

# Analysis of Nationwide Data to Determine the Incidence and Diagnosed Prevalence of Autosomal Dominant Polycystic Kidney Disease in the USA: 2013–2015

Cynthia Willey<sup>a</sup> Siddhesh Kamat<sup>b</sup> Robert Stellhorn<sup>b</sup> Jaime Blais<sup>b</sup>

<sup>a</sup>Division of Health Outcomes, College of Pharmacy, University of Rhode Island, Kingston, RI, USA;

<sup>b</sup>Otsuka Pharmaceutical Development and Commercialization, Inc., Princeton, NJ, USA

## Keywords

Autosomal Dominant Polycystic Kidney Disease · Epidemiology · Incidence · Prevalence

## Abstract

**Background:** This study addresses an important gap, as it is the first US nationwide, epidemiologic study of ADPKD incidence and prevalence. **Summary:** This 3-year, observational study utilized data from Truven Health MarketScan<sup>®</sup> administrative claims, as well as cross-sectional data from the National Ambulatory Medical Care Survey (NAMCS). We estimated the annual incidence and diagnosed prevalence using population-based data on over 170 million de-identified patients to provide the most current epidemiologic estimates available. The ADPKD-diagnosed prevalence was 4.3 per 10,000 in the NAMCS, which closely corresponded with age-adjusted rates from patients with either commercial insurance or employer-sponsored Medicare supplemental insurance. The annual incidence was 0.62 per 10,000. Both nationwide data sets indicate that approximately 140,000 patients are currently diagnosed in the USA. We also found significant differences by gender and age. Females are nearly twice as likely as males to be diagnosed in early adulthood, while the incidence in males was highest

in those aged 65 years or older. ADPKD appears more likely to be diagnosed in men after disease progression or the development of chronic kidney disease. **Key Messages:** Our results revealed striking age and gender differences in the incidence of ADPKD. Young women are diagnosed with ADPKD at nearly twice the rate of young men, perhaps due to the use of ultrasound in women during child-bearing years. This points to a need for increased recognition of ADPKD, with an emphasis on younger men in particular. ADPKD has been inaccurately perceived as a common condition based on misinterpretation of early epidemiologic data (1957) confirmed by our data and recent European data. ADPKD affects approximately 140,000 patients in the USA and meets the criterion for a rare disease. Our results indicate a need for further study of gender and ADPKD diagnosis, progression, management, and outcomes.

© 2019 The Author(s)

Published by S. Karger AG, Basel

## Background and Rationale

Autosomal dominant polycystic kidney disease (ADPKD) is a monogenic, hereditary, and rare kidney disorder, resulting in approximately 5% of total end-stage renal disease (ESRD) cases [1, 2] in the USA. ADPKD

places a substantial economic burden on healthcare systems, driven largely by ESRD expenditures [3–6]. In 2014, 31,794 ESRD cases were attributable to polycystic kidney disease (PKD), making it the 4th leading cause of ESRD behind diabetes, hypertension, and glomerulonephritis [7] in the USA. ADPKD differs from most other causes of ESRD because it can be detected early in life, decades before the onset of ESRD. Thus, epidemiologic data on the full clinical spectrum of affected patients may be beneficial for understanding the lifetime, clinical course of ADPKD and for planning interventions to delay or prevent ESRD.

To better understand the burden of ADPKD in the USA, current estimates of the prevalence and incidence of ADPKD are essential, but estimating ADPKD prevalence at the population level is challenging. At any given time, a significant fraction of ADPKD cases are asymptomatic and thus undiagnosed [8]. This feature reflects, in part, the progression of ADPKD over decades, with clinically detectable reductions in renal function typically appearing during the third to fourth decade of life [9–14]. The high cost of genetic testing and diagnostic imaging limits the feasibility of general population screening for ADPKD, and thus population-based data are more limited for ADPKD than for more common conditions.

### Objectives

Although several epidemiologic studies of ADPKD have recently been published in Europe [15–19], contemporary estimates of the incidence and prevalence in the USA are lacking. This study examines minimum or diagnosed prevalence, defined as the prevalence of diagnosed ADPKD cases divided by the population at a single point in time. This measure of disease burden is often used when data sources are limited to diagnosed cases, and it has been applied in other studies of autosomal dominant genetic disorders, such as Huntington's disease [20, 21]. The objectives of the present study were to estimate the annual incidence and prevalence of diagnosed ADPKD in the USA using patient level claims data and the National Ambulatory Medical Care Survey (NAMCS), and to compare these estimates with recently published epidemiologic studies and nationally generalizable, epidemiologic data resources.

## Methods

### Primary Data Sources

#### Truven Health MarketScan® Databases

Administrative claims data have been routinely used to estimate the prevalence and incidence of disease [22–24]. The MarketScan® Research Databases contain information on approxi-

mately 170 million de-identified patients and provide information reflecting real-world treatment patterns and encounters. These data include a collection of databases containing individual level health care claims, and hospital discharge information from a variety of fee-for-service, fully and partially capitated health plans. The data reflect the full range of health care settings in the USA such as physician offices, managed care organizations, and hospitals, as well as a full range of insurance plans including employer-sponsored commercial insurance, employer-sponsored Medicare Supplement for members who are eligible for Medicare, and state agency-sponsored Medicaid.

### National Ambulatory Medical Care Survey

The NAMCS, conducted by the Centers for Disease Control and Prevention, and the National Center for Health Statistics (NCHS), provides objective, nationally representative data about the use of ambulatory medical care services in the USA. Findings are based on a stratified random sample of visits to nonfederally employed physicians who are primarily engaged in direct patient care. The NAMCS contains information on patient symptoms, physician diagnoses, medications prescribed, diagnostic procedures, planned future treatment, and demographics for a random sample of patient visits collected during a randomly selected 1-week time period. Data is recorded primarily by US Census field representatives using an automated laptop-assisted instrument developed by the NCHS, although very rarely data were recorded by physicians or their staff.

### US Renal Data System

The United States Renal Data System (USRDS) is a national data system that collects, analyzes, and distributes information about chronic kidney disease and ESRD in the USA. The USRDS includes demographic and clinical information on all US treated ESRD patients who qualify for Medicare and who survive a minimum of 90 days on renal replacement therapy [2, 7, 25].

### Literature Review

We also performed a literature review of epidemiologic studies of ADPKD using the National Library of Medicine PubMed database to identify publications from January 1, 2000, to March 1, 2018. The search used the following Medical Subject Heading (MeSH) terms: ("polycystic kidney," or "autosomal dominant polycystic kidney disease," or "polycystic kidney disease") and ("incidence study," or "prevalence study," or "epidemiologic study"). The time period was chosen to include only contemporary, epidemiologic estimates based on case identification procedures utilizing modern diagnostic methods (such as ultrasound) for ADPKD detection.

### Study Design

#### Truven Health MarketScan® Sample

For this analysis the following 3 longitudinal MarketScan® databases were used: commercial claims or employer-sponsored private health insurance, employer-sponsored Medicare Supplement insurance, and Managed Medicaid. Two prevalence estimates were derived, i.e., one in which the numerator included patients of any age with at least 1 annual medical claim with an associated diagnosis of ADPKD or polycystic kidney disease (unspecified type) (ICD-9-CM: 753.13 and 753.12, respectively or ICD10: Q61.2, Q61.3) and a second estimate that specified 2 or more annual medical claims with an ADPKD diagnosis. ADPKD prevalence was examined separately in the combined commercial claims and

Medicare data, versus the Medicaid sample. The prevalence denominator included the number of patients enrolled into each health plan during the prevalence year. Incident cases in 2014 were identified as those without an ADPKD diagnosis claim during 2013, but with 2 or more ADPKD diagnosis claims in 2014, and continuous eligibility in 2013 to 2014. Similarly, incident cases in 2015 were those that were without a diagnosis code in 2014, and had 2 or more diagnosis codes in 2015, with continuous eligibility in 2014–2015. The incidence denominator included the number of patients enrolled into each health plan during the incidence year minus those who had already been identified as prevalent patients.

#### NAMCS Sample

We examined the 3 most recent years (2012–2014) of NAMCS data to determine the average annual period prevalence of diagnosed ADPKD in the USA for comparison with claims data analyses. The prevalence numerator included patients of any age with a primary, secondary, or tertiary diagnosis of ADPKD or polycystic kidney disease (unspecified type) during 2012–2014 (ICD-9-CM: 583.12 or 583.13). Due to the rarity of ADPKD, it was necessary to combine multiple years of data to obtain an adequate number of patient care visits for analysis. Sampling weights were applied to generate unbiased nationally representative estimates of outpatient visits throughout the USA. The combined sample from 2012 to 2014 provides information on 176,913 unweighted visits with a weighted size of  $n = 2,735,933,568$ .

#### US Renal Data System

This cross-sectional analysis examined the most recently available data to determine the number and rate of point prevalent, treated ESRD patients with ADPKD in the USA on December 31, 2014. ADPKD patients were identified with the code PKD as the “primary cause of renal failure,” as assessed by individual physicians and reported on the CMS 2728 form. The prevalence denominator was the US population at the midpoint of the year 2014 based on estimates from the US Census Bureau [26].

#### Statistical Methods

NAMCS survey data were analyzed using the sampled visit weight that is the product of the corresponding sampling fractions at each stage in the sample design. The sampling weights have been adjusted by NCHS for survey nonresponse within time of year, geographic region, urban/rural, and ownership designations, yielding an unbiased national estimate of visit occurrences, percentages, and characteristics. Because of the complex sample design, sampling errors were determined using SAS v9.3 with SUDAAN® from the Research Triangle Institute (RTI International). SUDAAN® procedures properly account for complex design features, such as correlated observations, clustering, weighting, and stratification which takes into account the clustered nature of the sample.

Additional reweighting of encounter data was performed to estimate patient level counts using a method based on the principles of a multiplicity estimator to convert the NAMCS provider statistics aggregated at the visit level to statistics aggregated at the patient level. This method was developed by the NCHS and has been validated by comparing the results of this method with national surveys utilizing patient level data [27]. The average annual prevalence was calculated by dividing the 3-year total number of estimated ADPKD patients by 3 and then dividing by the population at the midpoint of the 3-year period.

Descriptive statistics such as frequencies and percent were generated for both the NAMCS data and the Truven Health MarketScan® claims data. For the Truven data, stratified analyses by gender and age and year were also performed using SAS Enterprise Guide v7.1. Binomial proportion confidence intervals were calculated using the Wilson score interval [28]. Annual incidence and prevalence rates are presented as both crude and age-adjusted rates to allow examination of trends over time. For both incidence and prevalence analyses, all estimates are reported for the overall group, and stratified by age and by gender.

The USRDS data were accessed through the USRDS Renal Data Extraction and Referencing (RenDER) System, which is an online data querying application accessible through the USRDS website, allowing access to data on ESRD in the USA [25]. Prevalence rates were calculated using numerator data from the RenDER system and US population estimates from the US Census Bureau. The prevalence of ADPKD with ESRD was used to estimate the annual US prevalence using a method described in a previous publication [19]. This method applies an inflation factor derived from the ratio of total ADPKD prevalence to ADPKD with ESRD prevalence based on consistent results from population-based screening studies [29, 30].

#### Sensitivity Analyses

The prevalence of *diagnosed* ADPKD has not been previously estimated using a national sample of patient level, insurance claims data. However, the validity of our method is supported by a recent study that found ICD-10 coding in administrative data could identify adult patients who meet strict clinical criteria for ADPKD as assessed through medical chart review [31]. A sensitivity analyses was performed to compare algorithms for detecting ADPKD in large populations. For the sensitivity analysis, the Truven Health MarketScan® data claims data were analyzed using the following 3 approaches: (1) identifying patients based on a single ADPKD claim per year, (2) using a more stringent requirement of 2 or more claims per year to rule out coding errors or patients for whom the code may have been used for a rule-out diagnosis, and c) applying the additional criteria of an annual renal imaging study or specialist visit, with a buffer of 2 months at the beginning and end of the year.

Specialist visits were defined as a visit to either a nephrologist or a urologist. CPT codes to identify renal imaging included: 74181, MRI Abdomen W/O Contrast; 74183, MRI Abdomen W/O Contrast; 74178, CT Abd & Pelvis W W/O Contrast; 74177, CT Abd & Pelvis W/Contrast; 74176, CT Abd & Pelvis W/O Contrast; 74160, CT Abd W/ Contrast; 74150, CT Abd W/O Contrast; 74170, CT Abd W W/O Contrast; 76775, Ultrasound, retroperitoneal (e.g., renal, aorta, and nodes), real time with image documentation; and 76776, Ultrasound, transplanted kidney, real-time and duplex Doppler with image documentation.

## Results

### *Incidence and Prevalence from the Literature*

The literature review yielded 1,243 citations which were reviewed to select 3 population-based, US epidemiologic studies of ADPKD. Criteria used to select relevant studies included: national and regional population-based



studies, and adequate sampling or inclusion of power analyses. Studies were excluded from consideration based on the following criteria: studies of populations outside the USA, laboratory-based versus population-based studies, small-sample studies that focused on specific subgroups of ADPKD patients, and genetic studies of individual families. Of the studies that were excluded, most focused on small convenience samples of ADPKD patients that were not generalizable to a regional or national population. The resulting population-based studies [32–34] are described in detail in the Appendix.

Although the resulting studies provide valid epidemiologic data, the restrictive definitions for case inclusion yield information on only a limited subset of ADPKD patients in the USA. For example, 2 studies provided ADPKD incidence estimates that are limited to patients with advanced disease. Reule et al. [34] reported an annual incidence rate of 7.5 per million cases of ESRD with PKD as a primary diagnosis, and O'Shaughnessy et al. [33] reported the proportion of ESRD cases in the USA diagnosed with PKD (2.4%). Although both studies suggest that ADPKD is uncommon in the USA, a true population-based incidence rate that includes all diagnosed patients cannot be determined from the data presented.

The remaining study [32] also focused on patients with advanced disease, defined as prevalent cases of ADPKD who received inpatient care in 390 hospitals in the USA. Inpatient resource utilization was the primary focus of this study, and the data presented provide key data on healthcare utilization patterns in ADPKD patients. Despite the large, population-based sample, the prevalence of hospitalized cases cannot be determined because hospital catchment areas do not align with regional population data. Thus, this and other recently published US studies have limited utility for the descriptive epidemiology of ADPKD because they are not inclusive of the general population of patients and focus primarily on patients with more severe disease.

#### *Prevalence from National Data Sources*

ADPKD prevalence calculations from national data sources are summarized in the Appendix.

The average annual prevalence of diagnosed ADPKD in the USA from the NAMCS is 4.3 per 10,000. The annual prevalence of ADPKD in the USA was also found to be 4.3 per 10,000 using data from the USRDS, with adjustment for the proportion of ADPKD patients with ESRD from population-based screening studies in Europe [19, 29, 30]. The prevalence of hospitalized cases of ADPKD in the USA is much lower, at 0.56 per 100,000, as deter-

mined from the Agency for Healthcare Research and Quality, National Inpatient Sample [35].

#### *Incidence and Prevalence from Claims Data*

The annual age-adjusted prevalence of diagnosed ADPKD among commercially insured patients increased from 1.74 per 10,000 in 2013 to 2.10 per 10,000 in 2015, as shown in Table 1. A similar increase in age-adjusted diagnosed prevalence was evident for Managed Medicaid claims data from 2013 to 2014 (2.26 to 2.40 per 10,000), but prevalence rates then stabilized from 2014 to 2015 at 2.20 per 10,000. The average annual age-adjusted diagnosed prevalence of ADPKD in commercially insured patients was 2.06 per 10,000 and 2.28 per 10,000 in the Medicaid sample.

Large differences in diagnosed prevalence rates were evident by age (Table 1). Among commercially insured patients, prevalence is lowest in the youngest and oldest age groups, and it peaks between the ages of 50–64 years (3.29/10,000), consistent with the average age of ESRD [7]. Although a similar pattern was evident in the Managed Medicaid sample, peak prevalence occurred at a slightly younger age, i.e., among those aged 35–49 years (3.8 per 10,000). The prevalence in the Managed Medicaid sample declined after the age range of 35–49 years, as compared to after ages 50–64 years in the commercially insured group, suggesting that survival time might be shorter in this patient group. Analyses by year demonstrate that the age-adjusted ADPKD-diagnosed prevalence increased in the much larger, commercially insured population over time from 1.74 in 2013 to 2.10 in 2015.

The diagnosed prevalence also varied by gender (Table 2) and was higher among females as compared to males (2.18 vs. 1.94 per 10,000) in commercially insured patients. This difference is statistically significant, as the confidence intervals do not overlap (Table 1). Prevalence was also higher among females in the Medicaid sample (1.56 per 10,000 vs. 1.22 among males) (Table 1). Peak prevalence occurred in females aged 50–64 years, whereas peak prevalence occurred at older ages among males (65 years or older). These gender differences are consistent over time, as shown in analyses by year in Table 3.

We found striking age- and gender-related differences in ADPKD incidence and diagnosed prevalence (Table 2); females are much more likely to be diagnosed with ADPKD at an earlier age than males. The incidence within the 20- to 34-year age group was 6.19 for females, versus 3.52 per 100,000 for males. Males are more likely to be diagnosed at older ages; the incidence is 18% higher in males aged 50–64 years than in females of the same age and 80% higher in males aged 65 years or older than in older females.

**Table 1.** ADPKD annual diagnosed prevalence, and average prevalence by age and by gender for 2013–2015, Truven Health MarketScan® Claims and Encounters Database

	Commercial and Medicare Database (age 0–65 years or older)				Managed Medicaid Database (age 0–64 years)			
	sample size, <i>n</i>	crude rate per 10,000	age-adjusted rate per 10,000	95% CI (adjusted rate)	sample size, <i>n</i>	crude rate per 10,000	age-adjusted rate per 10,000	95% CI (adjusted rate)
Year								
2013	34,235,044	1.88	1.74	1.73–1.76	4,721,746	1.25	2.26	2.24–2.28
2014	35,809,429	2.12	1.97	1.96–1.99	7,067,028	1.46	2.40	2.38–2.42
2015	22,323,496	2.24	2.10	2.08–2.12	7,688,020	1.46	2.20	2.18–2.22
Age (years)			Prevalence OR				Prevalence OR	
0–19	21,828,302	0.39	1.00	0.37–0.42	12,434,819	0.51	1.00	0.48–0.56
20–34	17,445,743	1.29	3.30	1.24–1.35	2,774,436	1.86	2.95	1.70–2.02
35–49	20,584,592	2.73	6.96	2.66–2.80	2,096,597	3.80	5.70	3.54–4.07
50–64	23,866,372	3.29	8.40	3.22–3.37	2,170,942	3.63	5.89	3.39–3.90
65+	8,642,960	2.82	7.20	2.71–2.94	NA	NA		
Total	92,367,969	2.06	$p < 0.001$	2.03–2.09	19,476,794	1.41	$p < 0.001$	1.36–1.46
Gender			Prevalence OR				Prevalence OR	
Male	44,431,464	1.94	0.89	1.90–1.98	8,826,646	1.22	0.85	1.15–1.30
Female	47,936,505	2.18	1.00	2.13–2.22	10,650,148	1.56	1.00	1.49–1.64
			$p < 0.001$				$p < 0.001$	

**Table 2.** Average annual incidence and diagnosed prevalence of ADPKD by age and gender, Truven Health MarketScan® Commercial and Medicare Database, 2013–2015

Age, years	Females			Males		
	average annual sample size, <i>n</i>	average annual prevalence per 10,000	average annual incidence per 100,000 (2014–2015)	average annual sample size, <i>n</i>	average annual prevalence per 10,000	average annual incidence per 100,000 (2014–2015)
0–19	3,561,504	0.39	1.82	3,714,596	0.39	1.48
20–34	2,986,548	1.60	6.19	2,828,699	0.97	3.52
35–49	3,620,953	2.89	7.51	3,240,577	2.55	7.22
50–64	4,242,049	3.40	7.60	3,713,408	3.17	9.02
65+	1,567,780	2.36	7.45	1,313,207	3.37	13.46
Total	15,978,835	2.18	6.09	14,810,488	1.94	6.27
<i>p</i>		<0.001	<0.001		<0.001	<0.001

### Sensitivity Analyses

The proportion of identified cases using additional criteria of either an annual specialist visit or renal imaging study are presented in Table 4. When commercial claims or Medicare patients were identified using only 1 annual claim with an ADPKD diagnosis, 73% were found to have had either a specialist visit or a renal imaging study during the same year ( $\pm 2$  months). This proportion increased to 80% when patients were identified

based on 2 or more annual claims with an ADPKD diagnosis. Within the Medicaid population, these proportions were lower; for those identified with one annual diagnostic claim, the proportion was 58%, and for those identified with 2 or more annual diagnostic claims, the proportion increased to 65%. In both samples, the use of 2 or more annual codes is associated with higher retention in the sensitivity analyses.

**Table 3.** ADPKD diagnosed prevalence by year, age, and gender, Truven Health MarketScan® Claims and Encounters Database

Age, years	Commercial and Medicare Database (age 0–65 years or older)				Managed Medicaid Database (age 0–64 years)			
	females		males		females		males	
	crude rate/10,000	OR	crude rate/10,000	OR	crude rate/10,000	OR	crude rate/10,000	OR
<b>2013</b>								
0–19	0.34	1.0	0.36	1.0	0.42	1.0	0.46	1.0
20–34	1.48	4.42	0.82	2.30	1.84	4.37	1.90	4.13
35–49	2.62	7.83	2.34	6.60	3.84	9.13	3.62	7.88
50–64	3.15	9.41	2.94	8.27	3.98	9.46	3.28	7.15
65 or older	2.15	6.40	3.07	8.63	NA		NA	
Total	1.97	$p < 0.001^a$	1.64	$p < 0.001$	1.41	$p < 0.001$	1.05	$p < 0.001$
<b>2014</b>								
0–19	0.42	1.0	0.40	1.0	0.53	1.0	0.53	1.0
20–34	1.70	4.01	1.07	2.70	1.94	3.64	1.70	3.19
35–49	2.92	6.99	2.60	6.57	3.68	6.93	4.54	8.55
50–64	3.45	8.15	3.27	8.25	4.16	7.83	3.65	6.87
65 or older	2.51	5.93	3.51	8.86	NA		NA	
Total	2.21	$p < 0.001$	1.86	$p < 0.001$	1.60	$p < 0.001$	1.29	$p < 0.001$
<b>2015</b>								
0–19	0.43	1.0	0.44	1.0	0.56	1.0	0.54	1.0
20–34	1.63	3.75	1.02	2.33	1.90	3.38	1.70	3.16
35–49	3.24	7.46	2.79	6.37	3.58	6.36	3.79	7.05
50–64	3.69	8.51	3.36	7.68	3.59	6.38	3.09	5.74
65 or older	2.60	5.99	4.16	9.49	NA		NA	
Total	2.35	$p < 0.001$	1.93	$p < 0.001$	1.62	$p < 0.001$	1.27	$p < 0.001$

Values are based on continuously enrolled patients with 2 or more annual claims with an ICD9 or ICD10 code for ADPKD or unspecified PKD (ICD-9-CM: 753.12 or 753.13, ICD10: Q61.2, Q61.3). <sup>a</sup> Extended Mantel-Haenszel  $\chi^2$  for linear trend.

## Discussion

Our analyses of the NAMCS and patient claims data provide the first population-based US epidemiologic data on ADPKD since the publication of incidence data collected from 1935 to 1980 in Olmsted County (MN, USA) [8]. Our use of both the NAMCS and patient claims data allowed us to generate an overall diagnosed prevalence estimate that is generalizable to the US, and to provide insight into age and gender differences in ADPKD incidence and prevalence using a large, nationwide sample.

To our knowledge, patient claims data have not previously been used to estimate the diagnosed prevalence of ADPKD in the USA. Our analyses provide support for this method because of the consistent results between claims data and other sources of data, along with supporting sensitivity analyses. The diagnosed prevalence of ADPKD for

the US population was found to be 4.3 per 10,000 in the NAMCS survey sample, and age-adjusted prevalence rates in the commercial and Medicaid population were 2.06 and 2.28 per 10,000. The results show a close correspondence of 4.34 in the combined commercial and Medicaid populations versus 4.3 per 10,000 from national survey data.

Our US diagnosed prevalence estimate is consistent with our recently published prevalence estimate for Europe which was based on patient counts from national and regional registries, covering 19 EU countries and 81% of the EU population [19]. This approach yielded an ADPKD diagnosed point prevalence on December 31, 2012, of 3.96 per 10,000. Other recently published European studies [17, 36, 37] have yielded similar prevalence estimates ranging from 3.3 to 4.6 per 10,000. Also consistent with these results, a recent European meta-analysis of 8 epidemiologic studies yielded an average prevalence of

**Table 4.** Comparison between primary and sensitivity analyses, Truven Health MarketScan® Claims and Encounters Database for 2014

	Sensitivity criteria		
	cases, <i>n</i> (crude rate/10,000)	patients with a specialist <sup>a</sup> visit or procedure code for renal imaging, <i>n</i> (crude rate/10,000) <sup>b</sup>	proportion of cases retained in the sensitivity analysis, %
<i>Commercial and Medicare Database (ages 0–65+; n = 35,809,429)</i>			
Patients identified as cases with at least 1 ICD-9 claim for ADPKD	12,728 (3.56)	9,300 (2.6)	73.1
Patients identified as cases with at least 2 ICD-9 claims for ADPKD	7,580 (2.04)	6,086 (1.7)	80.3
<i>Managed Medicaid patients, age 0–64 years (n = 7,067,028)</i>			
Patients identified as cases with at least 1 ICD-9 claim for ADPKD	1,774 (2.51)	1,020 (1.44)	57.5
Patients identified as cases with at least 2 ICD-9 claims for ADPKD	1,029 (2.40)	664 (0.94)	64.5

CPT codes for renal imaging included: 74181, MRI Abdomen W/O Contrast; 74183, MRI Abdomen W W/O Contrast; 74178, CT Abd & Pelvis W W/O Contrast; 74177, CT Abd & Pelvis W/Contrast; 74176, CT Abd & Pelvis W/O Contrast; 74160, CT Abd W/ Contrast; 74150, CT Abd W/O Contrast; 74170, CT Abd W W/O Contrast; 76775, Ultrasound, retroperitoneal (e.g., renal, aorta, and nodes), real time with image documentation; 76776, Ultrasound, transplanted kidney, real-time and duplex Doppler with image documentation. The time frame was defined as either during 2014 or within 2 months of the beginning or end of the year.

<sup>a</sup> Defined as: nephrologist or urologist.

2.7 per 10,000 (95% CI = 0.73–4.67) and a point prevalence of 3.63 per 10,000 in the province of Modena (Italy) [38].

All recent prevalence estimates for the USA and Europe are much lower than a frequently cited prevalence rate of 1 in 400–1,000 [39] from a study published in 1957. Based on this early study, one might predict that as many as 400,000–700,000 Americans are affected with ADPKD, while our nationwide, population-based data indicate that only 140,000 patients are diagnosed with ADPKD. The discrepancy is related to the following 2 factors: (1) the rate from this early study, though often cited as “prevalence,” is actually a measure of lifetime morbid risk (LMR) which is not comparable with point or period prevalence rates, and (2) there is potential for overestimation and misclassification in this early study as cases were identified from autopsy notes and the card indices of radiology departments and private clinics during 1920 to 1953, before modern diagnostic imaging methods and genetic testing were available to rule out other cystic kidney diseases. Further epidemiologic study using survival analysis methods that adjust for competing risk is needed to refine and update this early, and often misunderstood, estimate of the LMR of ADPKD.

Finally, additional support for our prevalence estimates comes from the USRDS data set. Population-based

screening in some regions of Europe suggests that approximately 23% of diagnosed ADPKD cases are also diagnosed with ESRD [29, 30]. We applied this inflation factor to the US prevalence of ADPKD patients with ESRD and the resulting indirect prevalence estimate is nearly identical to the data-based estimate derived from the NAMCS (4.30 vs. 4.34 per 10,000). Thus, 3 different nationwide data sets, analyzed with different methods, consistently indicate that approximately 140,000 patients were diagnosed with ADPKD in the USA in 2017.

Our data is limited to diagnosed patients, and without widespread screening to detect asymptomatic patients, the total diagnosed and undiagnosed prevalence of ADPKD cannot be determined. However, it is possible to theoretically estimate the additional number of undiagnosed patients based on the results of recent European studies [17, 29]. These studies indicate that approximately 15% more cases might be detected after increased awareness and intensified population based screening [17, 19, 29]. We estimate that similar screening efforts in the USA could yield a total of 166,003 patients, with a screening prevalence of 5.08 per 10,000.

Although the prevalence of the PKD1/2 genotype in the population is certainly higher than our diagnosed prevalence measure, comparisons between our directly



measured prevalence and theoretical prevalence measures should be considered cautiously. For the purpose of understanding unmet needs, comparing our diagnosed, point prevalence estimate of 4.3 per 10,000 with theoretical estimates such as the LMR of Dalgaard [39] of 8 per 10,000 will overestimate the number of undiagnosed patients. The LMR applies to a hypothetical population where everyone survives to age 80 years and will overestimate the lifetime prevalence in living populations where the actual survival time is shorter. In addition, recent research suggests considerable phenotypic variability in families with ADPKD, with the identification of non- and incompletely penetrant alleles of PKD1 that produce atypical, mild, or severe disease depending on whether they are inherited in the heterozygous or homozygous state [40]. Population-based genomic data are needed to better elucidate the true prevalence of ADPKD and quantify unmet treatment needs. One limitation of our study is the reliance upon insurance claims to identify diagnosed ADPKD patients. For this reason, we compared our claims-based prevalence estimate to that obtained from a nationally representative physician survey of ambulatory care patients in the USA. The close correspondence between the 2 prevalence estimates supports the validity of claims data for epidemiologic studies of ADPKD. Other limitations are that claims databases are characterized by a limited duration of continuous enrollment and follow-up periods, and this type of data does not allow review of clinical information that may have occurred under a different health plan. This limitation is more likely to affect incidence versus prevalence estimates as these rely on identification of the first diagnosis of the condition. Future studies of ADPKD incidence with longer initial disease-free periods are needed.

Perhaps the most striking findings are significant differences in the incidence and prevalence of ADPKD by gender and age. As ADPKD is caused by mutations in either PKD1 or PKD2 genes, and has complete penetrance [10], all affected patients have the potential for diagnosis early in life. However, females are diagnosed during early adulthood at a much higher rate than males; the incidence of diagnosed ADPKD was 75% higher in females aged 20–34 years than in males. This difference may be related to the use of abdominal ultrasound during the first trimester of pregnancy, as this type of diagnostic imaging is also used to detect ADPKD.

To further explore the idea that the higher rate of diagnosis of ADPKD in young adult women may be related to ultrasound examinations during pregnancy, we compared the diagnosed prevalence of ADPKD in pregnant

and nonpregnant women in the commercial insurance population. Among all women aged 20–34 years ( $n = 8,959,645$ ), 1,046,007 had 1 or more ICD9 codes related to pregnancy<sup>1</sup>. The rate of ADPKD among nonpregnant women ( $n = 7,913,638$ ) was 1.49 per 10,000, while this rate was 2.41 per 10,000 among pregnant women (Fisher's exact test,  $p < 0.001$ ). These results provide support to the notion that ultrasound during pregnancy may be a pathway to the diagnosis of ADPKD in young adult women. Further research on this topic with a focus on incident cases and the timing of ultrasound exams is needed for a better understanding of how diagnostic imaging during pregnancy may impact the detection of ADPKD. This information would also provide insight into the extent to which ADPKD is underdiagnosed in the general population. Overall, we observed a pattern of later ADPKD diagnosis in males. The high incidence rate for men aged 65 years or older suggests that ADPKD may not be detected in many men until after the progression of chronic kidney disease. Delayed diagnosis in men may be related to the same factors that have been shown to delay the diagnosis of other chronic diseases, such as fewer preventive care doctor visits and delay in reporting symptoms [41–43]. The difference may also be related to gender differences in the experience of ADPKD-related pain; in a large study of quality of life in ADPKD patients, “females had more pain and reported greater interference with daily life than males for almost all of the pain questionnaire items” [44]. Thus, higher levels of pain in females may lead to earlier care seeking, detection, and diagnosis.

These age- and gender-related differences raise several questions for future research. Pregnancy screening for women and late-life development of chronic disease in men may serve as gateways to ADPKD diagnosis, but longitudinal research is needed to further explore these hypotheses. Our data also point to a need for more information on the relationship between earlier diagnosis and ADPKD progression and outcomes. For example, studies have shown an inverse relationship between the intensity of antihypertensive treatment and mortality in ADPKD [13, 37], but little is known about whether earlier diagnosis, such as during pregnancy, is protective against disease progression. In addition, several studies have reported more rapid disease progression in men, including more severe hypertension and earlier onset of ESRD [45–47].

<sup>1</sup> V22.x, supervision of normal pregnancy; V23.x, supervision of high-risk pregnancy; 651.03, twin pregnancy, antepartum condition, or complication; V91.00, twin gestation, monochorionic/monoamniotic; 642.63, eclampsia, antepartum condition, or complication; or 633, ectopic pregnancy.



Our study raises the question of whether later disease detection in men accounts for some of the observed differences in disease severity. More research on gender differences in the diagnosis, management, treatment, and progression of ADPKD is clearly needed.

## Conclusion

This study addresses an important gap in US ADPKD epidemiology, as it is the first nationwide epidemiologic study of ADPKD incidence and diagnosed prevalence. Several key points emerged from our findings. ADPKD

affects approximately 140,000 patients in the USA, which fulfills the criterion for a rare disease. We found significant differences in diagnosis and detection of ADPKD by gender and age. Females are nearly twice as likely to be diagnosed in early adulthood compared to males, perhaps due to the use of abdominal ultrasound during child-bearing years. Among males, the incidence was highest in the age group of 65 years or older, suggesting that ADPKD may not be detected in many men until after chronic kidney disease is diagnosed. Our results indicate a need for increased awareness of ADPKD and further study of gender and ADPKD progression, management, and outcomes.

## Appendix

Reference	Data source	Study population	Incidence estimates	Prevalence estimates	Epidemiologic strengths/limitations
Blanchette et al. [32]	MedAssets Health System Data of de-identified patient level claims for 390 hospitals in the USA	Hospital discharges between January 1, 2010, and December 31, 2012, with an ICD9 code of 753.12 or 753.13	NA	859/14,349,893 = 0.6 per 10,000 hospital admissions	Prevalence in the population could not be determined, as denominator data for catchment areas were unavailable
O'Shaughnessy et al. [33]	USRDS	All adult patients with PKD who initiated ESRD therapy between January 1, 1996, and December 31, 2011	$n = 36,272$ with a primary cause of ADPKD, % of adult ESRD = $36,272 / 1,536,217 = 2.4\%$	NA	Incidence rate is limited to patients with ESRD
Reule et al. [34]	USRDS	US PKD patients who initiated RRT (2001–2002) ( $n = 1,069,343$ )	1-year cumulative incidence = $2,141 / 286,297,074 = 7.5 / 1,000,000$	NA	Incidence rate is limited to patients with ESRD
<i>US Publicly Available National Databases</i>					
Centers for Disease Control National Center for Health Statistics	NAMCS (2012–2014)	Nationally representative sample of ambulatory care visits in the USA	NA	137,462/318,748,017 = 4.32/10,000	Generalizable to the US population; the number of diagnosed US patients in 2017 is estimated at 141,285
Agency for Healthcare Quality and Research	Healthcare Costs and Utilization (HCUP) Nationwide Inpatient Sample (2014)	Nationwide sample of ~20% of US hospital discharges	NA	1,785 discharges/318,748,017 = 0.56 hospitalizations per 100,000	Prevalence rate is limited to ADPKD patients in the USA with an inpatient stay during 2014
National Institutes of Health, National Institute of Diabetes and Digestive and Kidney Diseases	USRDS	National data registry on the ESRD population in the USA	2014 incidence of ESRD with PKD = $2,544 / 318,748,017 = 0.8$ per 100,000	ESRD with PKD = $31,544 / 318,748,017 = 9.89 / 100,000$ ; estimated US ADPKD prevalence = $4.30 / 10,000^a$	Direct prevalence calculation is limited to patients with ESRD; estimated prevalence consistent with NAMCS data; number of patients diagnosed in 2017 = 140,537

<sup>a</sup> Calculated from the ESRD prevalence rate using data on the proportion of ADPKD with ESRD from population-based screening studies (see Methods for additional details).

## Disclosure Statement

During the past 12 months, Dr. Willey has served as an epidemiologic consultant to Otsuka Pharmaceutical Development and Commercialization, Inc. Robert Stellhorn is a contractor for Otsuka Pharmaceutical Development and Commercialization, Inc. Siddhesh Kamat and Jaime Blais were employees of Otsuka Pharmaceutical Development and Commercialization, Inc. at the time the publication was submitted.

## Funding Source

Funding for statistical analyses of Truven Health MarketScan® and other databases and study oversight was provided by Otsuka Pharmaceutical Development and Commercialization, Inc.

## References

- Levy M, Feingold J. Estimating prevalence in single-gene kidney diseases progressing to renal failure. *Kidney Int.* 2000 Sep;58(3):925–43.
- Saran R, Robinson B, Abbott KC, Agodoa LY, Bhavé N, Bragg-Gresham J, et al. US Renal Data System 2017 Annual Data Report: Epidemiology of Kidney Disease in the United States. *Am J Kidney Dis.* 2018 Mar;71(3 3S1):A7.
- Blanchette CM, Iorga SR, Altan A, Seare JG, Fan Y, Rossetti S, et al. Healthcare resource utilization and costs associated with autosomal dominant polycystic kidney disease. *J Health Econ.* 2014;2:63–74.
- Lentine KL, Xiao H, Machnicki G, Gheorghian A, Schnitzler MA. Renal function and healthcare costs in patients with polycystic kidney disease. *Clin J Am Soc Nephrol.* 2010 Aug;5(8):1471–9.
- McEwan P, Bennett H, O'Reilly K, Robinson P. Quantifying The Impact of Treatment on The Public Health Burden of Adpdk: A UK Case Study Using The Adpdk Outcomes Model. *Value Health.* 2015 Nov;18(7):A763.
- Mennini FS, Marcellusi A, Russo S, Iorio A, Lanati EP, Robinson P. Adapting The Adpdk Outcomes Model to Predict Cost Consequence In Italian Patients With Autosomal Dominant Polycystic Kidney Disease (Adpdk) Treated With Jinarc (Tolvaptan). *Value Health.* 2015 Nov;18(7):A755.
- Saran R, Robinson B, Abbott KC, Agodoa LY, Albertus P, Ayanian J, et al. US Renal Data System 2016 Annual Data Report: Epidemiology of Kidney Disease in the United States. *Am J Kidney Dis.* 2017 Mar;69(3 Suppl 1):A7–8.
- Iglesias CG, Torres VE, Offord KP, Holley KE, Beard CM, Kurland LT. Epidemiology of adult polycystic kidney disease, Olmsted County, Minnesota: 1935–1980. *Am J Kidney Dis.* 1983 May;2(6):630–9.
- Blanchette CM, Liang C, Lubeck DP, Newsome B, Rossetti S, Gu X, et al. Progression of autosomal dominant kidney disease: measurement of the stage transitions of chronic kidney disease. *Drugs Context.* 2015 Apr;4: 212275.
- Grantham J, Nair V, Winklhofer F. Cystic diseases of the kidney. In: Bm B, editor. *Brenner & Rector's The Kidney*. Volume 2. Philadelphia (PA): WB Saunders Company; 2000. pp. 1699–730.
- King BF, Reed JE, Bergstralh EJ, Sheedy PF 2nd, Torres VE. Quantification and longitudinal trends of kidney, renal cyst, and renal parenchyma volumes in autosomal dominant polycystic kidney disease. *J Am Soc Nephrol.* 2000 Aug;11(8):1505–11.
- Orskov B, Rømming Sørensen V, Feldt-Rasmussen B, Strandgaard S. Improved prognosis in patients with autosomal dominant polycystic kidney disease in Denmark. *Clin J Am Soc Nephrol.* 2010 Nov;5(11):2034–9.
- Perrone R, Miskulin DC. Making an earlier diagnosis of ADPKD: implications for the treatment of hypertension. *Nephrol News Issues.* 2006 Jun;20(7):32, 35–36.
- Schrier RW, McFann KK, Johnson AM. Epidemiological study of kidney survival in autosomal dominant polycystic kidney disease. *Kidney Int.* 2003 Feb;63(2):678–85.
- Cornec-Le Gall E, Audrézet MP, Renaudineau E, Hourmant M, Charasse C, Michez E, et al. PKD2-Related Autosomal Dominant Polycystic Kidney Disease: Prevalence, Clinical Presentation, Mutation Spectrum, and Prognosis. *Am J Kidney Dis.* 2017 Oct; 70(4):476–85.
- de Almeida E, Sousa A, Pires C, Aniceto J, Barros S, Prata MM. Prevalence of autosomal-dominant polycystic kidney disease in Alentejo, Portugal. *Kidney Int.* 2001 Jun;59(6): 2374.
- Neumann HP, Jilg C, Bacher J, Nabulsi Z, Malinoc A, Hummel B, et al. Else-Kroener-Fresenius-ADPKD-Registry. Epidemiology of autosomal-dominant polycystic kidney disease: an in-depth clinical study for southwestern Germany. *Nephrol Dial Transplant.* 2013 Jun;28(6):1472–87.
- Neumann ME. The impending burden of kidney disease. *Nephrol News Issues.* 2015 Apr; 29(4):8.
- Willey CJ, Blais JD, Hall AK, Krasa HB, Makin AJ, Czerwiec FS. Prevalence of autosomal dominant polycystic kidney disease in the European Union. *Nephrol Dial Transplant.* 2017 Aug;32(8):1356–63.
- Morrison PJ, Harding-Lester S, Bradley A. Uptake of Huntington disease predictive testing in a complete population. *Clin Genet.* 2011 Sep;80(3):281–6.
- Tassicker RJ, Teltscher B, Trembath MK, Collins V, Sheffield LJ, Chiu E, et al. Problems assessing uptake of Huntington disease predictive testing and a proposed solution. *Eur J Hum Genet.* 2009 Jan;17(1):66–70.
- Cosmatos I, Matcho A, Weinstein R, Montgomery MO, Stang P. Analysis of patient claims data to determine the prevalence of hidradenitis suppurativa in the United States. *J Am Acad Dermatol.* 2013 Nov;69(5):819.
- Cragin LA, Laney AS, Lohff CJ, Martin B, Pandiani JA, Blevins LZ. Use of insurance claims data to determine prevalence and confirm a cluster of sarcoidosis cases in Vermont. *Public Health Rep.* 2009 May-Jun;124(3): 442–6.
- Nelson CA, Saha S, Kugeler KJ, Delorey MJ, Shankar MB, Hinckley AF, et al. Incidence of Clinician-Diagnosed Lyme Disease, United States, 2005–2010. *Emerg Infect Dis.* 2015 Sep; 21(9):1625–31.
- United States Renal Data System: USRDS RenDER, 2017.
- US Census Bureau. NP2014\_D1: Projected population by single year of age, sex, race, and Hispanic origin for the United States – 2014 to 2060. Suitland: US Census Bureau; 2014.
- Burt CW, Hing E. Making patient-level estimates from medical encounter records using a multiplicity estimator. *Stat Med.* 2007 Apr; 26(8):1762–74.
- Brown LD, DasGupta A. Interval estimation for a proportion. *Stat Sci.* 2001;16(2):101–17.
- Davies F, Coles GA, Harper PS, Williams AJ, Evans C, Cochlin D. Polycystic kidney disease re-evaluated: a population-based study. *Q J Med.* 1991 Jun;79(290):477–85.
- Heidland A, Bahner U, Deetjen A, Götz R, Heidbreder E, Schäfer R, et al. Mass-screening for early detection of renal disease: benefits and limitations of self-testing for proteinuria. *J Nephrol.* 2009 Mar-Apr;22(2):249–54.

- 31 Kalatharan V, Pei Y, Clemens KK, McTavish RK, Dixon SN, Rochon M, et al. Positive predictive values of International Classification of Diseases, 10th Revision coding algorithms to identify patients with autosomal dominant polycystic kidney disease. *Can J Kidney Health Dis.* 2016 Dec;3:2054358116679130.
- 32 Blanchette CM, Craver C, Belk KW, Lubeck DP, Rossetti S, Gutierrez B. Hospital-based inpatient resource utilization associated with autosomal dominant polycystic kidney disease in the US. *J Med Econ.* 2015 Apr;18(4):303–11.
- 33 O'Shaughnessy MM, Montez-Rath ME, Lafayette RA, Winkelmayer WC. Patient characteristics and outcomes by GN subtype in ESRD. *Clin J Am Soc Nephrol.* 2015 Jul;10(7):1170–8.
- 34 Reule S, Sexton DJ, Solid CA, Chen SC, Collins AJ, Foley RN. ESRD from autosomal dominant polycystic kidney disease in the United States, 2001–2010. *Am J Kidney Dis.* 2014 Oct;64(4):592–9.
- 35 Agency for Healthcare Research and Quality: HCUPnet. Rockville: Healthcare Cost and Utilization Project; 2017.
- 36 McGovern AP, Jones S, van Vlymen J, Saggar AK, Sandford R, de Lusignan S. Identification of people with autosomal dominant polycystic kidney disease using routine data: a cross sectional study. *BMC Nephrol.* 2014 Nov;15(1):182.
- 37 Patch C, Charlton J, Roderick PJ, Gulliford MC. Use of antihypertensive medications and mortality of patients with autosomal dominant polycystic kidney disease: a population-based study. *Am J Kidney Dis.* 2011 Jun;57(6):856–62.
- 38 Solazzo A, Testa F, Giovannella S, Busutti M, Furci L, Carrera P, et al. The prevalence of autosomal dominant polycystic kidney disease (ADPKD): A meta-analysis of European literature and prevalence evaluation in the Italian province of Modena suggest that ADPKD is a rare and underdiagnosed condition. *PLoS One.* 2018 Jan;13(1):e0190430.
- 39 Dalgaard OZ. Bilateral polycystic disease of the kidneys; a follow-up of two hundred and eighty-four patients and their families. *Acta Med Scand Suppl.* 1957;328:1–255.
- 40 Sandford RN. The diversity of PKD1 alleles: implications for disease pathogenesis and genetic counseling. *Kidney Int.* 2009 Apr;75(8):765–7.
- 41 Courtenay WH. Constructions of masculinity and their influence on men's well-being: a theory of gender and health. *Soc Sci Med.* 2000 May;50(10):1385–401.
- 42 Courtenay WH, Keeling RP. Men, gender, and health: toward an interdisciplinary approach. *J Am Coll Health.* 2000 May;48(6):243–6.
- 43 Smith JA, Braunack-Mayer A, Wittert G. What do we know about men's help-seeking and health service use? *Med J Aust.* 2006 Jan;184(2):81–3.
- 44 Miskulin DC, Abebe KZ, Chapman AB, Perrone RD, Steinman TI, Torres VE, et al.; HALT-PKD Study. Health-related quality of life in patients with autosomal dominant polycystic kidney disease and CKD stages 1–4: a cross-sectional study. *Am J Kidney Dis.* 2014 Feb;63(2):214–26.
- 45 Gabow PA, Johnson AM, Kaehny WD, Kimberling WJ, Lezotte DC, Duley IT, et al. Factors affecting the progression of renal disease in autosomal-dominant polycystic kidney disease. *Kidney Int.* 1992 May;41(5):1311–9.
- 46 Johnson AM, Gabow PA. Identification of patients with autosomal dominant polycystic kidney disease at highest risk for end-stage renal disease. *J Am Soc Nephrol.* 1997 Oct;8(10):1560–7.
- 47 Schrier RW, Brosnahan G, Cadnapaphornchai MA, Chonchol M, Friend K, Gitomer B, et al. Predictors of autosomal dominant polycystic kidney disease progression. *J Am Soc Nephrol.* 2014 Nov;25(11):2399–418.