Follow-up analyses were performed at 6 and 12 months following the index date, considering a temporal window of ±3 months for data extraction around each timepoint. From a source population of 3,290,365 individuals, a total of 623 patients diagnosed with SHPT undergoing HD were detected. Of these, the final analysis sample comprised 282 patients with available PTH information. NLP, natural language processing; SHPT, secondary hyperparathyroidism; CKD-HD, chronic kidney disease undergoing hemodialysis; PTH, parathyroid hormone.

Study Design

As shown in Figure 1, data were extracted and analyzed at two different time windows, namely, Baseline and Follow-Up. For all patients, an index date was defined as the timepoint at which inclusion criteria for both CKD-HD and SHPT overlapped within the study period (see section below); the baseline period was −6 months to +1 month from the index date. Follow-up analyses were performed at 6 and 12 months following the index date, using a ± 3-month window for data extraction around each timepoint. These windows were chosen to optimize data collection given the limitations in data availability when extracting real-life information and to compensate for possible differences in follow-up visits across patients.

Study Population and Analyzed Variables

The study sample included all CKD-HD patients in the source population with a diagnosis of SHPT, defined as PTH >300 pg/mL

and/or documented use of drugs for the management of SHPT, such as calcimimetics (cinacalcet or etelcalcetide), vitamin D, or vitamin D analogs (paricalcitol, calcifediol, cholecalciferol, calcitriol, and alfacalcidol). To guarantee the homogeneity and quality of the data, the analyses only included patients with available PTH values at both baseline and at least one timepoint during follow-up (Fig. 1). Variables extracted from EHRs included demographics, CKD-related clinical characteristics, comorbidities, and complications at baseline, as well as CKD mineral and bone disorder-related biochemical parameters (PTH, serum calcium, and phosphorus) and treatments at baseline, 6-month, and 12-month follow-up. Included patients were classified regarding levels of each analytical parameter into the following groups in the different timepoints: normal, lower than normal, and higher than normal levels. Patients were labeled as having "normal levels" if their laboratory values ranged as follows: serum PTH = 150–450 pg/mL; serum calcium = 8.4–9.4 mg/dL; and serum phosphate = 2.5–4.5 mg/dL. Regarding laboratory parameters, we performed a depuration for extreme, out-of-range values. Structured laboratory results were included in the analyses for 2 participating sites.

Extracting Clinical Information from EHRs

We used the *EHRead*® technology [21–25] to access and analyze the free-text, unstructured clinical information in EHRs. This technology enables the extraction of clinical information from EHRs via a combination of NLP [26] and other machine learning tools. In this process, the unstructured text is integrated into a structured database via automatic identification of sections in the EHRs as well as detection of key terms (and their synonyms) relevant to the study. Further details regarding clinical data extraction from EHRs are explained in online supplementary information.

In line with previously published procedures [21–24, 27], we assessed *EHRead*®'s ability to identify patient records that contained key variables related to SHPT and CKD-HD (see online suppl. information, evaluation of EHRead's performance). Briefly, this evaluation consisted of a comparison between *EHRead*'s reading output and an annotated corpus of EHRs by expert physicians (i.e., "gold standard"). The level of agreement between *EHRead*'s output and the gold standard was expressed in terms of accuracy (precision), recall, and their harmonic mean F-score (online suppl. Table S2). Both variables used to identify the study population (i.e., CKD-HD and SHPT) obtained high precision scores.

Statistical Analyses

Unless otherwise indicated, summary tables for continuous variables include the mean, standard deviation, median, minimum, maximum, and interquartile range for each variable. Missing data are also indicated. Frequencies are expressed as *n* (%). All analyses were conducted using R software [28].

Results

General Clinical Characteristics

EHRs from 3,290,365 individuals were processed. Of the 623 SHPT/CKD-HD patients at the index date, laboratory information on PTH levels was available for

Table 1. Demographics and CKD-related clinical characteristics at baseline

	SHPT/CKD-HD (n = 282)
Demographics	
Gender	
Female	89 (31.56)
Male	193 (68.44)
Age, years	
Mean (SD)	67.1 (15.4)
Median (Q1–Q3)	71.0 (59.0–79.0)
Missing	0
CKD-related clinical characteristics	
CKD etiology	
Diabetic nephropathy	81 (29.03)
Hypertensive/renal vascular disease	68 (24.37)
Tubulointerstitial disease	54 (19.35)
Glomerular disease	35 (12.54)
Cystic kidney disease	17 (6.09)
Systemic disease	10 (3.58)
CAKUT	9 (3.23)
Miscellaneous renal disorders	4 (1.43)
Other familial nephropathy	3 (1.08)
Unknown	89 (31.90)
Prior kidney transplantation	32 (11.35)
Time since the first occurrence of CKD, years*	
N	273
Mean (SD)	4.5 (5.9)
Median (Q1-Q3)	2.5 (0.5–5.8)
Missing	6

SHPT, secondary hyperparathyroidism; CKD-HD, chronic kidney disease undergoing hemodialysis; CKD, chronic kidney disease; CAKUT, congenital anomalies of the kidney and the urinary tract. Frequencies are shown as n (%). *Time since the first occurrence of CKD diagnosis in patients' EHRs until SHPT.

282 patients, comprising the final study population (Fig. 1).

The general demographic and CKD-related clinical characteristics of the study population at baseline are shown in Table 1. Most patients were male (68.44%; n = 193) with a mean age (±SD) of 67.1 (±15.4) years. The most common cause of CKD was diabetic nephropathy (29%; n = 81), followed by hypertensive/renal vascular disease (24.4%; n = 68), tubulointerstitial disease (19.3%; n = 54), and glomerular disease (12.5%; n = 35). Kidney transplantation was performed in 11.35% of patients (n = 32), with a median (Q1–Q3) time since the first occurrence of CKD of 2.5 (0.5–5.8) years.

Regarding comorbidities and complications at baseline, 83.7% (n = 236) of patients had been diagnosed with hypertension. Diabetes was also relatively common, appearing in more than half of patients (53.9%; n = 152), mostly as type 2 diabetes (T2D; 43.6% of the total sample; n = 123). Consequently, diabetes-related complications such as diabetic nephropathy (29.8%; n = 83) and diabetic foot (3.9%; n = 11) were reported. The most common cardiovascular-related comorbidities at baseline included heart failure (34.8%; n = 98), cardiovascular events (31.9%; n = 90), stroke (16.3%; n = 46), and myocardial infarction (15.2%; n = 43). Fractures were also reported in 20.9% (n = 59) as a SHPT complication. A complete list of comorbidities and complications is shown in Table 2.

Serum PTH, Calcium, and Phosphate

Figure 2a shows levels of PTH, calcium, and phosphorus at baseline, 6 months, and 12 months of follow-up. Of note, follow-up information at 12 months was lost in 15% of the population. Figure 2b shows the percentage of patients in the different evaluated groups (normal, lower, and higher than normal levels) for each specific parameter and at different timepoints. Briefly, approximately half of patients (52.8%; n = 149) showed normal PTH values at baseline; this trend remained relatively stable during follow-up (6-month: 51.8%, n = 132; 12-month: 60.4%, n = 96). Regarding calcium levels, 43.2% of

Ortiz et al.

Table 2. Comorbidities and complications at baseline

	SHPT/CKD-HD* (n = 282)
Comorbidities	
Hypertension	236 (83.7)
Diabetes	152 (53.9)
Diabetes type 2	123 (43.6)
Diabetes type 1	16 (5.7)
Heart failure	98 (34.8)
Cardiovascular event[a]	90 (31.9)
Hypercholesterolemia	59 (20.9)
Stroke	46 (16.3)
Ischemic stroke	22 (7.8)
Non-ischemic stroke	6 (2.1)
Myocardial infarction	43 (15.2)
NSTEMI/STEMI[b]	7 (2.5)
Valvular or vascular calcification	37 (13.1)
Peripheral artery disease	23 (8.2)
Hypothyroidism	19 (6.7)
Transient ischemic attack	14 (5.0)
Carotid artery disease	7 (2.5)
Liver impairment	1 (0.4)
Complications	
Diabetic nephropathy	84 (29.8)
Retinopathy	62 (22.0)
Fractures	59 (20.9)
Diabetic foot	11 (3.9)

*Data indicate single diagnostic labels found in patients' EHRs (i.e., if diagnostic information for a given condition exists multiple times for a single patient, the condition was counted only once). Of note, a single patient could have been diagnosed with more than one of the analyzed comorbidities. [a]Cardiovascular event includes stroke, myocardial infarction, unstable angina, and/or coronary revascularization. [b](Non/N) ST-segment elevation myocardial infarction.

patients (n = 117) showed values within normal range at baseline, 51.4% (n = 133) at 6 months, and 60% (n = 114) at 12 months. Finally, the percentage of patients with normal phosphate levels increased from 38.8% (n = 62) to 50.7% (n = 69) at 6 months and to 50% (n = 50) at 12 months post-baseline. A complete description of these data is included in online supplementary Table S3.

Medical Management of SHPT in Hemodialysis Patients

Figure 3 shows the use of selected SHPT-related medications across the study period. At baseline, 74.1% of patients (n = 209) were using SHPT-related medication, including at least one vitamin D/vitamin D analog (63.1%; n = 178), phosphate binder (46.8%; n = 132), and/or calcimimetics (9.6%; n = 27). Among phosphate binders, the most frequently prescribed drugs at baseline

were calcium carbonate (21.6%; n = 61) and sevelamer (20.6%; n = 58).

Changes in medication frequencies over time are displayed in Figure 3 and summarized in online supplementary Table S4. Over the follow-up period, the use of vitamin D and analogs decreased from 63.1% at baseline to 40% at 12 months. On the other hand, the use of calcimimetics and phosphate binders remained relatively stable during this period.

Discussion

The main findings of this study are that (1) NLP is a feasible tool to access and analyze the unstructured clinical information in the EHRs of patients with SHPT and CKD-HD in Spain; (2) NLP based on RWD allows us to address clinical problems; and (3) although after SHPT/CKD-HD diagnosis there is an improvement in the management of minerals and bone disorder parameters, almost half of patients do not achieve guideline-recommended analytical goals after 1 year, most frequently phosphorus levels.

NLP and machine learning are increasingly being used to extract and analyze clinical data from patients' EHRs in the broader context of CKD and HD. In a recent study, NLP identified HD-related symptoms with higher specificity and sensitivity than using International Classification of Diseases coding alone [29]. Crucially, NLP-extracted data from EHRs also identified novel clinical factors that predict outcomes in end-stage renal disease [30–32]. In this regard, NLP compared with traditional research methods presents some advantages. First, NLP represents a cost- and time-effective technology to extract and analyze large amounts of readily available real-world clinical evidence [18, 20, 33–35]. Second, NLP-based analyses of RWD may be repeated periodically over time in an efficient and feasible manner to explore the impact on health outcomes of new drugs, procedures, eHealth interventions, or any other interventions performed during routine clinical practice. In the context of SHPT, recent studies have pointed to stark differences in the characteristics and management of CKD patients undergoing HD [16] and highlight the pressing need for PTH RWD [6, 14].

From a source population of 3,290,365 individuals, our study sample comprised 282 SHPT/CKD-HD patients with available PTH values during the study period. The gender distribution, age, and CKD etiology are in line with the epidemiology data of kidney replacement therapy in Spain [36] as well as with previously reported
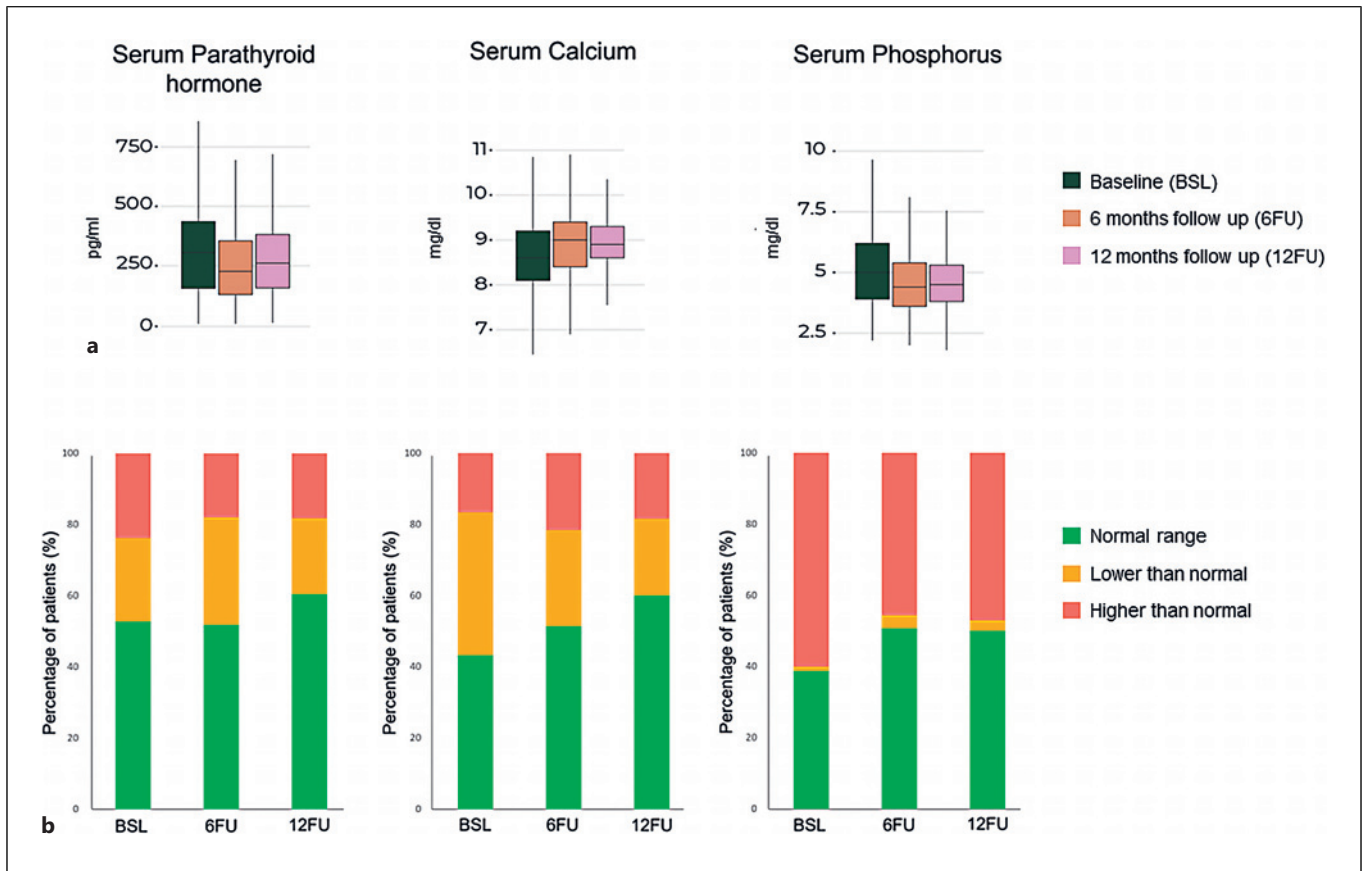
Fig. 2. Laboratory results at baseline and follow-up. **a** Laboratory values of serum PTH (pg/mL), calcium (Ca, mg/dL), and phosphate (P, mg/dL) in the SHPT/CKD-HD and PTH population at baseline (green), 6-month follow-up (orange), and 12-month follow-up (pink). The median is represented by a line within the colored boxes, enclosed by Q1 at the bottom and Q3 at the top. **b** Percentage of patients with normal levels of PTH (150–450 pg/mL), calcium (Ca, 8.4–9.4 mg/dL), and phosphate (P, 2.5–4.5 mg/dL) in green, and values outside normal ranges at baseline, 6-month follow-up, and 12-month follow-up (in yellow and red). See online supplementary Table S3 for complete data analysis. EHRs, electronic health records; SHPT/CKD-HD, secondary hyperparathyroidism and chronic kidney disease undergoing hemodialysis; PTH, parathyroid hormone.

results in larger European series of SHPT patients undergoing HD, including the prospective observational COSMOS [16], the retrospective observational Mimosa study in France [6], the global EVOLVE randomized trial aimed to assess the efficacy of cinacalcet [37], and the global observational DOPPS 2012–2015 database of SHPT patients on chronic HD [38]. In the present study, the main causes of CKD also matched those reported in COSMOS [16].

Cardiovascular diseases are the most common cause of death in patients with CKD with SHPT [39]. In our series, diabetes and cardiovascular diseases (i.e., hypertension, heart failure, and cardiovascular events) were the most frequent comorbidities at baseline in the study population, which was in line with prior observational studies and clinical trials, such as DOPPS, COSMOS, and

EVOLVE. Moreover, SHPT also has been associated with progressive bone disease and fractures [40].

Levels of laboratory mineral and bone disorder parameters at baseline agree with those previously reported in the DOPPS report for Spain (2012–2015) [38]. In addition, percentage of patients with normal levels at study entry falls within the range observed in the same report, while considering the slight differences in the range of analytical values reported as normal [38]. As explained before, KDIGO guidelines recommend maintaining calcium and phosphorus within normal ranges while PTH within 2–9 times the upper normal limit after SHPT diagnosis. According to the unstructured, free-text information in patients' EHRs, our results showed an improvement in the control of all analytical parameters during follow-up, as evidenced by the increase in the
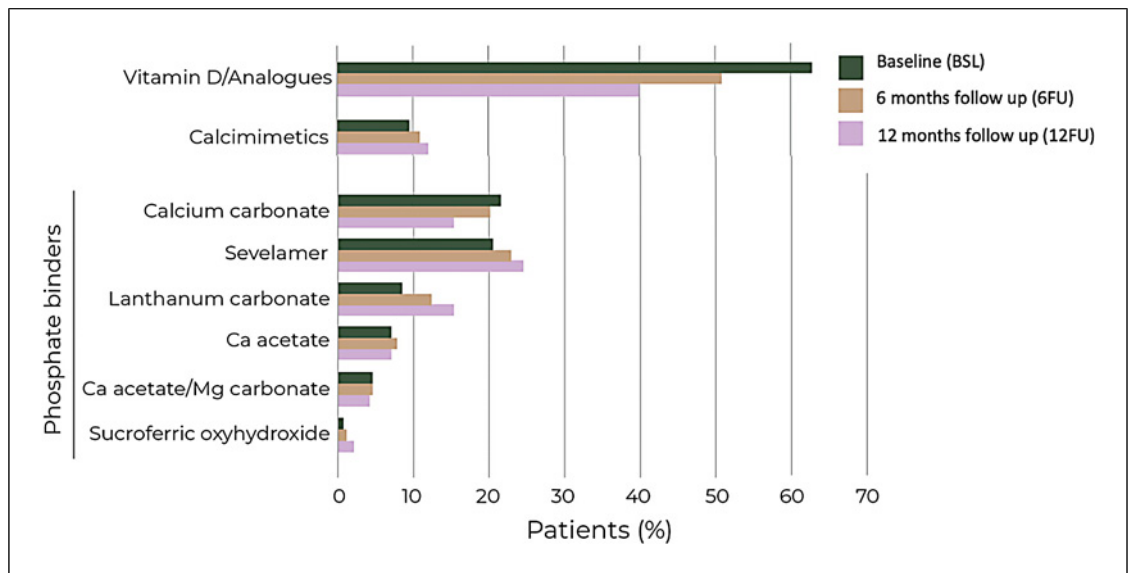
Fig. 3. Percentage of patients prescribed with SHPT-related medication at the index date (green), 6-month follow-up (orange), and 12-month follow-up (pink). Vitamin D/analogs include paricalcitol, calcifediol, cholecalciferol, calcitriol, and alfacalcidol. See online supplementary Table S4 for complete data analysis.

percentage of patients with levels within normal ranges at the consecutive timepoints (from 40 to 50% at baseline to 50–60% at 12-month follow-up). Moreover, recommendations for PTH in these patients allow maintenance of levels at upper normal limits, and the improvement in this specific biomarker could be even higher. However, despite the widespread use of SHPT-related medication already at baseline, only half of the patients reached "normal" median values during follow-up. Because extreme PTH values are associated with negative prognosis in hemodialyzed SHPT patients and scientific evidence supporting a definite PTH target is lacking (current recommendation grade 2C), additional studies linking laboratory values with treatment outcomes and disease prognosis are needed [41].

Regarding SHPT-related medications, vitamin D and its analogs and calcimimetics were used in ranges that fell within the previously reported [16, 38, 41, 42]. Notably, the use of etelcalcetide was underrepresented in our series likely related to the timeline of available data (etelcalcetide only became available during the last months of study time window). Unlike previous literature, only 46.8% of patients included in our study were prescribed phosphate binders at baseline, a number that remained relatively stable during follow-up [16, 38, 41–43]. However, the distribution of specific phosphate binders was in the range found in DOPPS and COSMOS.

Limitations

The results presented here must be interpreted considering several limitations intrinsic to the novel methodology used with RWD. First, the quality and accuracy of the data are limited by the extent to which physicians accurately describe patients' status in medical records. The lack of standardization in EHRs in terms of the type of collected data across disciplines, use of standard versus in-house medical terminology, omitted information, or misuse of EHR sections represent a methodological challenge and may have contributed to the heterogeneity of the data [44, 45]. Second, the lack of structured clinical information is particularly jeopardizing when laboratory results are a key to characterize the population of interest; combining both unstructured and structured information in future studies will lead to larger sample sizes and a more accurate depiction of the SHPT/CKD-HD population. Third, the availability of the desired variables at the selected timepoints across patients could not be guaranteed. Moreover, phosphate binder related data must be interpreted considering the low recall metric (0.43) for the variable "phosphate" obtained by our NLP system. While calculations regarding calcimimetics use could be optimized by having a complete access to hospital pharmacy records in future studies, an adequate representation of phosphate binder prescription requires the implementation of technological improvements in the NLP system and access to department-specific EHRs (e.g.,

specific software in HD units). Finally, in light of recent epidemiological data of HD in Spain [36], it is likely that some hemodialysis patients were missed in the initial screening. Indeed, the assessment of the NLP system revealed important differences in the recall and precision metrics of key study variables. Patients undergoing hemodialysis in out-of-hospital facilities may have been missed by our system since they may not have EHRs (or these might be incomplete) available in the reference participating hospitals.

Conclusion

This study represents the first attempt to apply NLP to extract and analyze real-life clinical information of hemodialyzed patients with SHPT in a multicenter setting. Unstructured data in EHRs allow for addressing clinical aspects which could be important for the patient's management. Based on this technology, it is possible to describe how well guideline recommendations regarding laboratory mineral and bone disorder parameters are achieved in this population. Despite an improvement in median values after the diagnosis and treatment of SHPT, an important percentage of patients remained in undesirable, out-of-guideline target ranges of PTH, calcium, and phosphorus after 1 year, especially for phosphorus. Clinicians should take these results into account when managing patients with SHPT and CKD-HD to improve outcomes.

In light of these promising results, future applications of machine learning tools to access EHR data will provide a large-scale clinical picture of SHPT, inform treatment strategies, improve patient safety, and guide health policy decisions in the hemodialyzed population [46]. In parallel, raising awareness among healthcare professionals about the importance of EHR completeness will undoubtedly boost the impact of medical research and ultimately improve patient care and hospital resource management [47–49].

Statement of Ethics

The present study was conducted within the scope of the SENEFRO-BD-SHPT project. This study was classified as a "non-post-authorization study" (EPA) by the Spanish Agency of Medicines and Health Products (AEMPS) and was conducted in compliance with legal and regulatory requirements and in agreement with generally accepted research practices described in the Helsinki Declaration in its latest edition, Good Pharmacoepidemiology Practices, and applicable local regulations. The study was approved by the Spanish Ethic Committee for Research with Medicinal Products (CEIm; ref ID #EO109-19_FJD). Because data were collected in a retrospective manner from EHRs in an anonymized and aggregated manner, patient consent was waived.

Conflict of Interest Statement

Alberto Ortiz has received consultancy, speaker fees, or travel support from AstraZeneca, Amicus, Amgen, Fresenius Medical Care, Bayer, Sanofi-Genzyme, Menarini, Kyowa Kirin, Alexion, Otsuka, and Vifor Fresenius Medical Care Renal Pharma and is Director of the Catedra Mundipharma-UAM of diabetic kidney disease and the Catedra AstraZeneca-UAM of chronic kidney disease and electrolytes. Jose Portoles has received consultancy, speaker fees, or travel support from Amgen, Astellas, Alexion, and Vifor Fresenius Medical Care Renal Pharma; none of them related with present study. Patricia de Sequera has received consultancy or speaker fees or travel support from AstraZeneca, Amgen, Fresenius Medical Care, Baxter, Nipro, Vifor Fresenius Medical Care Renal Pharma, and Astellas. Mariano Rodriguez has received consultancy or speaker fees or travel support from Amgen, Kiowa Kirin, and Viphor. Borja Quiroga has received consultancy or speaker fees or travel support from Amgen, Astellas, Sanofi-Genzyme, Otsuka, and Vifor Fresenius Medical Care Renal Pharma; none of them related with present study. Jose Portoles has received consultancy or speaker fees or travel support from Vifor Fresenius Medical Care, Astellas, Otsuka, and GSK during the last 3 years; none of them related with present study.

Author Contributions

Alberto Ortiz, Jose Portoles, and SAVANA Research Group (Jesus Barea and Maria Lopez) conceptualized and designed the study, analyzed the data, and interpreted the study results. The original manuscript was drafted by SAVANA Research Group (Jesus Barea and Maria Lopez); all authors revised the manuscript and contributed for intellectual content. Alberto Ortiz, Jose Portoles, Maria Dolores Pino-Pino, Patricia de Sequera, Borja Quiroga, Rocio Echarri, Mario Prieto Velasco, Rafael Díaz, Gonzalo Gómez Marqués, Pilar Sanchez Perez, Vicens Torregrosa, and Mariano Rodriguez contributed to different aspects of study completion, including the coordination of the medical experts who participated in the annotation of the "gold standard" corpus for the evaluation of the NLP system.

Data Availability Statement

Data cannot be shared publicly because of contractual obligations between SAVANA (the research company providing the NLP system used to extract and analyze the data), the study sponsor (Amgen/Spanish Society of Nephrology), and the participating hospital sites that allowed access to anonymized patient information and ultimately own the data. All data obtained, generated, and analyzed belong to a third party, namely, the participating hospitals in the study. Data extraction, database generation, and analysis were performed by SAVANA. All data were obtained via SAVANA, which had access to the anonymized, hospital-owned data. Thus, individual authors did not have special access privileges to the data. Further requests regarding data availability must be sent to SAVANA Institutional Data Access (Miren Taberna, Chief Scientific Officer; contact via e-mail: mtaberna@savanamed.com) for researchers who meet the criteria for access to confidential data.

References

1 Cunningham J, Locatelli F, Rodriguez M. Secondary hyperparathyroidism: pathogenesis, disease progression, and therapeutic options. Clin J Am Soc Nephrol. 2011;6(4):913–21.

2 Cozzolino M, Ureña-Torres P, Vervloet MG, Brandenburg V, Bover J, Goldsmith D, et al. Is chronic Kidney Disease-Mineral Bone Disorder (CKD-MBD) really a syndrome? Nephrol Dial Transplant. 2014;29(10):1815–20.

3 Block GA, Hulbert-Shearon TE, Levin NW, Port FK. Association of serum phosphorus and calcium x phosphate product with mortality risk in chronic hemodialysis patients: a national study. Am J Kidney Dis. 1998;31(4):607–17.

4 Goodman WG, Goldin J, Kuizon BD, Yoon C, Gales B, Sider D, et al. Coronary-artery calcification in young adults with end-stage renal disease who are undergoing dialysis. N Engl J Med. 2000;342(20):1478–83.

5 Ganesh SK, Stack AG, Levin NW, Hulbert-Shearon T, Port FK. Association of elevated serum PO(4), Ca x PO(4) product, and parathyroid hormone with cardiac mortality risk in chronic hemodialysis patients. J Am Soc Nephrol. 2001;12(10):2131–8.

6 Rottembourg J, Ureña-Torres P, Toledano D, Gueutin V, Hamani A, Coldefy O, et al. Factors associated with parathyroid hormone control in haemodialysis patients with secondary hyperparathyroidism treated with cinacalcet in real-world clinical practice: Mimosa study. Clin Kidney J. 2019;12(6):871–9.

7 Ketteler M, Block GA, Evenepoel P, Fukagawa M, Herzog CA, McCann L, et al. Diagnosis, evaluation, prevention, and treatment of chronic kidney disease-mineral and bone disorder: synopsis of the kidney disease: improving global outcomes 2017 clinical practice guideline update. Ann Intern Med. 2018;168(6):422–30.

8 Block GA, Martin KJ, de Francisco ALM, Turner SA, Avram MM, Suranyi MG, et al. Cinacalcet for secondary hyperparathyroidism in patients receiving hemodialysis. N Engl J Med. 2004;350(15):1516–25.

9 Friedl C, Zitt E. Role of etelcalcetide in the management of secondary hyperparathyroidism in hemodialysis patients: a review on current data and place in therapy. Drug Des Devel Ther. 2018;12:1589–98.

10 Bolasco P. Treatment options of Secondary Hyperparathyroidism (SHPT) in patients with chronic kidney disease stages 3 and 4: an historic review. Clin Cases Miner Bone Metab. 2009;6(3):210–9.

11 Tabibzadeh N, Karaboyas A, Robinson BM, Csomor PA, Spiegel DM, Evenepoel P, et al. The risk of medically uncontrolled secondary hyperparathyroidism depends on parathyroid hormone levels at haemodialysis initiation. Nephrol Dial Transplant. 2021;36(1):160–9.

12 Merle E, Roth H, London GM, Jean G, Hannedouche T, Bouchet JL, et al. Low parathyroid hormone status induced by high dialysate calcium is an independent risk factor for cardiovascular death in hemodialysis patients. Kidney Int. 2016;89(3):666–74.

13 Striker GE. Beyond phosphate binding: the effect of binder therapy on novel biomarkers may have clinical implications for the management of chronic kidney disease patients.. Kidney Int. 2009;76(114):S1–2.

14 Kidney Disease: Improving Global Outcomes KDIGO CKD-MBD Update Work Group. KDIGO 2017 clinical practice guideline update for the diagnosis, evaluation, prevention, and treatment of Chronic Kidney Disease-Mineral and Bone Disorder (CKD-MBD). Kidney Int Suppl. 2017;7(1):1–59.

15 Cannata-Andia JB, Fernandez-Martin JL, Zoccali C, London GM, Locatelli F, Ketteler M, et al. Current management of secondary hyperparathyroidism: a multicenter observational study (COSMOS). J Nephrol. 2008; 21(3):290–8.

16 Fernandez-Martin JL, Carrero JJ, Benedik M, Bos WJ, Covic A, Ferreira A, et al. COSMOS: the dialysis scenario of CKD-MBD in Europe. Nephrol Dial Transplant. 2013;28(7):1922–35.

17 Del Rio-Bermudez C, Yebes L, Poveda JL. Towards a symbiotic relationship between big data, artificial intelligence, and hospital pharmacy. J Pharm Pol Pract. 2020. 13(1):75.

18 Sheikhalishahi S, Miotto R, Dudley JT, Lavelli A, Rinaldi F, Osmani V. Natural Language processing of clinical notes on chronic diseases: systematic review. JMIR Med Inform. 2019;7(2):e12239.

19 Goldstein BA, Navar AM, Pencina MJ, Ioannidis JPA. Opportunities and challenges in developing risk prediction models with electronic health records data: a systematic review. J Am Med Inform Assoc. 2017;24(1):198–208.

20 Luo Y, Thompson WK, Herr TM, Zeng Z, Berendsen MA, Jonnalagadda SR, et al. Natural Language processing for EHR-based pharmacovigilance: a structured review. Drug Saf. 2017;40(11):1075–89.

21 Izquierdo JL, Almonacid C, González Y, Del Rio-Bermúdez C, Ancochea J, Cárdenas R, et al. The impact of COVID-19 on patients with asthma. Eur Respir J. 2021;57(3):2003142.

22 Ancochea J, Izquierdo JL, Medrano IH, Porras A, Serrano M, Lumbreras S, et al. Evidence of gender differences in the diagnosis and management of COVID-19 patients: an analysis of electronic health records using natural language processing and machine learning. J Women Health. 2020;30(3):393–404. In press.

23 Izquierdo JL, Soriano JB, Soriano JB. Authors' reply to: minimizing selection and classification biases comment on "clinical characteristics and prognostic factors for intensive care unit admission of patients with COVID-19: retrospective study using machine learning and natural language processing". J Med Internet Res. 2021;23(5):e29405.

24 Graziani D, Soriano JB, Del Rio-Bermudez C, Morena D, Díaz T, Castillo M, et al. Characteristics and prognosis of COVID-19 in patients with COPD. J Clin Med. 2020;9(10):3259.

25 Hernandez Medrano I, Tello Guijarro J, Belda C, Urena A, Salcedo I, Espinosa-Anke L. Savana: Re-using electronic health records with artificial intelligence. Int J Interactive Multimedia Artif Intelligence. 2018;4(7):8–12.

26 Wong A, Plasek JM, Montecalvo SP, Zhou L. Natural Language processing and its implications for the future of medication safety: a narrative review of recent advances and challenges. Pharmacotherapy. 2018;38(8):822–41.

27 Gomollón F, Gisbert JP, Guerra I, Plaza R, Pajares Villarroya R, Moreno Almazán L, et al. Clinical characteristics and prognostic factors for crohn's disease relapses using natural language processing and machine learning: a pilot study. Eur J Gastroenterol Hepatol. 2022. 34(4):389–397. in press.

28 Team RC. Statistical computing. Vienna, Austria: R Foundation for Statistical Computing. URL 2019.

29 Chan L, Beers K, Yau AA, Chauhan K, Duffy Á, Chaudhary K, et al. Natural language processing of electronic health records is superior to billing codes to identify symptom burden in hemodialysis patients. Kidney Int. 2020;97(2):383–92.

30 Zeng XX, Liu J, Ma L, Fu P. Big data research in chronic kidney disease. Chin Med J. 2018;131(22):2647–50.

31 Perotte A, Ranganath R, Hirsch JS, Blei D, Elhadad N. Risk prediction for chronic kidney disease progression using heterogeneous electronic health record data and time series analysis. J Am Med Inform Assoc. 2015;22(4):872–80.

32 Singh K, Betensky RA, Wright A, Curhan GC, Bates DW, Waikar SS. A concept-wide association study of clinical notes to discover new predictors of kidney failure. Clin J Am Soc Nephrol. 2016;11(12):2150–8.

33 Del Rio-Bermudez C, Medrano IH, Yebes L, Poveda JL. Towards a symbiotic relationship between big data, artificial intelligence, and hospital pharmacy. J Pharm Policy Pract. 2020;13(1):75.

34 Divita G, Carter M, Redd A, Zeng Q, Gupta K, Trautner B, et al. Scaling-up NLP pipelines to process large corpora of clinical notes. Methods Inf Med. 2015;54(6):548–52.

35 Neuraz A, Lerner I, Digan W, Paris N, Tsopra R, Rogier A, et al. Natural Language processing for rapid response to emergent diseases: case study of calcium channel blockers and hypertension in the COVID-19 pandemic. J Med Internet Res. 2020;22(8):e20773.

36 Patients SRoR. Dialysis and renal transplants in Spain: 2014 Report. Valencia (Spain): Oral presentation at: XLV National Congress of the Spanish Society of Nephrology; October 2015.

37 EVOLVE Trial Investigators; Chertow GM, Block GA, Correa-Rotter R, Drueke TB, Floege J, et al. Effect of cinacalcet on cardiovascular disease in patients undergoing dialysis. N Engl J Med. 2012;367(26):2482–94.

38 Cozzolino M, Shilov E, Li Z, Fukagawa M, Al-Ghamdi SMG, Pisoni R, et al. Pattern of laboratory parameters and management of secondary hyperparathyroidism in countries of europe, asia, the Middle East, and North America. Adv Ther. 2020;37(6):2748–62.

39 United States Renal Data System. 2019 USRDS Annual Data Report: Epidemiology of kidney disease in the United States. National Institutes of Health, National Institute of Diabetes and Digestive and Kidney Diseases Bethesda, MD, 2019. The data reported here have been supplied by the United States Renal Data System (USRDS) The interpretation and reporting of these data are the responsibility of the author(s) and in no way should be seen as an official policy or interpretation of the US Government. 2019.

40 Xu Y, Evans M, Soro M, Barany P, Carrero JJ. Secondary hyperparathyroidism and adverse health outcomes in adults with chronic kidney disease. Clin Kidney J. 2021;14(10):2213–20.

41 Tentori F, Wang M, Bieber BA, Karaboyas A, Li Y, Jacobson SH, et al. Recent changes in therapeutic approaches and association with outcomes among patients with secondary hyperparathyroidism on chronic hemodialysis: the DOPPS study. Clin J Am Soc Nephrol. 2015;10(1):98–109.

42 Fernández-Martín JL, Dusso A, Martínez-Camblor P, Dionisi MP, Floege J, Ketteler M, et al. Serum phosphate optimal timing and range associated with patients survival in haemodialysis: the COSMOS study. Nephrol Dial Transpl. 2019;34(4):673–81.

43 Russo D, Tripepi R, Malberti F, Di Iorio B, Scognamiglio B, Di Lullo L, et al. Etelcalcetide in patients on hemodialysis with severe secondary hyperparathyroidism. Multicenter study in "real life". J Clin Med. 2019;8(7):1066.

44 Chan KS, Fowles JB, Weiner JP. Review: electronic health records and the reliability and validity of quality measures: a review of the literature. Med Care Res Rev. 2010;67(5):503–27.

45 Wright A, McCoy AB, Hickman TT, Hilaire DS, Borbolla D, Bowes WA 3rd, et al. Problem list completeness in electronic health records: a multi-site study and assessment of success factors. Int J Med Inform. 2015;84(10):784–90.

46 Abdel-Kader K, Jhamb M. EHR-Based Clinical Trials: The Next Generation of Evidence. Clin J Am Soc Nephrol. 2020;15(7):1050–2.

47 Hong CJ, Kaur MN, Farrokhyar F, Thoma A. Accuracy and completeness of electronic medical records obtained from referring physicians in a Hamilton, Ontario, plastic surgery practice: a prospective feasibility study. Plast Surg. 2015;23(1):48–50.

48 Lai FW, Kant JA, Dombagolla MH, Hendarto A, Ugoni A, Taylor DM. Variables associated with completeness of medical record documentation in the emergency department. Emerg Med Australas. 2019;31(4):632–8.

49 Wu CHK, Luk SMH, Holder RL, Rodrigues Z, Ahmed F, Murdoch I. How do paper and electronic records compare for completeness? A three centre study. Eye. 2018;32(7):1232–6.