

Zombie statistics and the prevalence of rare diseases: A scary story

Donna Wix, ASA, MAAA
David V. Williams



Now that season 10 of the hit TV show, *The Walking Dead*, is on hiatus until 2020, our minds turn to zombies of a different sort. A zombie statistic is a false or misleading statistic, often reanimated from studies conducted many years ago and now printed as truth without citation.

Well-constructed prevalence studies on health conditions require large populations to produce statistically relevant results. The more rare the disease, the bigger the required population. The expense and effort to accurately assess the prevalence of a rare disease is beyond the reach of most study budgets. As a result, we often see zombie-like rare disease prevalence estimates that may not have a basis in reality. In addition, publications rarely segment prevalence rates by U.S. health insurance market (commercial, Medicaid, Medicare, individual) even though sometimes significant differences in prevalence may exist among different covered populations.

The U.S. incents the development of drugs to treat rare diseases through the Orphan Drug Act of 1983, which defines a rare disease as having fewer than 200,000 individuals affected. We used real-world data to calculate the age-adjusted prevalence rate for three diseases: rheumatoid arthritis, which is uncommon but not a rare disease by the Orphan Drug Act definition; ulcerative colitis, which could be considered a rare disease depending on which source is cited; and hemophilia, for which antihemophilic factor was approved in 2010 by the Food and Drug Administration (FDA) under the 1983 Orphan Drug Act. We then compared these results to zombie statistics commonly cited.

We queried administrative health claims databases of group commercial, individual, Medicare Advantage, and managed Medicaid populations with incurred claims during 2016 and 2017, consisting of over 926 million member months, searching for patients with ICD-10 codes describing each of the rare diseases. We calculated prevalence by age and then applied these results to the U.S. population to produce an estimate of the number of patients in the United States who are currently diagnosed with each disease.

Rheumatoid arthritis

According to the Arthritis Association approximately 1.5 million people in the United States have rheumatoid arthritis (RA).¹ Healthline.com reports 1.3 million.² Yet neither of these authoritative organizations names a source for their estimates. Both organizations indicate that RA often occurs later in life, yet neither provides estimates by age or specifically for the Medicare population.

A 2019 study of the incidence and prevalence of rheumatoid arthritis within members of an integrated healthcare delivery system (Aniket A. Kawatkar et al)³ and another paper, describing the cost-effectiveness of Sarilumab monotherapy for adults with active rheumatoid arthritis (Melanie D. Whittington et al),⁴ cite a Rochester, Minnesota, study published in 1999.⁵ These papers indicate that the prevalence of RA in the United States is 1.5 million, in the Kawatkar, and 1.8 million, in the Whittington. The Whittington study was conducted using medical records in Rochester, Minnesota, between 1955 and 1984, which was about the time of the dawn of the personal computer and well before the use of cell phones.

- 1 Arthritis Foundation. What Is Rheumatoid Arthritis? Retrieved December 1, 2019, from <https://www.arthritis.org/about-arthritis/types/rheumatoid-arthritis/what-is-rheumatoid-arthritis.php>.
- 2 Healthline. Rheumatoid Arthritis by the Numbers: Facts, Statistics, and You. Retrieved December 1, 2019, from <https://www.healthline.com/health/rheumatoid-arthritis/facts-statistics-infographic#1>.
- 3 Kawatkar, A.A., Gabriel, S.E., & Jacobsen, S.J. (March 2019). Secular trends in the incidence and prevalence of rheumatoid arthritis within members of an integrated health care delivery system. *Rheumatology International*. Retrieved December 1, 2019, from <https://link.springer.com/article/10.1007/s00296-018-04235-y>.
- 4 Whittington, M.D., McQueen, R.B., Ollendorf, D.A. et al. (January 2019). Assessing the value of Sarilumab monotherapy for adults with moderately to severely active rheumatoid arthritis: A cost-effectiveness analysis. *Journal of Managed Care and Specialty Pharmacy*. Retrieved December 1, 2019, from <https://www.jmcp.org/doi/pdf/10.18553/jmcp.2019.25.1.080>.
- 5 Gabriel, S.E., Crowson, C.S., & O'Fallon, W.M. (March 1999). The epidemiology of rheumatoid arthritis in Rochester, Minnesota, 1955-1985. *Arthritis and Rheumatism*. Retrieved December 1, 2019, from <https://onlinelibrary.wiley.com/doi/epdf/10.1002/1529-0131%28199904%2942%3A3%3C415%3A%3A%3A%3AID-ANR4%3E3.0.CO%3B2-Z>.

The table in Figure 1 shows the age-adjusted RA prevalence rate per 100,000 people by payer as calculated from our administrative claims data. Note the considerable variance by market.

Additionally, the prevalence is substantially higher than the cited numbers described above. We included a 95% confidence interval, which means that, as an example, we are 95% confident we would find between 2,579 and 2,635 (2,607 +/- 28) people with RA from a population of 100,000 Medicare beneficiaries.

FIGURE 1: RHEUMATOID ARTHRITIS

MARKET	INSURED POPULATION ⁶	PEOPLE WITH RA	PREVALENCE PER 100,000 PEOPLE	95% CONFIDENCE INTERVAL
Medicare	42,200,000	1,100,000	2,607	+/- 28
Commercial and Tricare	159,600,000	1,500,000	940	+/- 5
Individual	21,200,000	200,000	943	+/- 37
Medicaid	66,000,000	400,000	606	+/- 49
Total	289,000,000	3,200,000	1,107	

Ulcerative colitis

Up until December 2019, the Centers for Disease Control and Prevention (CDC) stated that ulcerative colitis (UC) affects 37 to 246 people per 100,000 of the population. The CDC's source, published in 2004, is a study by Edward V. Loftus, Jr.⁷ This paper is a review of selected registries in North America, Europe, and Asia, with prevalence dates ranging from 1980 to 1997. The prevalence rate endpoints of the CDC range come from two different locations two decades apart: a northern Alberta, Canada, population in 1981 is the source of the 37/100,000 rate and an Olmstead County, Minnesota, population in 2001 is the source of the 246/100,000 rate.

The American College of Gastroenterology (ACG) Clinical Guideline for Ulcerative Colitis in Adults,⁸ published in 2019, states that nearly 1 million Americans are affected with UC, though it doesn't note a source.

6 Kaiser Family Foundation (2017). Health Insurance Coverage of the Total Population. State Health Facts. Retrieved December 1, 2019, from <https://www.kff.org/other/state-indicator/total-population/?currentTimeframe=0&sortModel=%7B%22colId%22:%22Location%22,%22sort%22:%22asc%22%7D>.

7 Loftus, E.V. (May 2004). Clinical epidemiology of inflammatory bowel disease: incidence, prevalence, and environmental influences. *Gastroenterology*. Retrieved December 1, 2019, from <https://www.sciencedirect.com/science/article/abs/pii/S0016508504004627?via%3Dihub>.

8 Rubin, D.T., Ananthakrishnan, A.N., Siegel, C.A. et al. (March 2019). ACG Clinical Guideline: Ulcerative Colitis in Adults. *American Journal of Gastroenterology*. Retrieved December 1, 2019, from <https://www.healthline.com/health/rheumatoid-arthritis/facts-statistics-infographic#1>.

The table in Figure 2 shows the age-adjusted UC prevalence rate per 100,000 population by payer that we calculated from our administrative claims data. The overall prevalence rate of 253 per 100,000 people is similar to the CDC's high end of the range, with the Medicare population having the highest prevalence rate (474/100,000) and Medicaid having the lowest (121/100,000). The total prevalence does not quite reach 1 million as quoted in the 2019 clinical guideline.

FIGURE 2: ULCERATIVE COLITIS

MARKET	INSURED POPULATION ⁹	PEOPLE WITH UC	PREVALENCE PER 100,000 PEOPLE	95% CONFIDENCE INTERVAL
Medicare	42,200,000	200,000	474	+/- 11
Commercial and Tricare	159,600,000	400,000	251	+/- 3
Individual	21,200,000	50,000	236	+/- 21
Medicaid	66,000,000	80,000	121	+/- 23
Total	289,000,000	730,000	253	

Hemophilia

The CDC estimates that 20,000 males in the United States suffer from hemophilia.¹⁰ This statistic is based upon a study by J.M. Soucie et al¹¹ published in 1998 and based on an "active surveillance system in six states."

In a recently published paper by Alfonso Iorio et al,¹² a total prevalence rate was reported at 20.9 per 100,000 males. In this study, the authors applied a steady state analysis to hemophilia registry data.¹³

9 Kaiser Family Foundation, Health Insurance Coverage of the Total Population, op cit.

10 Healthline. Rheumatoid Arthritis by the Numbers: Facts, Statistics, and You. Retrieved December 1, 2019, from <https://www.healthline.com/health/rheumatoid-arthritis/facts-statistics-infographic#1>.

11 Soucie, J.M., Evatt, B., & Jackson, D. (December 1998). Occurrence of hemophilia in the United States: The Hemophilia Surveillance System Project Investigators. *Am J Hematol*. Retrieved December 1, 2019, from <https://www.ncbi.nlm.nih.gov/pubmed/9840909>.

12 Iorio, A., Stonebraker, J.S., Chambost, H. et al. (October 15, 2019). Establishing the prevalence and prevalence at birth of hemophilia in males: A meta-analytic approach using national registries. *Annals of Internal Medicine*. Retrieved December 1, 2019, from <https://annals.org/aim/article-abstract/2749729/establishing-prevalence-prevalence-birth-hemophilia-males-meta-analytic-approach-using>.

13 Inserro, A. (September 9, 2019). Prevalence of hemophilia worldwide is triple that of previous estimates, new study says. In Focus Blog. Retrieved December 1, 2019, from <https://www.ajmc.com/focus-of-the-week/prevalence-of-hemophilia-worldwide-is-triple-that-of-previous-estimates-new-study-says->.

The table in Figure 3 shows the age-adjusted hemophilia prevalence rate per 100,000 male population by payer that we calculated from our administrative claims data. The overall prevalence rate is similar to the cited values described by Iorio, but we show a higher estimate of the number of males with hemophilia. Once again, there are differences by market, with Medicare and Medicaid having the highest rates among the four markets shown.

FIGURE 3: HEMOPHILIA

MARKET	INSURED POPULATION ^{14, 15, 16, 17, 18}	MALES WITH HEMOPHILIA	PREVALENCE PER 100,000 PEOPLE	95% CONFIDENCE INTERVAL
Medicare	19,300,000	5,000	26	+/- 11
Commercial and Tricare	78,200,000	10,000	13	+/- 3
Individual	21,200,000	4,000	19	+/- 9
Medicaid	27,700,000	8,000	29	+/- 17
Total	146,400,000	27,000	18	

Conclusion

Prevalence rates in rare diseases vary in accuracy, may become outdated with new advances in treatment, and are seldom broken into payer types found in the U.S. market. A large administrative claims survey of the prevalence of rare diseases may increase insight into the way individuals are treated for disease and may aid in providing funding for treatments. There is no reason to be afraid or to use old and undocumented prevalence estimates, even in rare diseases. A wealth of real-world data is waiting to be explored.

- 14 Kaiser Family Foundation, Health Insurance Coverage of the Total Population, op cit.
- 15 Kaiser Family Foundation (2017). Distribution of Nonelderly Adults With Medicaid by Gender. State Health Facts. Retrieved December 1, 2019, from <https://www.kff.org/medicaid/state-indicator/distribution-by-gender-4/?currentTimeframe=0&sortModel=%7B%22colId%22:%22Location%22,%22sort%22:%22asc%22%7D>.
- 16 AHIP (May 2019). Medicare Advantage Demographics Report, 2016. Retrieved December 1, 2019, from https://www.ahip.org/wp-content/uploads/MA_Demographics_Report_2019.pdf.
- 17 Kaiser Family Foundation (2017). Distribution of Nonelderly Adults With Employer Coverage by Gender. State Health Facts. Retrieved December 1, 2019, from <https://www.kff.org/other/state-indicator/distribution-by-gender-3/?currentTimeframe=0&sortModel=%7B%22colId%22:%22Location%22,%22sort%22:%22asc%22%7D>.
- 18 Kaiser Family Foundation (2017). Marketplace Plan Selections by Gender. State Health Facts. Retrieved December 1, 2019, from <https://www.kff.org/health-reform/state-indicator/marketplace-plan-selections-by-gender-2/?dataView=1¤tTimeframe=2&sortModel=%7B%22colId%22:%22Location%22,%22sort%22:%22asc%22%7D>.

Data sources

Milliman 2016 and 2017 consolidated data sets

Milliman Standard Demographic Assumptions – Commercial Insurance

The table in Figure 4 shows the portion of the U.S. population represented by members who were continuously enrolled in 2016 and 2017, by region.

FIGURE 4: PORTION OF POPULATION CONTINUOUSLY ENROLLED, 2016-2017, BY REGION

NORTHEAST	7.2%
MIDWEST	3.0%
SOUTH	4.2%
WEST	3.7%
Total	4.4%

Methodology

We calculated a prevalence rate for each disease, health insurance market, and age range band, and applied it to the number of people insured in each market and age range band.

Prevalence rate = numerator/denominator

Where:

Numerator is the number of people estimated to have a particular condition as described below. For hemophilia, only males were considered.

Denominator is the number of people who were continuously enrolled in 2016 and 2017, by health insurance market and age band. For hemophilia, only males were considered.

For each disease, health insurance market, and age range band, we found people who had one or more of the diagnosis codes shown in the table in Figure 5 within the first five diagnosis codes listed on claims records. The table in Figure 6 contains the age range bands used by the health insurance market. We removed people from the study who were not continuously enrolled during 2016 and 2017. We also removed people who did not have at least one inpatient admit, one outpatient visit (excluding radiology and lab), or two professional visits (again excluding radiology and lab) during 2016 and 2017.

FIGURE 5: ICD-10 DIAGNOSIS CODES

DISEASE	ICD-10 DIAGNOSIS CODES
Rheumatoid arthritis	M05, M06, M08, M120, M45, M488X1, M488X2, M488X3, M488X4, M488X5, M488X6, M488X7, M488X8, M488X9
Ulcerative colitis	K5180, K51811, K51812, K51813, K51814, K51818, K51819, K5190, K51911, K51912, K51913, K51914, K51918, K51919
Hemophilia	D66 and D67

FIGURE 6: AGE RANGE BANDS BY HEALTH INSURANCE MARKET

MARKET	AGE RANGE BANDS
Medicare	0-64, 65-74, 75-84, 85 and over
Commercial and Tricare	0-24, 25-29, 30-34, 35-39, 40-44, 45-49, 50-54, 55-59, 60-64, 65 and over
Individual	0-17, 18-34, 35-44, 45-54, 55-64, 65 and over
Medicaid	0-20, 21-26, 27-45, 46-64, 65 and over

POPULATION WITH EACH DISEASE

To estimate the number of insured people with each disease, we multiplied the prevalence rate for each health insurance market and age range band by the number of people insured in that market and age band. We averaged 2016 and 2017 population counts by health insurance market type. We have excluded uninsured from this study. We applied the commercial prevalence rate to the Tricare population because our data did not allow for separate calculation of prevalence within the Tricare segment.

Limitations

We relied on the diagnosis codes in the administrative claims data. Inaccurate diagnosis codes could change the prevalence rates we've calculated. Many rare diseases have not been assigned ICD-10 diagnosis codes and therefore an analysis of this type is not currently possible. Health insurance market populations may not represent each state in the same proportion as their insured populations. Prevalence rates can change over time as new tests and treatments are developed. Different methodologies could also produce different prevalence rates. We assumed that the distribution of people with each disease resembles a normal distribution when calculating the confidence interval.



Milliman is among the world's largest providers of actuarial and related products and services. The firm has consulting practices in life insurance and financial services, property & casualty insurance, healthcare, and employee benefits. Founded in 1947, Milliman is an independent firm with offices in major cities around the globe.

milliman.com

CONTACT

Donna Wix
donna.wix@milliman.com

David V. Williams
david.williams@milliman.com