VIA MEDICA

REVIEW ARTICLE

Int Marit Health
2019; 70, 3: 187–192
DOI: 10.5603/IMH.2019.0029
www.intmarhealth.pl

ISSN 1641–9251

How to calculate incidence rates from proportionate data

Olaf Chresten Jensen[1, 3, 4], Agnes Flores[2], Despena Andrioti Bygvraa[1], Fereshteh Baygi[1], George Charalambous[3]

[1]Centre of Maritime Health and Society, Institute of Public Health, University of Southern Denmark, Esbjerg, Denmark
[2]Caja Seguro Social, Rep. of Panamá, Vacamonte, Panama
[3]Graduate School, Frederick University "Nicosia", Cyprus
[4]University of Panamá, Faculty of Health, Panama

ABSTRACT

This paper describes the methodological aspects of calculation of incidence rates from incomplete data in occupational epidemiology. Proportionate measures in epidemiological studies are useful e.g. to describe the proportion of slips, trips and falls compared to other types of injury mechanisms within single age-strata. However, a comparison of proportions of slips, trips and falls among the different age-strata gives no meaning and can hamper the conclusions. Examples of a constructed example and some selected studies show how estimates of incidence rates can be calculated from the proportionate data by applying estimates of denominators available from other information. The calculated examples show how the risks based on the incidence rates in some cases differ from the risks based on the proportionate rates with the consequence of hampering the conclusions and the recommendations for prevention. In some cases the proportionate rates give good estimates of the incidence rates, but in other studies this might cause errors. It is recommended that estimates of the incidence rates should be used, where this is possible, by estimation of the size of the population. The paper is intended to be useful for students and teachers in epidemiology by using the attached Excel training file.

(Int Marit Health 2019; 70, 3: 187–192)

Key words: epidemiology, injury, occupational, proportionate, maritime

DEFINITIONS

The basic measurements in epidemiology are incidence and prevalence rates. The unique feature of the incidence and prevalence rates is that they express a risk that is needed to know for prioritisation of the prevention. Further, the risk expressed as the incidence and prevalence is comparable between different populations and between segments of the population, *e.g.* between the age groups [1]. To obtain the incidence and prevalence rates we need two types of information, one is the denominator, the number of persons or better the sum of person-time at risk and the number of incidents, injuries or diseases, the numerator. In most studies we have access to both types of data but sometimes we have no access to the denominator, the number of persons (time) at risk and then we can only calculate the proportionate rate of e.g. per cent of slips trips and falls. The proportionate measures are useful as the first step of the analysis of the risk but they are not comparable directly within different age-groups, which is the problem we address here.

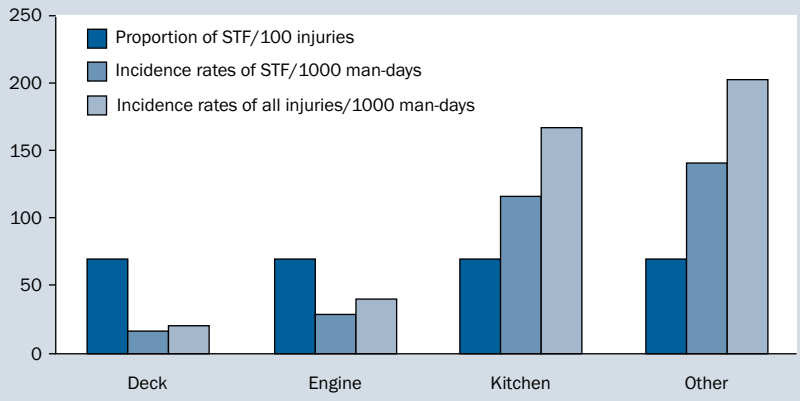
Proportionate measurements or, with other words, the percentages in epidemiology are very useful to study the overall distribution of the variables. However the use of only the percentages is of limited use for identification of the relative risks needed for the prioritisation of the prevention. As the percentages do not present the relative risks, they cannot be used for comparison of the risks in different strata of the population and for comparison with other studies. Still, the proportionate rates are used and give meaning, *e.g.*, for comparison of the percentages of slips, trips, and falls *within* the single-study strata, such as the age-groups. Also the proportional mortality ratio, which is the percentage

Dr. Olaf Chresten Jensen, Centre of Maritime Health and Society, Institute of Public Health, University of Southern Denmark, N Bohrs Vej, 6700 Esbjerg, Denmark, e-mail: ocj@health.sdu.dk

$$(P_{\text{Injury type y in stratum x}}) \times IR_{\text{(All injuries in stratum x)}} = IR_{\text{(Injury type y in stratum x)}}$$

P = The proportion of a specific type of disease or injury y in stratum x
IR = Incidence rate per 10^n man-days

Figure 1. Mathematical relationship between the proportions and the incidence rates



Constructed example		Deck	Engine	Kitchen	Other
1	Man-days at risk	50000	25000	6000	5000
2	Total number of injuries	1000	1000	1000	1000
3	Number of STF-related injuries	700	700	700	700
4	Proportion of STF/100 injuries	70%	70%	70%	70%
5	Incidence rates of STF/1000 man-days	14	28	116.7	140
6	Incidence rates of all injuries/1000 man-days	20	40	166.7	200

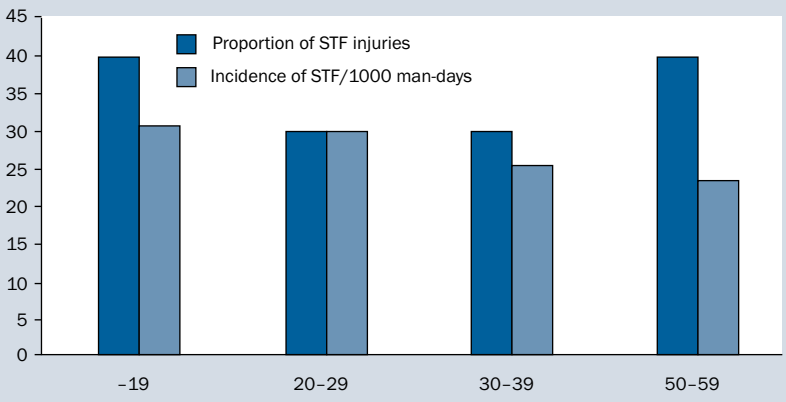
The man-days at risk, the number of injuries and the proportions of slips, trips, and falls (STF) in four departments are chosen in order to illustrate the problem most clearly. Incidence rate of STF in row 5 for Deck = (700/50000) × 1000 = 14. Incidence rates of all injuries in row 6 for Deck = (1000/50000) × 1000 = 20. Based on row 6, the incidence rate of all injuries per 1000 man-days = 20 and the proportion of STF = 70% (0.70); the incidence rate of STF can be recalculated = 0.70 × 20 = 14/1000 man-days.

Figure 2. Constructed example of slips, trips and falls' proportions and incidence rates

of all deaths due to a specific disease in the population during a time period, is used with good meaning. The incidence rate ratio defined as the ratio of two incidence rates gives meaning. However, a direct comparison of the proportions of slips, trips, and falls *among* the strata such as for example different age groups provides meaningless information or can even produce erroneous results. Incidence rates and the relative risks expressed by the incidence rate-ratios are needed to prioritise the prevention. While the proportions, especially on injury data, are still presented in scientific publications, the limitations of its use have not so far been questioned. As we have observed some potential errors related to this problem in different studies, we found it important to inform the readers interested in maritime medicine about these methodological aspects of injury risk analysis. The constructed example and some examples from published studies show how the comparable incidence rates for the relative risks can be calculated from the proportions by applying estimates of the populations at risk in the specific strata of the populations. Based on the constructed example and the examples from the published studies, a mathematical relationship of the proportions and the incidence rates has been explained as in Figure 1. Statistical 95% confidence levels are calculated by comparing 2 person-time rates using *e.g.* the Open-Epi programme [2].

CONSTRUCTED EXAMPLE

The constructed example gives a clear illustration of how the incidence rates can be calculated from proportionate data by applying estimates of the size of the populations at risk (Fig. 2). The denominators, *e.g.*, national population data, are applied (in this case the number of man-days at



	–19	20–29	30–39	50–59
Number of man-days at risk — estimated	13000	8000	8000	12000
Total number of injuries*	1000	800	700	700
Incidence rates of all injuries/1000 man-days	77	100	88	58
Number of STF-related injuries*	400	240	210	280
Proportion of STF-related injuries	0.40	0.30	0.30	0.40
Incidence of STF-related injuries/1000 man-days	31	30	26	23

The proportions of slips, trips, and falls (STF) in four age groups are based on the study data. The man-days at risk and the total number of injuries are estimated in order to illustrate the problem most clearly. Incidence rate of all injuries, age –19 = (1000/13000) × 1000 = 77 per 1000 man-days. The incidence rate of STF injuries, age –19 can be calculated = 0.40 × 77 = 31 per 1000 man-days.

Figure 3. Slips, trips and falls-related injury proportions and incidence rates by age [3]

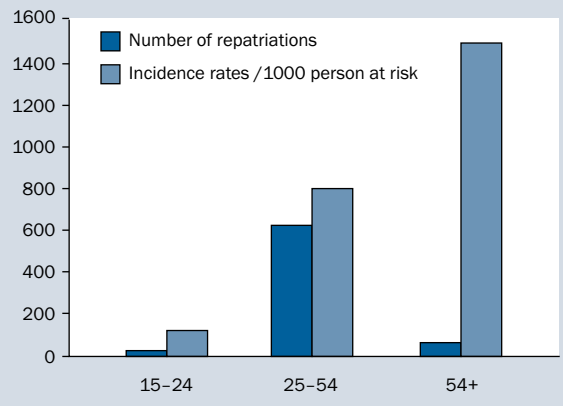
risk). In this constructed example the sizes of the populations at risk in four departments are selected to illustrate the problem most clearly and the proportions of slips, trips, and falls (STF) are chosen to be equal 70% also for the best illustration.

The incidence risks of STF-related injuries based on man-days across the departments are significantly different from the proportionate data; percentages are almost equal in the four groups, while the incidence rates vary significantly. The number of cases and the proportions of the specific sub-groups of cases need to be available together with estimates of the population data. The risks in the strata are calculated by multiplying the incidence rate for all types of injuries in the specific strata (*e.g.*, all injuries on Deck = 20/1000 man-days) by the proportion of STF. In this example the proportion of STF-related injuries was chosen to be 70%, and the incidence rate of STF injuries in this strata = 0.7 × 20 = 14/1000 man-days. Concerning the ethical issues, no personal information is included, so approval from the Ethics Committee or written informed consent was not necessary.

STUDY EXAMPLES

The first example is an analysis of the proportional STF-related injuries in a sample of data for 582 cases of injured commercial fishermen from an emergency department. Information about the number of all injuries and the number of STF-related injuries was available, but there was no information about the population at risk [3]. The proportions of injuries from STF by age was U-shaped and constituted around 40% for men under 20 and over 50, and around 20% for those between these ages. These estimates do not reflect the relative risk for STF-related injuries in the age groups. However, incidence rates for comparison can be calculated by using estimates of the person-time in the age groups as shown in Figure 3.

This type of methodological problem was seen in a study about repatriations of seafarers from the sea. The numbers of repatriations are taken from Figure 3 in the article [4]. There is no information about the age structure among the Filipino seafarers, only the number of seafarers and the number of repatriations. In an attempt to calculate the best estimates of the incidence rates, an estimate of the population of seafarers at risk was calculated by multiplying the percentages for each age group. The Filipino Population Pyramid 2017 for men was applied to calculate the approximate incidence rates [5]. The pattern of the numbers is clearly different from the incidence rates of repatriations with the impact on the results and conclusions (Fig. 4). In the same study, the



Age-groups	15–24	25–54	54+
Persons at risk estimated Filipino Population Pyramid 2017	25000	80000	5000
Number of repatriations	30	650	75
Incidence rates/1000 persons at risk	120	813	1500

The numbers of repatriations in the age-groups are based on the study data. The persons at risk are estimated based on the Filipino Population Pyramid 2017. Incidence rate in the age group 15–24 = (30/25000) × 1000 = 120 repatriations per 1000 persons at risk. There are clear differences between the number of repatriations and repatriation incidence rate in relation to age.

Figure 4. Number of repatriations compared to incidence rates by age groups [4]

length of stay on board and the repatriation rates of seafarers were estimated by the absolute numbers of repatriations related to 150 and 250 days stay on average on board [4]. The relationship between the incidence rate-ratio and the days on board was very different from the relationship between the absolute number of repatriations and days spent on board (Fig. 5). The different results have an important impact on the recommendations given concerning the obligatory routine health examinations in shipping and the safety.

DISCUSSION

This is to our knowledge the first contribution to the solution of the methodological problems with proportionate risk estimates. There is some methodical similarity but this is not the same as the indirect and direct age-standardised methods that were used *e.g.* in the studies of fatal injury in seafaring and fishing [6, 7]. In these studies the observed numbers of fatalities for each age-group was summed up and compared with the expected numbers to yield the standardised mortality ratios. The expected number of fatalities in the reference population was calculated in the age-group by multiplication of the fatality incidence rate with the number of person-years in the index group (seafarers and fishermen, respectively) in the calendar periods. In the direct standardisation we calculate the observed and the expected number of fatalities by multiplication of the fatality rate in the age groups with a standard selected risk population. For both methods, we need to have the information about the incidence rates in each age-group. Our method presented here is a step before, as we don't have the incidence rates and this is what we want to obtain. To do this we need the overall incidence rate of all injuries to calculate the age-specific incidence rates by multiplication of the per cent of slips, trips and falls in the age-groups by the overall incidence rate of all injuries. That also means that the method of age-standardised cancer ratio or more specific for the topic, age-standardised injury ratio, would only be applicable here as the second step for rate-ratio calculations by standardisation or by the logistic regression analyses.

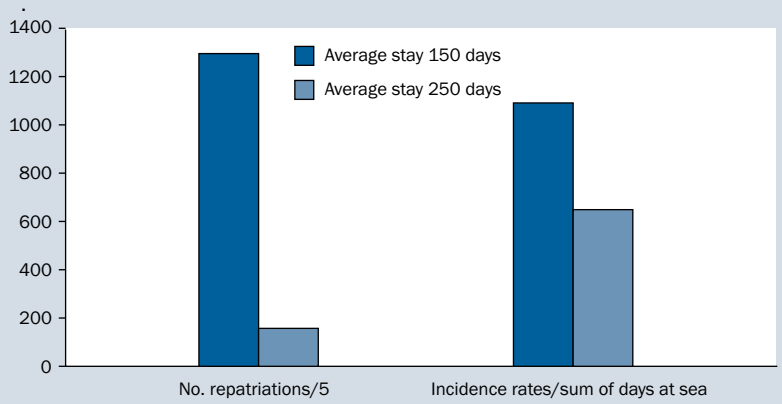
In the study example on slips, trips, and falls-related injuries on fishing vessels, the information on the denominators in the age groups was not available (Fig. 3). The methodological weakness of the study using only proportionate measures of STF in the age groups was not discussed either [3]. However, an estimation of the distribution of the number of days at sea in the age groups among the fishermen could have been applied based on an earlier epidemiological study on injuries among fishermen [8]. The numbers in Figure 6 are adjusted to show that the risk differences of STF-related injuries in the age groups may be even greater than the differences seen in the proportionate measures.

In a Mexican study the overall incidence rates of fatal industrial injuries declined from 1980 to 1995. However, the risks of fatal injuries in the industrial branches (15% in construction, 14% in oil and gas production, and 11% in farming) were based on the proportionate measures and then less useful for the prevention [9].

In a study of repatriations of seafarers, those from India were more frequently repatriated than the Philippines based on proportionate rates [10]. But if information of the populations at risk were applied to calculate estimates of the incidence rates, then the Philippines might have less repatriation incidence rates than the Indians.

In a recent study from Latin America on snakebites, the data were collected from hospital emergency room records. The authors argue that farmers are most affected [11]. But as the occupations of the patients were not registered in the hospital no relative risks for agricultures or other occupations were calculated. Further, when proportionate measures for each age group and no incidence rates are given, this reduces the usefulness of these data for the prevention.

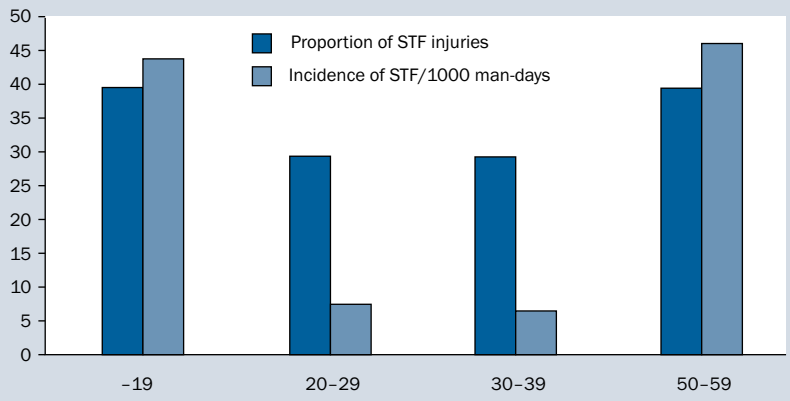
Use of only proportionate rates for age groups was lately also seen in a study among Finnish fishermen [12]. The lack of the age distribution of the population is men-



Average stay at sea	Numbers of seafarers repatriated	Numbers of seafarers	Days at risk	Risk/100000 days
150	6612	388963	58344450	11
250	840	51840	12960000	6

The numbers of repatriations depending on average stay on board and the persons at risk per 100.000 days on board are estimated based on the study data. The contrasts of the numbers of repatriated seafarers in the two groups are clearly different from the incidence rates.

Figure 5. Numbers and incidence rates of repatriations by average stay at sea [4]



Age-groups	–19	20–29	30–39	50–59
Man-days at risk	9000	30000	30000	6000
Total number of injuries	1000	800	700	700
Number of STF-related injuries	400	240	210	280
Proportion of STF-related injuries	40	30	30	40
Incidence of STF/1000 man-days	44	8	7	47
Incidence rates of all injuries/1000 man-days	111	27	23	117

The incidence rates in Figure 3 are recalculated by using an estimation of the distribution of the number of days at sea in the age groups based on an previous study [8].
The risk differences of slips, trips, and falls (STF)-related injury expressed as the incidence rates in the age groups may be even greater than the differences seen in the proportionate measures.

Figure 6. Proportions and incidence rates of slips, trips, and falls-related injuries by age groups [3]

tioned but the impact of the lack of incidence rates was not discussed.

In a review of studies of fishery health research in Latin America, only studies on occupational diseases and no studies of occupational injuries in fishing were found [13]. The lack of studies on occupational injuries in the Latin American fishery with 2–3 million, mainly small-scale fishermen is probably due the absence of register data regarding occupational diseases and injuries in the countries [14]. This situation will be improved when the infrastructure and data systems are modernised and the needed data sources to produce incidence rate studies are available.

RECOMMENDATIONS

The proportionate rates can only give some unprecise estimates of the risks. Supervisors, teachers, reviewers and editors should advice and help the students to get estimates of the populations at risk for calculation of the incidence rates where this is possible. Further, when the training in epidemiology is globally widespread and the availability of data for registers of the workforce is improved, the use of only injury proportionate rates will be reduced and incidence rates be more frequently used. Finally, the public health researchers from the universities can help and inspire governments, industries, Maritime Authorities, Insurances and Unions to compile valid and complete data for registers of the workforce as well as effective reporting systems of diseases and injuries to calculate comparable incidence rates and trends over time.

Excel file for training is available from this Dropbox: https://www.dropbox.com/s/53txv6jjn3p786i/Excel%20Table%201-3%20for%20training.xlsx?dl=0.

REFERENCES

1. Incidence (epidemiology). In: Wikipedia [Internet]. 2019 [cited 2019 Aug 2]. Available from: https://en.wikipedia.org/w/index.php?title=Incidence_(epidemiology)&oldid=886878644.
2. Dean AG, Sullivan KM, Soe MM, et al. OpenEpi: Open Source Epidemiologic Statistics for Public Health [Internet]. Open-Epi. 2013 [cited 2019 Mar 16]. Available from: https://www.openepi.com/Menu/OE_Menu.htm.
3. Jensen OC. Non-fatal occupational fall and slip injuries among commercial fishermen analyzed by use of the NOMESCO injury registration system. Am J Ind Med. 2000; 37(6): 637–644, indexed in Pubmed: 10797507.
4. Abaya A, Rivera J, Roldan S, et al. Does long-term length of stay on board affect the repatriation rates of seafarers? International Maritime Health. 2018; 69(3): 157–162, doi: 10.5603/imh.2018.0025.
5. Population Pyramids of the World from 1950 to 2100 [Internet]. PopulationPyramid.net. [cited 2019 Mar 10]. Available from: https://www.populationpyramid. net/ philippines. ; 2017.
6. Brandt LP, Kirk NU, Jensen OC, et al. Mortality among Danish merchant seamen from 1970 to 1985. Am J Ind Med. 1994; 25(6): 867–876, indexed in Pubmed: 8067363.
7. Jensen OC. Mortality in Danish fishermen. Bull Inst Marit Trop Med Gdynia. 1996; 47(1-4): 5–10, indexed in Pubmed: 9101046.
8. Jensen OC. Work related injuries in Danish fishermen. Occup Med (Lond). 1996; 46(6): 414–420, doi: 10.1093/occmed/46.6.414, indexed in Pubmed: 8987374.
9. Fullerton L, Olson L, Crandall C, et al. Occupational injury mortality in New Mexico. Ann Emerg Med. 1995; 26(4): 447–454, doi: 10.1016/s0196-0644(95)70113-3, indexed in Pubmed: 7574127.
10. Lefkowitz RY, Slade MD, Redlich CA. Risk factors for merchant seafarer repatriation due to injury or illness at sea. Int Marit Health. 2015; 66(2): 61–66, doi: 10.5603/IMH.2015.0016, indexed in Pubmed: 26119673.
11. Pecchio M, Suárez JA, Hesse S, et al. Descriptive epidemiology of snakebites in the Veraguas province of Panama, 2007-2008. Trans R Soc Trop Med Hyg. 2018; 112(10): 463–466, doi: 10.1093/trstmh/try076, indexed in Pubmed: 30165536.
12. Kaustell KO, Mattila TE, Ahvonen A, et al. Occupational injuries and diseases in fish farming in Finland 1996-2015. Int Marit Health. 2019; 70(1): 47–54, doi: 10.5603/IMH.2019.0007, indexed in Pubmed: 30931517.
13. Jensen O, Flores A, Bygvraa DA, et al. A Review of Epidemiological Studies in Latin American Fishing. J Agromedicine. 2019 [Epub ahead of print]: 1–10, doi: 10.1080/1059924X.2019.1639575, indexed in Pubmed: 31293219.
14. Barreto SM, Miranda JJ, Figueroa JP, et al. Epidemiology in Latin America and the Caribbean: current situation and challenges. Int J Epidemiol. 2012; 41(2): 557–571, doi: 10.1093/ije/dys017, indexed in Pubmed: 22407860.