The National Economic Burden of Rare Disease in United States in 2019

Grace Yang
   The Lewin Group   https://orcid.org/0000-0002-1184-2305

Inna Cintina
   The Lewin Group

Anne Pariser
   National Institutes of Health

Elisabeth Oehrlein
   National Health Council

Jamie Sullivan
   EveryLife Foundation for Rare Diseases

Annie Kennedy (✉ akennedy@everylifefoundation.org)
   EveryLife Foundation for Rare Diseases   https://orcid.org/0000-0002-0987-675X

Abstract

Background: To provide a comprehensive assessment of the total economic burden of rare diseases (RD) in the U.S. in 2019.

We followed a prevalence-based approach that combined the prevalence of 379 RDs with the per-capita direct medical and indirect costs, to derive the national economic burden by patient age and type of RD. To estimate prevalence and the direct medical cost of RD, we used claims data from three sources: Medicare 5% Standard Analytical File, Transformed Medicaid Statistical Information System, and Optum claims data for the privately insured. To estimate indirect and non-medical cost components, we worked with the rare disease community to design and implement a primary survey.

Results: There were an estimated 15.5 million U.S. children (N=1,322,886) and adults (N=14,222,299) with any of the 379 RDs in 2019 with a total economic burden of $997 billion, including a direct medical cost of $449 billion (45%), $437 billion (44%) in indirect costs, and $111 billion (11%) in non-medical costs.

The top drivers for excess medical costs associated with RD are hospital inpatient care and prescription medication; the top indirect cost categories are labor market productivity losses due to absenteeism, presenteeism, and forced early retirement.

Conclusions: Our findings highlight the scale of the RD economic burden and call for immediate attention from the scientific communities, policy leaders, and other key stakeholders such as health care providers and employers, to think innovatively and collectively, to identify new ways to help improve the care, management, and treatment of these often-devastating diseases.

Background

It is estimated that there are more than 7,000 rare diseases (RDs) affecting about 30 million Americans. While the exact cause for many RDs remains unknown, for a large number of RDs the origin can be traced to mutations in a single gene, contributions from multiple genetic factors, and/or a combination of genetic and environmental factors. Besides the direct medical costs associated with RD, there are significant costs to society and individuals, including indirect costs associated with productivity losses and non-medical costs such as medical care not covered by insurance and necessary modifications of homes or vehicles for accessibility. Many individuals with RDs have high medical needs requiring that they miss work, retire early, and utilize the assistance of a caregiver for activities for daily living. As such, the economic burden of RD is likely to be significant, for patients, unpaid family caregivers, and to society.

There is limited evidence on RD prevalence and economic burden in the United States (U.S.) because of a lack of a national registry or all-payer database enabling the calculation of disease prevalence or cost. Although data sources such as the Europe-based Orphanet contain prevalence estimates for many RDs, these data may or may not apply to the U.S. population and do not always contain details such as the

age-specific prevalence. A literature review also found that most RD related cost studies are Europe-based and often focus on a small number of high cost RDs." Cost of illness information guides health-care decision-makers in quantifying the impact of diseases at a population level. This can inform resource allocation, as well as healthcare cost projections. However, data on the burden of RDs in the U.S. is scarce. To address this knowledge gap, we estimated the economic burden of RD in the U.S. in 2019. We aimed to provide the most comprehensive estimate of the total economic burden of RD from a societal perspective and expand the evidence base in those less well-understood cost components and caregiver burden.

Methods

To estimate the national economic burden by patient age and type of RD, we took a prevalence-based approach. This approach combined the RD prevalence with per-person disease-attributable excess cost (i.e., a difference in the annual per-person costs between the RD sample and matched controls without RD). Direct medical costs were captured through an administrative claims-based analysis the indirect productivity losses for persons with RD and their caregivers, non-medical costs associated with RD, and expenditures for daily non-medical care were captured via survey that we designed and implemented— the National Economic Burden of Rare Disease Survey (hereafter, the Survey). This survey was one of the largest surveys conducted so far covering multiple communities representing 379 RDs. We mapped 379 RDs to 16 disease groups (see additional file 2) representing a body system and corresponding to the International Classification of Disease 10th Edition (ICD-10) diagnosis coding system, which was done in several steps. First, we identified the ICD-10 codes for each disease. Then, we mapped codes and added decision tree logic to the most granular (i.e., "leaf nodes"). Finally, we ran the data against the Orphanet codes to see whether there was agreement and whether we could identify any codes in Orphanet that were unidentifiable using the ICD. The entire process was reviewed several rounds by a designated technical advisor group consisting clinical experts, RD researchers, and family advocates. Figure 2 shows the cost calculation steps and data sources. The study received approval by an Institutional Review Board.

RD Prevalence

Because there is not a single dataset that addresses prevalence of multiple RDs and associated costs, we relied on claims data from three sources: 2019 Medicare Standard Analytical File (SAF, Medicare 5%), 2016 Transformed Medicaid Statistical Information System (TMSIS, Medicaid), and 2018 Optum de-identified Normative Health Information System (dNHI, a large, geographically diverse commercial claims database).

We estimated RD prevalence based on U.S.-specific claims: the Medicare 5% sample for the Medicare population, TMSIS claims for Medicaid population including children covered under the Children's Health Insurance Program (CHIP) and adults <65, and the dNHI claims for the privately insured population <65. In each database, we identified patients with 379 different RDs using ICD-10 codes and categorized them

into 16 RD groups for adults and 7 RD groups for children (<18 years), as the sample size was too small to report estimates individually for all 16 RD groups. This allowed us to increase the sample size and to be consistent with the disease groups used in the survey study.

Direct Medical Costs

To quantify the annual excess medical costs associated with RD, we compared the average total cost of people with RD with that of a matched comparison group with similar characteristics including age, gender, race/ethnicity, and insurance, but without any RD, using a 10:1 control to case ratio.

Medical costs based on claims data included primary payer paid amount, patient out-of-pocket expenses (e.g., copay, co-insurance, deductibles), and any third party paid amount. We estimated the direct medical costs of RD by types of healthcare service: acute and non-acute inpatient stay; outpatient care; physician office visit; durable medical equipment (DME); other ancillary, outpatient-based drug administration; retail prescription drug use; and caregiver payments (by Medicaid). Since the Medicare 5% data do not include Part-D claims, we used commercial per-person prescription cost to impute the Medicare per-person prescription cost for each age and disease group. All costs estimates were expressed in 2019 dollars. Average direct medical cost was calculated for each of the 16 RD groups for adults and the 7 RD groups for children. We used the group average across all RDs to compare with the comparison group to derive the average RD-attributable cost, to avoid double counting.

Indirect Costs and Non-Medical Costs

We worked with broad coalition of patient advocacy organization partners to design an online survey to estimate cost due to: reduced labor market participation, productivity loss for those in the labor force, non-medical costs of RD (such as the cost of hiring professional non-medical caregivers to assist with daily living, necessary home modification costs), and disability benefits. The respondents were persons with RD. However, the family member most familiar with the health of the person with the RD could respond to the survey, if the health of the person with RD prevented accurate self-reporting, or if the person was a minor.

We took convenience sample approach and disseminated the survey to the RD communities via partner networks of more than 200 partner-patient advocacy organizations, reaching a broad range of RDs and large patient sample.

We received 3,484 responses with 1,399 being fully completed. The survey asked about the respondents' disease and 581 RDs were reported. After removing misspellings, RD alternative names, and diseases that were not rare (e.g., cancer) or represented a protein, there were 379 unique RDs.

Additional file 3 provides the breakdown of the respondents' self-description and shows that 57% of the respondents were people with a RD, and 41% of responses were from a family caregiver. About 28% of the responses represent children (<18 years); the rest were adults, with those above age 65 representing 14% of all responses (additional file 4). People with RDs were predominantly white (87%), followed by multi-

racial individuals (4%). About 77% of people with RD had at least one caregiver (a primary caregiver) and about half had both a primary and a secondary caregiver; 23% of the RD sample did not rely on a caregiver.

Given the survey sample size, the number of RD groups, age stratification, and the number of indirect and non-medical cost components, it was not feasible to calculate reliable cost estimates for all RD groups-age strata. Where the strata sample size permitted, we calculated average costs for that strata; if strata sample size was <5 observations, we reported the average costs across all RD groups. The mapping of cost components to RD group-age strata is reported in additional file 5; cost component calculations are detailed below.

RD may increase the likelihood that severe functional impairment or disability will prevent them from working, or limit employment opportunities and reduces earnings. [3,4] Our survey indicated that among the working age (18-64) persons with RD, 43.8% are in the labor market, as compared to the national labor force participation rate of 63.1% among the U.S. adult population. To ensure that the early termination of employment was related to RD, we calculated labor market employment related earnings loss due to RD as the count of persons with RD, who have retired or stopped working in the past 12 months and indicated that RD played a major role in their decision, multiplied with the median annual earnings by job status (full-time versus part-time) obtained from the 2019 American Community Survey public use microdata sample. We used medians rather than averages, as medians are less likely affected by outliers. As the full-time/part-time status of persons with RD before retirement was unknown, we used the allocation of full-time to part-time job status among currently working persons with RD. Then, we calculated earnings loss due to early retirement for those who retired due to RD as a weighted average between those assumed working full-time before retirement and those working part-time before retirement.

We calculated two measures of reduced labor market productivity for those who are employed: absenteeism, (increased workdays missed due to illness), and presenteeism, (illness-related poorer work performance while on the job). We asked about the number of days in an average working month during 2019 the person with RD and the caregivers missed work or felt less productive while at work because of RD. Based on responses to these two questions and the average daily earnings calculated from the self-reported annual earnings, we calculated the productivity loss due to absenteeism by multiplying the number of days missed with the daily earnings and then annualized the total loss. Presenteeism was calculated similarly, with an adjustment factor applied to each day felt unproductive, reflecting that an unproductive day is not equivalent of a total loss of a whole day's value. The adjustment factor was obtained from the responses to productivity self-assessment scale: i.e., on days when feel less productive, on average how productivity of the person with RD and caregivers was affected on a scale from 0 to 10, where 0 represents "not at all", 1-3 "mildly", 4-6 "moderately", 7-9 "markedly", and 10 represents "extremely". We translated these responses into the reduced overall productivity (e.g., 0 corresponds to 100% productivity, 10 corresponds to 0% productivity or the full reduction in productivity). Daily earnings were calculated from the annual earning brackets that was applicable to the respondent in

2019 (categorical responses were converted into numerical values based on the mid-point of each earnings category: everyone who indicated earnings "less than $1,000" were assigned earnings of $500, etc.).

Additionally, RD may affect patients and the caregivers' ability to participate in various social activities using their leisure time.The challenge of quantifying social productivity is the measuring the time forgone from social activities and in the proper valuation of the time forgone. Although one could argue that forgone leisure time visiting family and friends also creates economic loss, we focused on activities that directly involved volunteering and provide a conservative estimate of the social productivity loss. We asked about the number of hours the person with RD and the caregivers spent in a typical week before and after RD started having a significant impact, on the following social activities: performing voluntary or charity work; providing help to family/friends/neighbors unrelated to personal care or care for person with RD; participating in a political or community-based organization.

We compared self-reported volunteering hours before RD with the average national annual volunteering hours obtained from the Current Population Survey (CPS) Volunteer Supplement that measures the population's participation in volunteer activities (2017). The national average volunteering hours are generally lower than the self-reported volunteering hours (e.g., 1.9 hours per week versus 12.1 hours per week for person with RD before RD). Therefore, we took a conservative approach in our calculations by calculating the percentage of people volunteered and average hours volunteered from CPS and multiplied with the estimated percentage productivity loss from the Survey (calculated as the difference between before and after hours divided by before hours) for the three activities combined. Productivity loss due to forgone volunteering activities was calculated as volunteering hours affected per year times $27.20, which is a dollar value per volunteering hour according to the Independent Sector.

The non-medical costs calculated included expenses of purchasing formal care (e.g., adult day care and personal aides) and necessary modification to homes, purchases of adapted motor vehicles or car modifications for accessibility, medical foods, dietary supplements, specialty clothing (e.g., compression stockings), and increased travel costs for medical visits. We also asked about healthcare services not covered by insurance such as experimental treatments, alternative or non-traditional treatments (alternative therapies, massage therapy, acupuncture), and over-the-counter drugs. We estimated cost of non-medical components and medical out-of-pocket costs by multiplying the weighted percentage of families who responded as having such expenses and the average expense per-family per-year.

To capture the overall economic burden of RDs, it is always an important policy perspective to be able to identify the extent to which individuals are transitioning into public programs, and what the potential costs to public programs are due to any specific condition/disease, particularly if these costs are avoidable. For example, the Social Security Disability Insurance (SSDI) and the Supplemental Security Income (SSI) are considered as transfer payments (i.e., a cost to one person is a benefit to another person). Therefore, these components may inform on the extent of government budgetary burden due to a specific disease. We asked respondents whether the person with RD had received SSI, SSDI, or other

types of disability income, in 2019. While we estimated the average and total disability income due to RD, these costs were excluded from the overall burden estimates, as these funds could have been used for healthcare payments or non-medical expenses already captured in other cost components.

Results

An estimated 15.5 million individuals in the U.S. have at least one of the 379 RDs were included in this study in 2019. The estimated RD prevalence of 379 RDs is reported in Table 1.

Table 1
Rare Disease Prevalence for 379 RDs Included in the Study by Age and Insurance Coverage (in 2019)

	No. of Persons Estimated to Have RD	Population	Prevalence
Age			
<18	1,322,886	71,580,109	1.8%
18-64	8,371,639	182,528,781	4.6%
≥65	5,850,660	51,822,242	11.3%
Insurance			
Commercial	7,124,610	188,738,510	3.8%
Medicaid	1,582,062	57,833,466	2.7%
Medicare	6,838,513	59,359,156	11.5%
Total	**15,545,185**	**305,931,132**	**5.1%**

Source: Analyses of 2018 dNHI claims, 2019 Medicare claims, and 2016 Medicaid claims, and Census population projection for 2019.

Census population projection for 2019.

The estimated total economic burden of 379 RDs in 2019 was $966 billion, including a direct medical cost of $418 billion and an additional $548 billion in indirect and non-medical costs. **Figure 1** shows the estimated total economic burden of RD by cost components. The direct medical cost of RD represents almost half of the total burden (45%), followed by absenteeism losses (15%), and presenteeism losses (14%). Overall, labor market productivity losses due to RD, including absenteeism, presenteeism, and earnings losses from forced retirement, represent roughly the same proportion of the total burden as direct medical costs (42%). Non-medical costs and healthcare cost not covered by insurance represent about 11% of the total burden.

Medical Direct Costs

Table 2 shows that the working-age population with commercial health coverage bears the majority of the medical cost of RD (47%). The Medicare population with RD accounts for about 43% of total excess costs, and the Medicaid population with RD accounts for the remaining 10%. The estimated $449 billion for medical direct cost translate into a per-person excess cost of $28,913, implying that an average person with RD has an annual medical cost that is $28,913 more than a comparison person without RD. Per-person direct medical cost of RD decreases with age, with children on average having an annual excess cost of $32,037 when they have RD; $29,647 for working-age adults with RD, and $27,157 for RD patients age 65 and older. Across different insurance coverages, the highest per-person excess costs are for the privately insured ($29,910), followed by cost for the Medicare beneficiary population with RD ($28,185); the lowest per-person excess costs are for persons with RD who are on Medicaid ($27,573).

Table 2
Direct Medical Cost of Rare Diseases by Age, Insurance Coverage, and Type of Service

	Total Excess Medical Cost due to RD		Per Capita
	(in $ Millions)	Percentage of the Total	(2019 $s)
Age			
<18	42,381	9.4%	32,037
18-64	248,198	55.2%	29,647
≥65	158,884	35.3%	27,157
Insurance			
Medicaid	43,621	9.7%	27,573
Commercial	213,094	47.4%	29,910
Medicare	192,747	42.9%	28,185
Type of service			
Inpatient	143,000	31.8%	9,199
Prescription medication	79,466	17.7%	5,112
Outpatient	62,032	13.8%	3,990
Other ancillary	48,974	10.9%	3,150
Outpatient prescription administration	47,567	10.6%	3,060
Physician	31,372	7.0%	2,018
Non-acute inpatient	30,759	6.8%	1,979
Durable medical equipment	4,401	1.0%	283
Caregiver	1,890	0.4%	122
Overall	**449,462**	**100%**	**28,913**

Source: RD prevalence calculated from claims data (2018 dNHI, 2019 Medicare 5%, and 2016 Medicaid) combined with the 2019 Census population projections; direct medical cost estimates also based on three claims databases. Other ancillary services include telehealth, ambulance transportation via land, air or water, mobile unit services, etc.

Hospital inpatient care and prescription medication are the two largest cost categories, representing 32% and 18% of the total direct medical cost, respectively. DME and caregiver costs represent the two smallest categories. Caregiver costs were only covered by Medicaid; however, we estimated them as an average value for the entire RD population, to be consistent with other cost categories.

Additional file 1 shows that the highest per-person excess medical cost for children is for "Lysosomal storage diseases" - $132,757, followed by "Other endocrine or metabolic disorders ($72,285). For adults, the two most expensive RD group based on per-person excess costs are "Neoplasms" ($58,025) and "Lysosomal storage diseases" ($56,087).

Indirect Costs and Non-Medical Costs

The estimated total indirect and non-medical cost of RD is $548 billion in 2019, with $343 billion to persons with RD and another $205 billion to unpaid caregivers. Table 3 shows that the estimated total indirect and non-medical cost of RD is $64 billion for children (age <18) and $484 billion for adults (age ≥18).

Table 3
Total Indirect and Non-Medical Costs by Cost Components in 2019 (in Million $s)

	Age <18			Age ≥18		
	Person with a RD	Primary Caregiver	Secondary Caregiver	Person with a RD	Primary Caregiver	Secondary Caregiver
Indirect Productivity Loss						
Forced retirement	NA	850	1,083	88,877	38,462	6,823
Absenteeism	NA	10,265	14,497	59,853	50,176	14,024
Presenteeism	NA	10,176	7,232	74,741	40,453	5,367
Social productivity loss in volunteer work	494	550	424	8,226	4,035	89
Non-Medical Costs and Medical Costs not Covered by Insurance						
Healthcare services not covered by insurance	2,172	NA	NA	35,750	NA	NA
Paid daily care	1,482	NA	NA	7,477	NA	NA
Necessary home modification	1,682	NA	NA	8,709	NA	NA
Special equipment at home or on a personal/family vehicle	1,865	NA	NA	21,677	NA	NA
Transportation costs	1,305	NA	NA	19,127	NA	NA
Home schooling	454	NA	NA	NA	NA	NA
Missed schooling	2,298	NA	NA	NA	NA	NA
Special education	7,199	NA	NA	NA	NA	NA
Overall in $ Millions	18,951	21,840	23,237	324,437	133,125	26,303
Grand Total (to society)	**64,028**			**483,865**		

Source: RD prevalence calculated from claims data (2018 dNHI, 2019 Medicare 5%, and 2016 Medicaid) combined with the Census population projection for 2019; indirect and non-medical costs estimated from the Survey data.

Among indirect and non-medical costs associated with RD, absenteeism to both persons with RD and their caregivers is nearly $149 billion (27% of the $548 billion), followed by presenteeism cost ($138 billion, 25%), and losses due to forced retirement ($136 billion, 25%). For adults, the costs of absenteeism for caregivers are about the same as those for the person with RD ($64 billion versus $60 billion). However, when the costs to caregivers for children with RD are taken into account, the costs of absenteeism for the caregivers surpass those for the person with RD ($89 billion versus $60 billion);

losses associated with presenteeism are slightly smaller for caregivers compared to the person with RD ($63 billion across all caregivers for children and adults with RD versus $75 billion for persons with RD).

The total non-medical cost, including medical costs not covered by insurance, is $111 billion. Healthcare services not covered by insurance (e.g., experimental, alternative, or non-traditional treatments; over-the-counter drugs; dental surgeries) represent the largest share (34%), followed by costs of necessary specialized equipment for home or motor vehicle (21%), and transportation cost (18%).

Table 4 shows that the average indirect and non-medical cost per child with RD is $14,326 for person with RD only and $48,400 for person with RD combined with caregiver costs. Per capita indirect and non-medical costs among adults are $22,812 per adult person with RD and $34,022 when person with RD is combined with caregiver burden.

Table 4
Per Capita Indirect and Non-Medical Costs by Cost Component (in 2019 $s)

	Age <18			Age ≥18		
	Person with a RD	Primary Caregiver	Secondary Caregiver	Person with a RD	Primary Caregiver	Secondary Caregiver
Indirect Productivity Loss						
Loss due to forced retirement	NA	642	818	6,249	2,704	480
Absenteeism	NA	7,759	10,959	4,208	3,528	986
Presenteeism	NA	7,692	5,467	5,255	2,844	377
Social productivity loss in volunteer work	373	416	321	578	284	6
Non-Medical Costs and Medical Costs not Covered by Insurance						
Healthcare services not covered by insurance	1,642	NA	NA	2,514	NA	NA
Paid daily care	1,121	NA	NA	526	NA	NA
Necessary home modification	1,271	NA	NA	612	NA	NA
Special equipment at home or on a personal/family vehicle	1,409	NA	NA	1,524	NA	NA
Transportation costs	986	NA	NA	1,345	NA	NA
Home schooling	344	NA	NA	NA	NA	NA
Missed schooling	1,737	NA	NA	NA	NA	NA
Special education	5,442	NA	NA	NA	NA	NA
Overall	14,326	16,509	17,565	22,812	9,360	1,849
Grand Total	**48,400**			**34,022**		

Source: RD prevalence calculated from claims data (2018 dNHI, 2019 Medicare 5%, and 2016 Medicaid) combined with the Census population projection for 2019. Indirect and non-medical costs estimated from the Survey data.

Discussion

We showed that RD significantly affected individuals living with RDs, as well as their unpaid caregivers, employers, and payers. Data demonstrated that diagnosis, care navigation, and treatment of RD included

significant costs to individual households and the health system. Our finding can help guiding health-care decision-makers in setting up and prioritizing health-care policies and interventions related to RDs.

Our estimate of the overall economic burden is based on a subset of 379 RDs and, therefore, represents a lower bound estimate and is not generalizable to RDs not included in this study. The previous burden estimates are generally limited to a specific RD or a small group of RDs and U.S.-based cost estimates are very scarce. With the exception of cystic fibrosis and hemophilia, which are relatively well studied, information on the economic cost for other conditions (e.g., Duchenne muscular dystrophy, fragile X syndrome, juvenile idiopathic arthritis, mucopolysaccharidosis, scleroderma, Prader-Willi syndrome, histiocytosis, epidermolysis bullosa) is very limited.

Previous studies that target a specific disease produce cost estimates with very wide ranges. For example, Kawalec and Malinowski's (2015) systematic review of the indirect costs related to psoriatic arthritis, showed that per person indirect cost range was $1,694 - $12,318 using the friction cost approach and $1,751 - $50,270 (in 2013 $s) using the human capital approach. In our study, psoriatic arthritis is included in the "Diseases of the skin and subcutaneous tissue" group, with estimated indirect costs of $22,066 per adult with such conditions. We were not able to identify the U.S.-based studies that estimated the overall burden of RD, but a few studies focused on the burden of specific diseases. For example, aggregate inpatient costs among eleven genetic orphan diseases in the U.S. were roughly $1 billion in 2016, with cystic fibrosis and sickle cell disease having the largest costs ($414.8 million and $338.2 million). Therefore, any comparison between the findings of this study and any literature should consider these differences.

Most people living with RDs in this study were between the ages of 18 and 65. While commercial payers bore the largest share of the direct medical cost, indirect and non-medical costs paid by families were greater than direct costs. Significant productivity losses associated with absenteeism and presenteeism were experienced by employers. These productivity losses included $135 billion from adults with RD whose disease progression and diagnoses required time away from the workplace and $152 billion from their caregivers. Additionally, disability based government supplemental income programs provide disability support ($115 billion) to persons with RD whose ability to participate in labor market is significantly affected by RD.

Our study has several limitations. First, RD prevalence and direct medical cost estimates relied on one diagnosis code, which may include false positive cases (i.e., persons with a singular diagnosis code who later ruled out as having a RD). While we captured as many RDs as possible, this may have resulted in an overestimate of the prevalence, and/or an underestimate of the direct medical costs (as patients with one diagnosis code may not be as severe as those with multiple diagnosis codes of RD in a given year).

Additionally, not every RD was matched to a definitive ICD-10 code. Some RDs we matched to the closest ICD-10 code, which might have captured a broader group of people than intended. Conversely, in some instances the absence of a specific ICD-10 code for a RD required mapping to a disease category that was not rare, thus eliminating it from inclusion in the direct cost analysis. The lack of ICD code specificity

in the majority of rare diseases often adds to the overall rare disease burden by impacting availability of disease data to inform payer decision making, diagnostic and treatment algorithms, and burden of public health surveillance.

Second, the indirect and non-medical costs were based on the self-reported data. Since it is a convenience sample, there may be selection biases if people who are more severely affected by a RD are more likely to respond to the survey than those with less severe RD. Conversely, only the experience of individuals with confirmed RD diagnoses and the ability to participate in a rigorous online, self-report survey are represented in the survey sample.

It should be noted that this study focuses on RD burden in 2019, and does not reflect the cumulative costs associated with living with a RD over time.

Finally, due to small sample sizes, we could not break down burden estimates by desired population strata (e.g., sex, race/ethnicity). For indirect and non-medical costs, we aggregated up disease groups with ultra-small sample size.

Despite these limitations, this study is the largest and most comprehensive effort so far in the U.S. to measure the societal impact of many RDs at once, encompassing various cost components (such as cost of social productivity loss and cost of special education), and including estimates of indirect productivity loss for family caregivers whose lives are significantly affected by RD.

Conclusions

Together, the 379 RDs included in this study represent an economic burden that far surpasses some of the costliest chronic diseases studied in the U.S., both due to a high prevalence (i.e., more than 15 million individuals with these rare diseases) and a higher per-person cost. The findings of this study highlight the scale of the RD burden and call for immediate attention from the scientific communities, policy leaders, and other key stakeholders such as health care providers and employers, to think innovatively and collectively, to identify new ways to help improve the care and treatment of RD. These findings demonstrate that the RD community has a significant unmet need with tremendous public health impact, one that requires urgent support to advance research and development of resources for prevention, management, and ultimately, cures of these often-devastating diseases. All these steps may lead to significant individual and societal benefit.

Abbreviations

CHIP: Children's Health Insurance Program

dNHI: Optum de-identified Normative Health Information System

ICD-10: International Classification of Disease 10th Edition

RD: Rare Disease

SAF: Medicare Standard Analytical File

SSI: Supplemental Security Income

SSDI: Social Security Disability Insurance

TMSIS: Transformed Medicaid Statistical Information System

U.S.: United States

Declarations

Ethics approval and consent to participate

The study received an approval by the WCG New England Institutional Review Board (IRB). The survey instrument and consent form were reviewed by the IRB (IRB Tracking number 20201796).

Consent for publication

Survey participant consented to completion of online survey. All data reported are aggregated and no individual data are reported in the manuscript.

Availability of data and materials

The data that support the findings of this study were obtained using standard contracts and data use agreements for the analysis of all datasets. The private, Medicare, and Medicaid claims datasets for this study are proprietary to Optum and CMS and, therefore, cannot be shared without a data use agreement with Optum and CMS, respectively. Parties interested in the survey data need to contact the EveryLife Foundation for Rare Diseases for permission.

Competing interests

The authors declare that they have no competing interests.

Funding

This study was supported by the EveryLife Foundation for Rare Diseases.

Authors' contributions

GY contributed to conceptual framework, review of results, and writing of the manuscript. IC analyzed and interpreted survey data, calculated indirect impacts, and contributed to writing. AK, EO, AP, and JS performed review and assignment of individual rare diseases into larger grouping, review and assignment

of ICD-10 codes to individual rare diseases, and contributed in writing the manuscript. All authors read and approved the final manuscript.

References

1. Haendel M, Vasilevsky N, Unni D, et al. How many rare diseases are there? Nat Rev Drug Discov. 2020;19(2):77-78; doi: 10.1038/d41573-019-00180-y.

2. National Human Genome Research Institute. Rare Diseases FAQ. January 10, 2020. https://www.genome.gov/FAQ/Rare-Diseases. Accessed April 19, 2021.

3. Forshaw R. What is the cost of living with rare disease? https://raremark.com/articles/what-does-it-cost-to-live-with-a-rare-disease--572. Accessed January 29, 2021.

4. Therapies, Shire Human Genetic. Rare Disease Impact Report: Insights from patients and the medical community. Shire Human Genetic Therapies, Tech. Rep (2013).

5. Orphanet is a European website providing information about orphan drugs and rare diseases. It contains content both for physicians and for patients. Its administrative office is in Paris and its official medical journal is the Orphanet Journal of Rare Diseases published on their behalf by BioMed Central. Web link is available at: https://www.orpha.net/consor/cgi-bin/index.php

6. Chevreul K, Brigham KB, Gandré C, Mouthon L; BURQOL-RD Research Network. The economic burden and health-related quality of life associated with systemic sclerosis in France. Scand J Rheumatol. 2015;44(3):238-46; doi: 10.3109/03009742.2014.976653

7. López-Bastida J, Oliva-Moreno J, Linertová R, Serrano-Aguilar P. Social/economic costs and health-related quality of life in patients with rare diseases in Europe. Eur J Health Econ. 2016;17 Suppl;1:1-5; doi: 10.1007/s10198-016-0780-7.

8. Floyd DN, Langham S, Séverac HC, Levesque BG. The economic and quality-of-life burden of Crohn's disease in Europe and the United States, 2000 to 2013: a systematic review. Dig Dis Sci. 2015;60(2):299-312; doi: 10.1007/s10620-014-3368-z.

9. Greenberg D, Mohamed Ibrahim M, Boncz I. What Are the Challenges in Conducting Cost-of-Illness Studies?. Value in health region issues. 2014;4:115–116; 10.1016/j.vhri.2014.08.003.

10. Angelis A, Tordrup D, Kanavos P. Socio-economic burden of rare diseases: a systematic review of cost of illness evidence. Health Policy. 2015;1;119(7):964-79.

11. Kawalec P, Malinowski KP. The indirect costs of psoriatic arthritis: systematic review and meta-analysis. Expert Rev Pharmacoecon Outcomes Res. 2015;15(1):125–132; doi 10.1586/14737167.2015.965154

12. MIT NEWDIGS Research Brief 2019 F205-v038-Pediatric Gene Therapy Launches. May 24, 2019. https://newdigs.mit.edu/sites/default/files/FoCUS%20Research%20Brief_2019F205v038b.pdf

13. Harrington M, Hareendran A, Skalicky A, et al. Assessing the impact on caregivers caring for patients with rare pediatric lysosomal storage diseases: development of the Caregiver Impact Questionnaire. J Patient Rep Outcomes. 2019;3(44):1-14; doi 10.1186/s41687-019-0140-3.

14. Independent Sector. The Current Estimated National Value of Each Volunteer Hour Is $27.20. July 20, 2019. https://independentsector.org/value-of-volunteer-time-2020/. Accessed January 21, 2021.

Figures
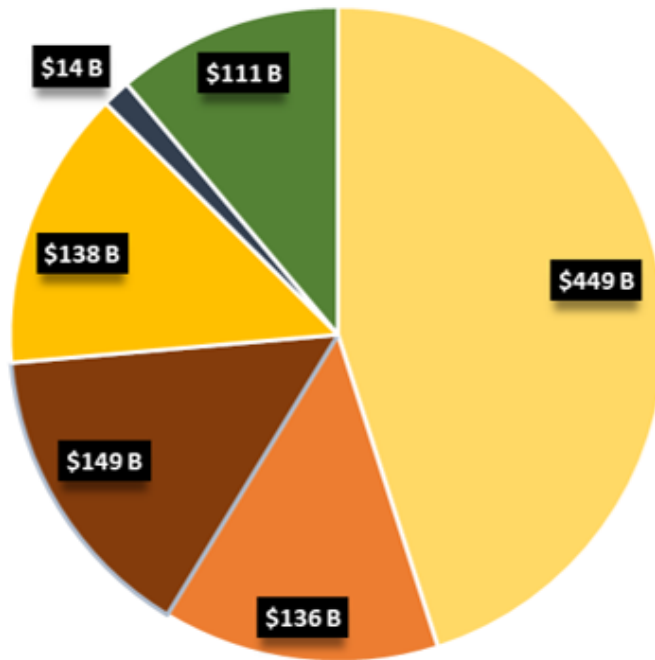
Figure 1

Total Economic Burden of Rare Disease in the U.S. in 2019: $997 Billion
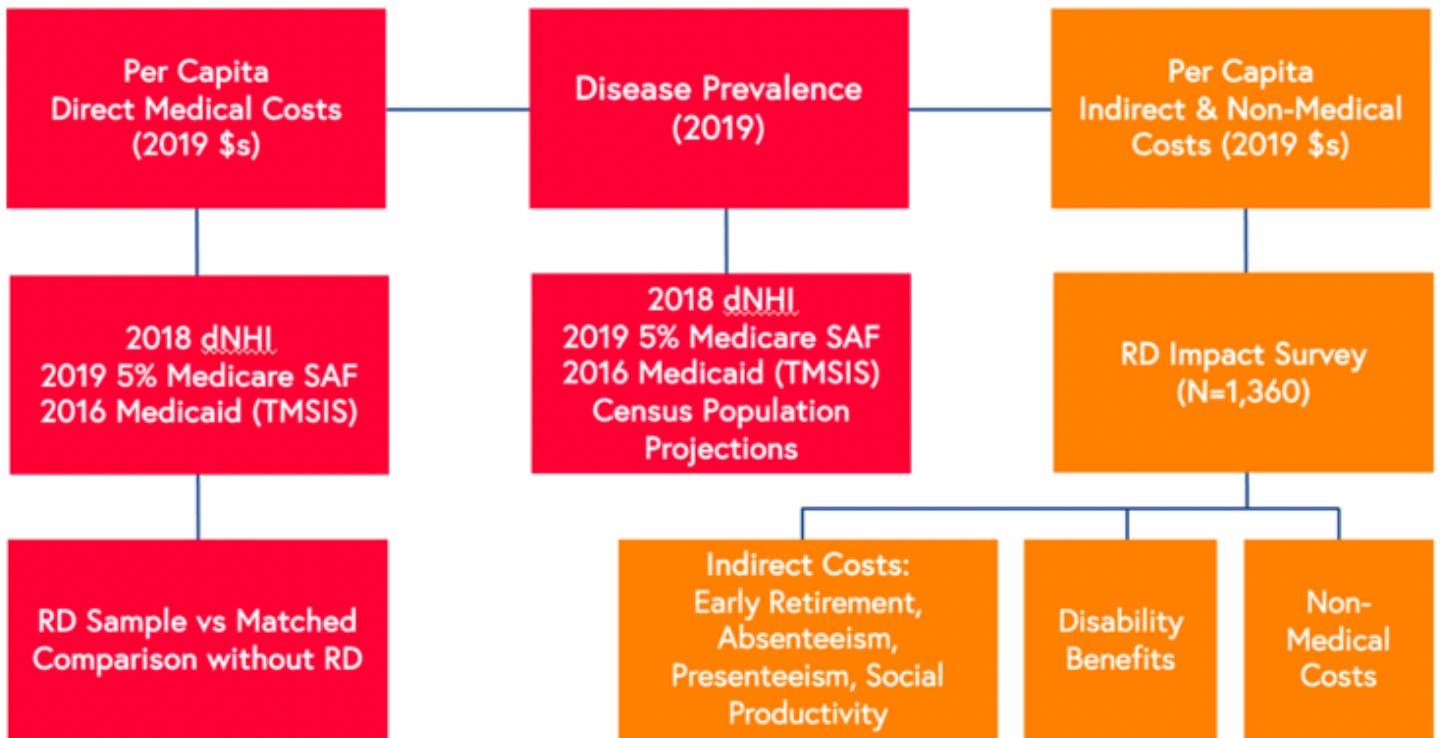
Figure 2

Flow Chart of Cost Calculation and Data Sources

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Additionalfile1.docx
- Additionalfile2.docx
- Additionalfile3.docx
- Additionalfile4.docx
- Additionalfile5.docx