Statistical Analysis of Case Recovery Rate

Edith Uzoma Umeh*, Cyprian Anene Oyeka
Department of Statistics, Nnamdi Azikiwe University, Awka

Abstract
This paper propose to develop statistical measures of case recovery rate assuming that patient recovery is influenced by two factors acting simultaneously, that is, to develop a two way analysis of case recovery rate for a disease. The specific rates which are useful in identifying segments of the population that are at elevated risk of a given disease and would need a more aggressive medical and health interventionist measures to alleviate the disease compared with the average number of the general population at risk were estimated. Case recovery rate has shown to possibly serve as a measure of morbidity of a disease in a population especially if the disease is not virulent. It is shown that the smaller this rate is, the more likely is the disease morbid in the population and vice versa. It is also shown under certain condition that case recovery rate when used as a measure of survivorship or morbidity from a disease must be used and interpreted with great caution because of definitional problems. The propose method was illustrated with some data.
Keywords: Case fatality, Recovery rate, Crude rate, Adjusted and Unadjusted rates, Treatment, Survivorship, Morbidity

1. Introduction

The concept of case fatality rate, the number of deaths due to a certain disease per infected population at risk during a specific time period is often used as a gauge of the virulence of a disease and of the success of its treatment and management strategies. However, the concept of case recovery rate have been defined as the number of persons who recovered from a certain disease per population at risk marked by the cessation of the relevant symptoms during a specified period of time. (Zar 1984, 1996; Fleiss, 1981, Cohen 1988). It may also be used as a measure of the success of the disease treatment and management program. (Crawford et al 1992, Yeates and Taylor 1997). This may be a more relevant measure especially for life threatening highly virulent severe and acute disease outbreak including AIDS/HIV and Ebola.

(Daniel 1983; Cole and MacMahon 1971; Cook,, Doll. and Fellingham 1969; Hennekens and Buring 1987, Cicchetti et al 1991, 1993 , Rourke, et al 1992). The prognosis of a disease is usually dependent on several factors including the demographic and socio-economic characteristics of the patient as well as the length of time that elapsed between disease onset and hospitalization and available medical and health management technologies. These variables are usually associated with the occurrence of the event of interest as well as with one another. **(**Kelly and Cowling 2013; *Taubenberger Jeffrey and David 2006;* Li et al 2008; King 2008, Cornfield 1951, Crawford, and Allan 1997 **).**

In this paper, we propose to develop statistical measures of case recovery rate assuming that patient recovery is influenced by two factors acting simultaneously, that is, we propose to develop a two way analysis of case recovery rate for a disease.

1. 1: Propose Method

Let A with 'a' levels or strata and B with 'b' levels or strata be any two factors influencing the outcome of a certain disease during a specified time period. Let n_{ij} be the number of persons at the ith level of factor A and jth level of factor B in the population at risk who contacted the disease within a specified time period, for $i = 1,2,,\ldots,a \ and \ j = 1,2,\ldots,b. \ Let \ f_{ij}$ be the corresponding number of persons who recovered from the disease within the specified period of time.. These two numbers may be written as the pair (f_{ij}, n_{ij}). The factor A by factor B case specific recovery rate $r_{ij,}$, at the ith level of factor A and jth level of factor B is then given

$$r_{ij} = \frac{f_{ij}}{n_{ij}}, \ i = 1,2,\ldots,a; j = 1,2,\ldots,b \quad \ldots\ldots\ldots\ldots\ldots\ldots 1$$

Let $n_{i.} = \sum_{j=1}^{b} n_{ij}$ be the number of persons in the population at risk infected by the disease at the ith level of factor A for all levels of factor B and Let $n_{.j} = \sum_{i=1}^{a} n_{ij}$ the number of persons in the population at risk infected at the jth level of factor B for all levels of factor A. The total number of persons in the infected population at risk of recovery or the infected population is $n = \sum_{i=1}^{a} n_{i.} = \sum_{j=1}^{b} n_{.j} = \sum_{i=1}^{a}\sum_{j=1}^{b} n_{ij}$

Let $p_{i.} = \frac{n_{i.}}{n}$ be the proportion of the population at risk infected at the ith level of factor A and $p_{.j} = \frac{n_{.j}}{n}$ be the proportion of the infected population at risk infected in the jth level of factor B. Also let $p_{ij} = \frac{ni_j}{n}$ be the proportion of the population at risk infected in the ith level of factor A and the jth level of factor B where

$$\sum_{i=1}^{a} p_{i.} = \sum_{j=1}^{b} p_{.j} = \sum_{i=1}^{a}\sum_{j=1}^{b} pi_j = 1$$

Let $f_{i.} = \sum_{j=1}^{b} f_{ij}$; $f_{.j} = \sum_{=1}^{a} f_{ij}$ ……………2

be respectively the number of persons who recovered from the disease at the ith level of factor A during the specified time period and the jth level of factor B who recovered from disease during the specified time period.

Let $f = \sum_{i=1}^{a} f_{.i.} = \sum_{ij=1}^{b} f_j = \sum_{i=1}^{a}\sum_{ij=1}^{b} f_{ij}$ ……………3

be the total number of persons in the population at risk who recovered from the disease during the specified time period.

Then the unadjusted factor A specific recovery rate at the ith level of factor A, namely $r_{i.}$ is calculated as

$$ri. = \frac{f_{i.}}{n_{i.}} = \frac{\sum_{j=1}^{b} f_{ij}}{n_{i.}} = \sum_{j=1}^{b} \frac{pi_j ri_j}{p_{i.}} \qquad …………………4$$

Similarly, the unadjusted specific recovery rate at the jth level of factor B is given as

$$r_{.j} = \frac{f_{.j}}{n_{.j}} = \frac{\sum_{i=1}^{a} fi_j}{n_{.j}} = \sum_{i=1}^{a} \frac{pi_j ri_j}{p_{.j}} \qquad …………………5$$

Also, the overall total or crude unadjusted recovery rate from the disease during the specified time period is

$$r = \frac{f}{n} = \frac{\sum_{i=1}^{a} n_{.i} r_{i.}}{n} = \frac{\sum_{j=1}^{b} n_{.j} r_{.j}}{n} \qquad … … … … … … … … … 6$$

The specific rates in equations 1-3 are useful in identifying segments of the population that are at elevated risk of a given disease and would therefore be in need of more aggressive medical and health interventionist measures to alleviate the disease compared with the average number of the general population at risk.

Further interest here is also often to obtain rates adjusted for the two factors of classification. As noted above, factors A and B are usually closely associated with the outcome of the disease as well as with each other. There are therefore likely to have confounding effects on the estimated rates. The rates standardization or adjustment is aimed at removing some of these effects. If the distribution of the population by the levels of the factors of classification is available as in the present case, then the method of standardization that readily suggests itself is the direct method.

To apply the direct method of standardization, we use the overall distribution of the population at risk over the categories of factor B namely $p.j$; $j = 1, 2, \ldots, b$ as the standard population for adjusting the specific rates for the levels of factor A. Similarly, we use the overall distribution of the population at risk over the levels of factor A namely $pi.$; $i = 1, 2, \ldots, a$, as the standard population to apply in adjusting its specific rates for the levels of factor B.

Thus, the direct standardized or adjusted recovery rate for the ith level of factor A namely $r_{i.}$ direct is calculated as

$$r_{i.}.direct = \sum_{j=1}^{b} p_{.j} r_{ij}; \quad j, 1, 2, \ldots, b \qquad … … … … … … … 7$$

Similarly, the direct adjusted rate for the jth level of factor B namely $r_{.j}$ direct is given as

$$r_{.j}\ direct = \sum_{i=1}^{a} p_{i.} r_{ij}, \quad i = 1, 2, \ldots, a \qquad … … … … … … … 8$$

The direct adjusted crude rate, r adj. direct resulting from equations 7 and 8 may be equal but are not necessarily equal to the unadjusted crude rate r obtained from equation 6 .

The format of the calculation is presented in the following table.

Journal of Natural Sciences Research www.iiste.org
ISSN 2224-3186 (Paper) ISSN 2225-0921 (Online)
Vol.6, No.12, 2016

Table1. Schematic Table Illustrating the Calculation of Specific and Direct Adjusted Rates in a Two - Way Classification

Factor A	1	2	3	…	b	Proportion of Popn $p_{i.}$	Unadjusted rate$(r_{i.})$	Direct adjusted rate $r_{i.}$ Direct
1	r_{11}	r_{12}	r_{13}	…	r_{1b}	$p_{1.}$	$r_{1..}$	$r_{1.}$ direct
2	r_{21}	r_{22}	r_{23}	…	r_{2b}	$p_{2.}$	$r_{2.}$	$r_{2.}$direct
3	$r_{31.}$	$r_{32.}$	r_{33}	…	$r_{3b.}$	$p_{3..}$	$r_{3..}$	$r_{3..}$direct
.	.							
A	$r_{a1.}$	r_{a2}	$r_{a3.}$	…	$r_{ab.}$	$p_{a..}$	$r_{a..}$.	$r_{a..}$direct
Proportion of Popn $p_{.j}$	$p_{.1}$	$p_{.2}$	$p_{.3}$		$p_{.b}$		$r_{.j}$	
Unadjusted rate $r_{.j}$	$r_{a1.}$	$r_{a2.}$	$r_{a3.}$		r_{ab}	$p_{a..}$	$r_{a..}$	$r_{a..}$direct
Direct adjusted rate $r_{.j}$ Direct	$r_{.1.}$direct	$r_{.2.}$direct	$r_{.3.}$direct	…		$p_{.b.}$direct		

Table 2 : Distribution of Meningitis Patient and Number who recovered within two weeks of Treatment by Age of patients and type of Medication. The 1st entry is the number of people who recovered and the 2nd entry is the number of meningitis patients treated for 2 weeks

Age	Natural herbs	Standard Drug	New Drug	Total
< 10	5; 72	10; 130	13; 150	28; 352
10-19	5;76	4; 59	7; 85	16; 220
20-29	6; 114	4; 70	9; 108	19; 292
30-39	3; 53	4;49	5;56	12;158
40-49	4; 67	8; 105	4;52	16;224
50+	4; 60	5;63	7;78	16'201
Total	29; 442	35; 476	45;529	107;1447

The age by type of treatment specific rate are obtained by applying equation 1 to table 2. The results are presented in Table 3.

Table 3: Age by Type of Treatment Meningitis Recovery Rate per 1000 Patients

Age	Natural herbs	Standard Drug	New Drug	Total
< 10	69.4	76.9	86/7	79.5
10-19	65.8	67.8	82.4	72.7
20-29	52.6	57.1	83.3	65.1
30-39	56.6	81.6	89.2	76.9
40-49	59.7	76.2	76.9	71.4
50+	66.7	5;63	89/7	79.6
Unadj type of treatment specific (r.j)	65.6	73.5	85/1	73.9 (r)

As expected, recovery is dependent on type of medication, increasing from 65. 6 for natural herbs to a high level of 35.1 per thousand patients treated with new drug. However, age specific recovery rates do not seem to have a consistent pattern with age. This finding will seem to indicate the need to develop appropriate strategies to re-orientate patients who tend to prefer use of natural herbs.

Using equation 5, we calculate the direct adjusted specific recovery rates by age groups using the distribution of patients across types of treatment as the standard population. The calculation is shown in Table 4

Table 4: Calculation of Direct Adjusted Specific Recovery Rates by Age using Distribution of Patients by Types of Treatment as Standard Population (in Thousands**)**

Age	Natural Herbs	Standard Drug	New Drug	Unadjusted Rate $(r_{i.})$	Direct Adjusted $(r_{i.}\ direct)$
$\ll 10\ r_{1j}$	69.4	76.9	86.7	79.5	
$p_{.j}r_{1j}$	21.2	25.3	31.7		78.2
$10-19\ r_{2j}$	65.8	67.8	82.4	72.7	
$p_{.j}r_{2j}$	20.1	22.3	30.2		72.6
$20-29\ r_{3j}$	52.6	57.1	83.3	65.1	
$p_{.j}r_{3j}$	16.0	18.8	30.5		65.3
$30-39\ r_{4j}$	56.6	81.6	89.2	76.9	
$p_{.j}r_{4j}$	17.3	26.8	32.6		76.8
$40-49\ r_{5j}$	59.7	76.2	76.9	71.4	
$p_{.j}r_{5j}$	18.2	25.1	28.2		71.5
$50+\ r_{6j}$	66.7	79.4	89.7	79.6	
$p_{.j}r_{6j}$	20.3	26.1	32.8		79.2
Total r_{tj}	65.6	73.5	85.1	73.9	
p_{6j}	20.0	24.2	31.2		75.3
Proportion of Patients in Treatment Type	0.305	0,329	0.366		

Note that adjusted of age specific recovery rate for type of treatment slightly reduces these rates but not significantly, indicating that the type of treatment the patient receives may not be importantly associated with age of patient.

Finally, the direct adjusted specific recovery rate by type of treatment using the distribution of affected patients by age as standard population is calculated from equation 6 as the results are presented in Table 5.

Table 5: Calculation of Direct Adjusted Specific Recovery Rates by Treatment using the Distribution of Patients by Age as Standard Population (in thousands)

Age	Natural Drug		Standard Drug		New Drug		Total		Proportion of Patients in Age Group (Standard Popn)
	r_{i1}	$p_{i,}r_{i1}$	r_{i2}	$p_{i,}r_{i2}$	r_{i3}	$p_{i,}r_{i3}$	r_{it}	$p_{i,}r_{it}$	
$\ll 10$	69.4	16.9	76.9	18.7	86.7		79.5	19.3	0.243
$10-19$	65.8	10.0	67.9	10.3	82.4		72.7	11.1	0.152
$20-29$	52.6	10.6	57.1	11.5	83.3		65.1	13.2	0.2
$30-39$	56.6	6.2	81.6	8.9	89.2		76.9	8.4	0.109
$40-49$	59.7	9.3	76. 2	11.8	76.9		71.4	11.1	0.155
$50+$	66.7	9.3	79.4	11.0	89.7		79.6	11.1	0.139
Unadjusted Specific $r_{.j}$	65.6		73..5		85.1				
Direct Adj Specific rate $r_{i1}\ direct$		62.3		72.2				74.2	

It is observed that when type of treatment specific recovery rates is adjusted for age of patients, their values are slightly reduced for each treatment type, although the observed patient still remains.

1.2: Conclusion

We have in this paper proposed the concept of case recovery rate defined as the number of infected persons who recovered from a specific disease per population of persons infected with the disease within a specified time period as a public health indicator. Both the unadjusted and adjusted factor specific rates help the public health worker to determine those factors and combination of factors that would need special public health and medical attention.

As noted earlier, case fatality rate and case recovery rate may be used to measure the success of a disease management program in controlling disease. However, unlike case fatality rate, case recovery rate can also serve as a measure of morbidity of a disease in a population especially if the disease is not virulent. The smaller this rate is, the more likely is the disease morbid in the population but if the recovery rate is high, the chances that the disease is morbid may be small.

Never the less, unlike case fatality rate which has an element of finality, case recovery rate when used as a measure of survivorship or morbidity from a disease must be used and interpreted with great caution because of definitional problems. Recovery is often a continuous process that may be prolonged especially with increasing age and compounded by other associated conditions that may affect recovery. However, if recovery is identified with certain physical factors such as resumption of usual economic and social activities with the cessation of some well defined primary symptoms, then the resumption of these activities and the cessation of these symptoms may be regarded as necessary and fairly sufficient indication of recovery. Under these conditions, case recovery rate may serve as a good public health indicator.

References

Centers for Disease Control and Prevention (2009): FoodNet Facts and Figures - Case Fatality Ratios. CDC 2009

Cicchetti, D.V, and Rourke, B.P (1991); Methodological and Biostatical foundations of Clinical Neuropsychology and Medical and Health Disciplines (2nd ed. Psychology Press, Taylor and Francis group, Hove and Newyork

Cicchetti, D.V, Shawalter, D., Rourke, B.P and Fuerst, D. (1992); A computer program for analyzing Ordinal Trends with dichotomous Outcomes: Applications to Neuropsychological Research. The Clinical Neuropsychologist Vol 6, 4 458-463

Cicchetti, D.V, Shawalter, D. and Weder, B.E,. (1993); A computer program for performing meta analysis when the outcome variable is dichotomous: Relevance to Neuropsychology and Biomedical Research . The Clinical Neuropsychologist 4, 454-459

Cohen , J. (1988); Power analysis for the behavioural Sciences (2nd ed). Englewood Cliffs NJ Eribaum Associates

Cole, P. and MacMahon, B. (1971): Attributable risk percent in case-control studies. Br. J. prev. soc. Med., 25, 242-244

Cook, P., Doll, R. and Kellinger, S.A. (1969): A mathematical model for the age distribution of cancer in man. Int. J. Cancer, 4, 93-1 12

Cornfield, J. (1951): A method of estimating comparative rates from clinical data. Applications to cancer of the lung, breast and cervix. J. Natl Cancer Inst., 11, 1269-1275

Crawford, J.R, Moore, J.W, and Cameron I.M (1992): Verbal fluency : A NART- based equation for the estimation of parameter performance , British Journal of Clinical Psychology 31, 321-329

Crawford, J.R, and Allan K.M (1997): Estimating pre-morbid WAIS-R-IQ with demographic variables: Regression equations derived from a UK sample. The Clinical Neuropsychologist 11, 192-197

Daniel, W. W.(1983): Biostatistics: A foundation for analysis in the health sciences, New York, Wiley

Hennekens CH, Buring JE. Epidemiology in Medicine, Lippincott Williams & Wilkins, 1987

Kelly, Heath; Cowling, Benjamin J (2013).: Case Fatality: Rate, Ratio, or Risk? Epidemiology: July 2013 - Volume 24 - Issue 4 - p 622–623

King, John W (April 2, 2008). "Ebola Virus". eMedicine. WebMd. Retrieved 2008-10-06

Li, F C K; B C K Choi; T Sly; A W P Pak (June 2008). "Finding the real case-fatality rate of H5N1 avian influenza". Journal of Epidemiology and Community Health **62** (6): 555–559.

National Center for Emerging and Zoonotic Infectious Diseases(NCEZID) 2015 http://www.cdc.gov/ncezid/

Rourke, B.P, Costta, L., Cicchetti, D.V, Adams, K.M. and Plasterk, K.J (1992): Methodological and Biostatistical foundations of Clinical neuropsychology, Lisse, The Netherlands: Swets and Zeitlienger

Taubenberger, Jeffery K.; David M. Morens (January 2006). "1918 influenza: the mother of all pandemics". Emerging Infectious Diseases (Coordinating Center for Infectious Diseases, Centers for Disease Control and Prevention) **12** (1): 15–22.

Yeates, K.O and taylor, H, g. (1997): Predicting Pre-Morbid neuropsychological functioning following pediatric traumatic brain injury. Journal of Clinical and \experimental *neuropsychology, 19, 825-837*

Zar, ,J..H. (1984): Biostatistical Analysis (2nd ed.) Englewood Cliffs, NJ, Prentice- Hall

Zar, ,J..H. (1984): Biostatistical Analysis (3rd ed.) London, Prentice- Hall