HHS Public Access

Author manuscript

Stat Med. Author manuscript; available in PMC 2015 September 30.

Published in final edited form as:

Stat Med. 2012 November 30; 31(27): 3295–3298. doi:10.1002/sim.5340.

Confidence intervals and statistical testing for ratio measures of percent change‡

Carla A. Winston[*,†], **Andrew N. Hill**, **Michael P. Chen**, **Nong Shang**, and **José E. Becerra**

Division of Tuberculosis Elimination, Centers for Disease Control and Prevention (CDC), Atlanta, GA, USA

Abstract

In public health and medical research, ratio measures of percent change relative to baseline are often used to express a change in disease incidence. Estimating variance becomes more complex when the comparison is to an expectation based on previous data (*E*), rather than to an observed value (*O*). In 2009, the decline in reported tuberculosis (TB) cases was the largest single-year decrease since national TB surveillance began in 1953. To investigate the 2009 TB decline compared with expected counts, we analyzed TB cases reported to the Center for Disease Control and Prevention's National Tuberculosis Surveillance System. We log-transformed case counts for 2000–2008, and performed linear regression stratified by patient and clinical characteristics. We calculated relative declines from expectation as $(O-E)/E$ for patient subgroups, and constructed 95% confidence intervals for TB declines. We then formulated a *Z*-score test statistic comparing declines across patient subgroups under the null hypothesis that the difference of the two ratio measures was zero. We illustrate our methods by comparing 2009 declines from expectation for US-born versus foreign-born patients. Predicted values and confidence intervals assessed the magnitude of unexpected TB declines within patient groups, while statistical tests comparing ratio measures evaluated relative TB declines across groups.

Keywords

variance analyses; estimation techniques; infectious disease reporting; linear models; Poisson distribution; maximum likelihood estimates

1. Introduction

In public health and medical research, measures of percent change relative to baseline are often used to express a change in disease incidence, for instance, an increase in disease because of exposure to an infectious agent, or decrease in disease because of a clinical intervention or public health campaign [1–3]. Time is a key predictor for such analyses. The baseline incidence of disease is measured or assumed to be known for comparison with

‡Note: This article is a US Government work and is in the public domain in the USA.

*Correspondence to: Carla Winston, Division of Tuberculosis Elimination, Centers for Disease Control and Prevention [CDC], Mailstop E-10, 1600 Clifton Road NE, Atlanta, GA, USA.. †*cwinston@cdc.gov*.

The authors have no competing interests. The findings and conclusions are those of the authors and do not necessarily represent the views of the Centers for Disease Control and Prevention.

increasing or decreasing incidence over time. Variance of the appropriate baseline disease incidence measure is calculated based on sample or population data.

Estimating variance for relative change measures becomes more complex when the comparison is to an expectation based on previous data, rather than to an observed value. In 2009, the decline in reported tuberculosis (TB) cases was the largest single-year decrease since national TB surveillance began in 1953. To investigate the 2009 decline compared with expected counts, we analyzed TB cases reported to the Center for Disease Control and Prevention (CDC) National Tuberculosis Surveillance System.

2. Estimation of percent change without accounting for trends

Since 1953, US reporting areas have submitted information on incident TB cases to the CDC's National Tuberculosis Surveillance System [4]. Individual case-report data, including characteristics on TB patient country of origin, have been reported to the National Tuberculosis Surveillance System since 1993. In 1993, 29% of TB cases diagnosed in the United States were foreign-born. The proportion of TB cases among foreign-born persons has steadily grown, to 59% in 2009 [4].

The incidence of TB has declined among both US-born and foreign-born persons since 1993; however, TB declines among the foreign-born have been less substantial than among the US-born. While global TB control has led to incidence declines worldwide [5], TB infection prevalence remains high in many countries from which persons diagnosed with TB disease in the United States emigrate. Risk of progression from latent infection to TB disease remains high among foreign-born persons, even after more than five years since US arrival [6].

In the United States, TB case counts in 2009 were the lowest since national reporting began, showing the greatest single-year decrease ever. Although case counts declined an average of 425 cases per year since 2000, the observed decline in 2009 was 1361 cases fewer than in 2008 [4]. The decline in TB case counts was reported as a 14.8% decrease for US-born persons and 10.5% decrease for foreign-born persons [7]. In the next section, we compare the relative decrease in US-born and foreign-born cases, accounting for prior TB trends (Figure 1; data publicly available at http://www.cdc.gov/tb/statistics/reports/2009/table5.htm [Accessed 9/13/2011]).

3. Methods: percent change accounting for trends

We log-transformed observed case counts for 2000–2008, and performed linear regression (overall R^2 = 0:99) to calculate predicted values for 2009 counts. For US-born and foreign-born subgroups, we compared the observed count in 2009 to the expected value based on prediction. We reported the relative decline of observed counts in 2009 (O) from expected values (E) as a percent change by multiplying $(O - E)/E$ by 100. We constructed 95% confidence intervals around percent declines by patient sub- group. We then formulated a hypothesis test for the difference of the relative declines in TB cases among US-born and foreign-born persons. We constructed a Wald test statistic [8] for the null hypothesis that the

difference in relative TB declines between US-born and foreign-born persons in 2009 was zero.

4. Model for decline in cases and formulation of test statistic

For both the US-born and foreign-born populations, we constructed a linear regression model for log cases from 2000–2008 and assumed a Poisson distribution of cases in 2009. Specifically, with Y_{ij} denoting counts,

$$\log \quad Y_{ij} \quad = \quad \beta_{i0}+\beta_{i1}x_j+\epsilon_{ij}, \quad j \quad = \quad 0,\ldots,8$$

$$Y_{i9}\tilde{}\text{Poi} \quad (\lambda_i)\,,$$

where x_0 = 2000, … x_8 = 2008, and index i refers to US-born (i = 0) or foreign-born (i = 1). Within each population, the error terms are normally distributed with zero mean and common variance σ_i^2. The parameters of interest are the relative declines in cases, where x_9 = 2009

$$\Delta_i \quad = \quad \lambda_i/\exp\left(\beta_{i0}+\beta_{i1}x_9\right)-1.$$

For brevity, we denote the observed count for 2009 by $O_i = Y_{i9}$ and the predicted count for 2009, based on regression estimates for 2000–2008 data, by $E_i \quad = \quad \exp\left(\hat{\beta}_{i0}+\hat{\beta}_{i1}x_9\right)$, where the $\hat{\beta}$ terms denote the regression-based maximum likelihood estimators (MLEs). Then O_i and E_i are statistically independent and are distributed as

$$O_i\tilde{}N\left(\lambda_i,\lambda_i\right),$$

$$\log \quad E_i\tilde{}N\left(\beta_{i0}+\beta_{i1}x_9,\sigma_i^2x_9^T\left(X^TX\right)^{-1}x_9\right),$$

where x_i =(1, x_i)^T and X is the 9×2 design matrix with rows x_i^T, i = 0, …, 8. Approximate normality of O_i results from the normal approximation for a Poisson distribution with large mean (case counts ~ 10^3). The MLEs of Δ_i are (O_i − E_i)/E_i.

Confidence intervals (CI) for Δ_i at 100(1 − α)% level are obtained by inverting the Wald test statistic for Δ_i. These are given by

$$\left(O_i/E_i-1\right)\pm z_{\alpha/2}\left(O_i/E_i\right)\left(O_i^{-1}+\hat{\sigma}_i^2x_9^T\left(X^TX\right)^{-1}x_9\right)^{1/2},$$

where $\hat{\sigma}_i^2$ are the regression-based MLEs of the error variances σ_i^2.

Similarly, the Wald test yields a *Z*-score test statistic for the difference of the relative percent declines $\Delta_1 - \Delta_2$

$$Z = [(O_1/E_1) - (O_2/E_2)]/\mathrm{SE},$$

$$\mathrm{SE}^2 = (O_1/E_1)^2 \left(O_1^{-1} + \hat{\sigma}_1^2 x_9^T \left(X^T X\right)^{-1} x_9\right) + (O_2/E_2)^2 \left(O_2^{-1} + \hat{\sigma}_2^2 x_9^T \left(X^T X\right)^{-1} x_9\right).$$

5. Results: percent change in TB cases by US and foreign origin of birth

Among the US-born, the observed count of 4571 TB cases was compared with an expected 4918, for a difference of 347 fewer cases than expected in 2009. Among foreign-born persons, the observed count of 6854 cases in 2009 was 855 fewer cases than the expected 7709 count.

We calculated the relative decline from expected case counts, $(O - E)/E$, by TB patient origin. Estimated percent declines and 95% confidence intervals for the US-born and foreign-born were −7.1% (−11.4, −2.8) and −11.1% (−14.7, −7.4), respectively. We calculated the *Z*-test statistic comparing US-born and foreign-born relative declines from expected as $Z = 1.66$ ($p = 0.10$).

6. Discussion

Because predicted values were obtained from regression rather than a known population estimate, we were unable to use typical variance calculations for observed compared with expected, sometimes described as standardized morbidity or mortality ratios [9]. In the case of standardized mortality ratios, the expected count is calculated from such a large sample, or from a complete population census, as to have no variance. The observed value is then assumed for variance calculations to be distributed as a Poisson random variable with parameter equal to the expected count. This assumption was inappropriate because our expected count, *E*, was a predicted value based on the random variable of observed counts, *O*; *E* has variability around its estimate. Instead, we calculated variance around *O* assuming a Poisson with parameter *O*, and used the log-transformed normal variance of *E*. To assess the robustness of our assumption, we examined the test statistics, which resulted when we used exact expressions for mean and variance from the log-normal distribution for *E*. Numerical differences were negligible; therefore, in the interests of simplicity, we chose to present Wald statistics.

Our example describes differences by TB patient country of origin. The magnitude of the foreign-born decline relative to expectation (855 cases) was more than twice the magnitude of the US-born decline (347 cases). This reflects the demographics of TB cases diagnosed in the United States, of whom almost 60% are foreign-born. However, percent declines relative to expectation were useful for standardizing across groups, to account for differences in TB trends. Accounting for 2000–2008 trends in declining TB incidence was valuable in comparison to interpreting declines comparing 2009 to 2008 cases only, because failure to

account for differences in trends yielded estimates of a 14.8% decrease among US- born persons and a 10.5% decrease among foreign-born persons in 2009, compared with 7.1% and 11.1% declines, respectively [7]. When analyzing deviation from the expected, the 11.1% decline among the foreign-born was not significantly different from the 7.1% decline among the US-born; both patient strata by origin had significant declines in TB incidence in 2009. Predicted values and confidence intervals evaluated the magnitude of unexpected TB declines within groups, while statistical tests of percent change relative to expected counts were used to compare groups.

Acknowledgements

The authors thank local health department staff who reported data included in these analyses. The authors also acknowledge the Division of Tuberculosis Elimination Weekly Statistics Meeting Workgroup, and Surveillance, Epidemiology, and Outbreak Investigations Branch. Routine data in these analyses are in the public domain.

References

1. Greenberg, RS.; Daniels, SR.; Flanders, WD.; Eley, JW.; Boring, JR. Medical Epidemiology. 3rd. Lange/McGraw-Hill; New York: 2001. p. 35-38.
2. Rothman, KJ.; Greenland, D. Modern Epidemiology. 2nd. Lippincott-Raven; Philadelphia: 1998. p. 51-52.
3. Zhou F, Harpaz R, Jumaan AO, Winston CA, Shefer A. Impact of varicella vaccination on health care utilization. JAMA. 2005; 294(7):797–802. DOI: 10.1001/jama.294.7.797. [PubMed: 16106004]
4. Centers for Disease Control and Prevention: Reported tuberculosis in the United States, 2009. U.S. Department of Health and Human Services, CDC; Atlanta, GA: Oct. 2010
5. World Health Organization: Global tuberculosis control: WHO report 2010. World Health Organization; Geneva: Nov. 2010
6. Cain KP, Benoit SR, Winston CA, MacKenzie WR. Tuberculosis among foreign-born persons in the United States. JAMA. 2008; 300(4):405–412. DOI: 10.1001/jama.300.4.405. [PubMed: 18647983]
7. Winston C, Pratt R, Armstrong L, Navin TR. Decrease in reported tuberculosis cases – United States, 2009. MMWR. 2010; 59(10):289–294. [PubMed: 20300055]
8. Shao, J. Mathematical Statistics. 2nd. Vol. 433. Springer-Verlag; New York: 2003.
9. Xu, JQ.; Kochanek, KD.; Murphy, SL.; Tejada-Vera, B. Deaths: Final data for 2007. National vital statistics reports. Vol. 58. Vol. 130. National Center for Health Statistics; Hyattsville, MD: May. 2010

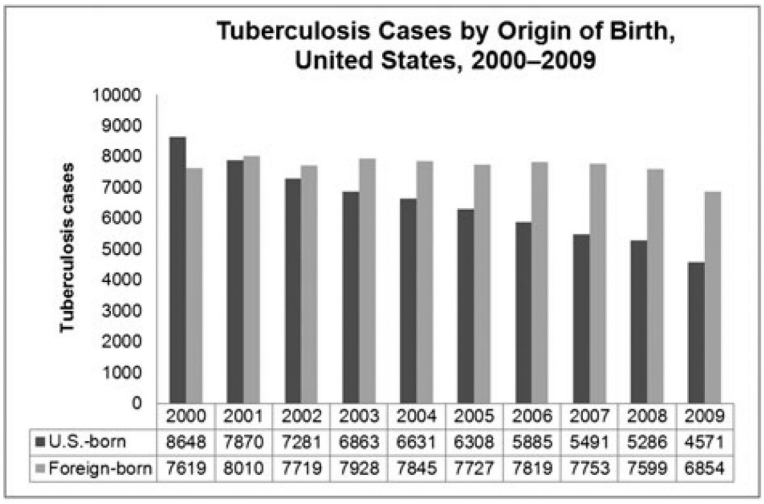

Figure 1.
Tuberculosis cases by origin of birth, United States, 2000–2009.