Methods to assess the contribution of diseases to disability using cross-sectional studies: comparison of different versions of the attributable fraction and the attribution method

C Palazzo, Renata T. C. Yokota, J Ferguson, J Tafforeau, Jean-François Ravaud, Herman van Oyen, Wilma J. Nusselder

To cite this version:

C Palazzo, Renata T. C. Yokota, J Ferguson, J Tafforeau, Jean-François Ravaud, et al.. Methods to assess the contribution of diseases to disability using cross-sectional studies: comparison of different versions of the attributable fraction and the attribution method. International Journal of Epidemiology, Oxford University Press (OUP), 2019, 48 (2), pp.559-570. 10.1093/ije/dyy222 . hal-03480158

HAL Id: hal-03480158
https://hal-cnrs.archives-ouvertes.fr/hal-03480158

Submitted on 21 Jan 2022

International Journal of Epidemiology, 2019, 559–570
doi: 10.1093/ije/dyy222
Advance Access Publication Date: 29 October 2018
Original article

Methods

Methods to assess the contribution of diseases to disability using cross-sectional studies: comparison of different versions of the attributable fraction and the attribution method

Clémence Palazzo ,[1,2]* Renata T C Yokota,[2,3] John Ferguson,[4] Jean Tafforeau,[2] Jean-François Ravaud,[5] Herman Van Oyen[2,6] and Wilma J Nusselder[1]

[1]Department of Public Health, Erasmus MC, Rotterdam, The Netherlands, [2]Epidemiology and Public Health, Sciensano, Brussels, Belgium, [3]Department of Sociology, Interface Demography, Vrije Universiteit Brussel, Brussels, Belgium, [4]HRB Clinical Research Facility, National University of Ireland Galway, Galway, Ireland, [5]INSERM, CNRS, EHESS, Université Paris Descartes, IFRH, CERMES3, Villejuif, France and [6]Department of Public Health and Primary Care, Ghent University, Ghent, Belgium

*Corresponding author. Department of Public Health, Erasmus MC, Rotterdam, The Netherlands. E-mail: clemence.palazzo@aphp.fr

Editorial decision 10 September 2018; Accepted 4 October 2018

Abstract

Background: This study aims to illustrate the differences between approaches proposed for apportioning disability to different diseases in a multicausal situation, i.e. the unadjusted attributable fraction (AF), the adjusted AF, the average AF and the attribution method (AM). This information is useful to better interpret results obtained from cross-sectional data and help policy makers decide on public health strategies.

Methods: Data for 29 931 individuals, representative of the French household population, who participated in the 2008–09 cross-sectional Disability-Health Survey, were included. Disability was defined as any limitation reported with the Global Activity Limitation Indicator. Unadjusted AFs were calculated using Levin's formula. Adjusted AFs were estimated for each disease by calculating predicted probabilities of disability for each individual in the dataset, under the assumption that the individual is unexposed to this specific disease (logistic model). Average AFs are based on the same methodology, but have the additional advantage that the average AFs for different diseases sum to the total AF associated with eliminating all diseases. AM accounts for competing risks and partitions total disability prevalence into additive contributions of different diseases and background disability (additive model).

Results: All methods obtained similar results with respect to the estimates of the disease contribution to disability prevalences and to ranking of the diseases, except unadjusted AFs, as the method ignores multimorbidity. Confounders other than diseases, such as age and gender, should be accurately taken into account.

Conclusions: Conceptual differences, strengths and limitations of the different approaches were discussed.

Key words: Disability, contribution, attribution, diseases, attributable fraction

Key Messages

- Attributable fractions estimate the proportion of disability that theoretically can be avoided by removing a disease in the population, whereas the attribution method acknowledges competing risks and partitions total disability into additive contributions of different diseases and background.
- Because the methods account for multimorbidity, the adjusted estimates of the attributable fractions and the attribution method showed very close ranking and contributions of diseases. The Levin's unadjusted attributable fraction is hampered because it does not take into account multimorbidity.
- Age should be always considered when assessing the contributions of diseases to disability.
- Other variables that can modify the contribution of diseases to disability should be accurately taken into account with stratification, adjustment or inclusion of interactions terms in the model.
- The R packages *averisk* and *addhaz* are available to facilitate the use of the average attributable fraction and the attribution method using cross-sectional data.

Introduction

Identifying which diseases cause disability and quantifying their impact is important in promoting and monitoring population health.[1] A challenge is to measure the contribution of diseases to disability in case of multimorbidity, which occurs in 50% of the population aged 50 years and over.[2]

Two methods are used to quantify the contributions of specific diseases to disability in surveys: the attributable fraction (AF) and the attribution method (AM). The AF represents the proportion of disability in the population avoidable if a particular disease is somehow eliminated. Attributable fractions (AF) were originally unadjusted for coexisting risk factors,[3] but adjusted versions (adjusted and average AFs) were developed.[4,5] One way to obtain an adjusted version of the AF is to use a logistic model as proposed by Bruzzi *et al.*,[4] predicting the total number of disability cases that would have been observed in the dataset under the scenario that no individual had the disease of interest, but with the values of all other risk factors left unchanged. The average AF is defined as the expected proportion of disability preventable by the additional elimination of the disease of interest, when disorders are sequentially eliminated from the population in a randomly chosen order.[6–11] The average AF is valid in a multimorbidity framework.[11] Unlike the adjusted AF, the average AF has the property that the sum of the average AFs for the separate diseases under investigation is guaranteed to be less than 100%, and equal to the combined AF for all exposures (i.e. the AF for being exposed for at least one disease). The AM is analogous to the competing risks setting in the mortality analysis[12,13]; it attributes each disability case reported in a survey to a single cause, taking into account multimorbidity, and acknowledges a background disability, that is disability in individuals who do not report any disease.

The aim of this study was to illustrate the differences between approaches previously proposed for apportioning disease risk to different exposures in a multicausal situation, i.e. the unadjusted attributable fraction (unadjusted AF), adjusted AF, average AF and the attribution method (AM), by estimating the contribution of diseases to disability using the French Disability Health Survey data.

Methods

Disability-Health Survey (DHS)

The 2008–09 DHS is a national population-based representative survey [http://www.cmh.ens.fr/greco/enquetes/XML/lil-0459.xml]. We used the data of the 29 931 subjects with complete data (age range 0–106 years) living in private households (HSM). The HSM methodology has been described elsewhere.[14,15] Each respondent was assigned a weight reflecting the probability of being investigated and answering the questionnaire.

Definition of chronic diseases groups and disability

Self-reported diseases were gathered in nine groups (Supplementary Table S1, available as Supplementary data at *IJE* online). The disability indicator was the Global Activity Limitation Indicator (GALI). This is a self-reported single-item question to assess longstanding health-related activity limitations and participation restrictions[16,17]: 'For at least the past 6 months, to what extent have you been limited, because of a health problem, in activities people usually do? Severely limited, limited but not severely or not limited at all?' People were considered as disabled if they were limited or severely limited.

Unadjusted attributable fraction

The unadjusted AF was calculated using Levin's formula[3] (formula 1):

$$AF = Pe(PR-1)/(1+Pe(PR-1)) \qquad (1)$$

where Pe is the prevalence of the disease and PR the prevalence ratio, i.e. the ratio of being disabled when comparing individuals with and without the disease estimated from cross-sectional data. This attributable fraction is unadjusted for covariates or for the presence of other diseases.

Adjusted attributable fraction

The adjusted attributable fraction for each disease can be estimated by calculating predicted probabilities of disability for each individual, under the assumption that the individual is unexposed to this specific disease.[7] This requires a probabilistic model for disability adjusting for the presence or absence of a number of diseases, and possible other confounders if appropriate. The predicted probability of disease for individual i ($\hat{p}_i$) based on a logistic model is (formula 2):

$$\hat{p}_i = \frac{1}{1+\exp(-(\hat{\alpha}+\Sigma_j \hat{\beta} x_{ij}))} \qquad (2)$$

where $\hat{\alpha}_i$ represents the estimate for the intercept of the logistic regression model, x_{ij} a dummy variable (0 or 1) for the presence of disease j within person i and $\hat{\beta}$ the parameter vector for the diseases included in the model. An estimator of adjusted AF for the burden of disability due to disease j can be derived by calculating predicted probabilities from (2), setting x_{ij} to zero for all individuals, and leaving the values of the other covariates unchanged. The sum of these predicted probabilities: $E = Ew_ip_i$, with w_i the survey weight, is proportional to the expected number of disability cases that might be observed if the disease was removed from the population. The adjusted attributable fraction of the disease, AF_a, is then estimated by subtracting these expected cases from the total weighted observed cases, $O = \Sigma w_i y_i$, y_i being the indicator function for disability in person i, and dividing the result by O (formula 3):

$$\hat{A}F_\alpha = \frac{O-E}{O} \qquad (3)$$

Average attributable fraction

The combined AF for a collection of diseases can be thought of as the reduction in the burden of disability that would result from sequential elimination of all the diseases from the population, in any order. Adjusted AF assumes that the disease in question is the first of the diseases to be removed. However, in a scenario where all diseases are sequentially eliminated, the reduction in the burden of disability due to the removal of a particular disease from the population (known as the sequential attributable fraction[18]) depends on its position in the elimination order, the impact of removing the disease generally being smaller if many diseases have already been eliminated from the population. Average AF prevents this 'first removal' bias, present with the adjusted AF, by averaging all the sequential AFs for the same disease, corresponding to all possible elimination orders.[18,19]

Attribution method

The AM partitions disability prevalence into additive contributions of chronic diseases, taking into account multimorbidity and the fact that individuals can be disabled even in the absence of any disease ('background disability'). The background can represent the age effect, disability causes that were not included in the analysis, underreported and undiagnosed conditions and disability that is not associated with any condition.

In our analysis, we assumed that: (i) the distribution of disability by cause is entirely explained by the conditions that are still present at the time of the survey and the background; (ii) the cause-specific cumulative rates of disability for each disease were proportionally equal in the time preceding the survey; (iii) all persons in the same age group are exposed to the same background cumulative rate of disability; (iv) diseases and background act as independent competing causes; and (v) the start of the time at risk for disability is the same for all diseases.

The attribution method is based on the binomial additive hazard model (formula 4):

$$\begin{aligned}\hat{p} &= \exp(-\hat{\eta}_i)\\ \hat{\eta}_i &= \hat{\alpha}_a + \sum_{d=1}^{m}\hat{\beta}_j x_{ij}\end{aligned} \qquad (4)$$

where p_i is the estimated probability that individual i is disabled; η_i is the overall cumulative hazard rate of disability (linear predictor) for each individual i; α_a is the cumulative disability rate for background and is specific for each age group a; β_j are the disease-specific cumulative rates of disability (or disabling impacts) for each disease j ($1, \ldots, m$); and x_{ij} is the indicator variable for each disease j and individual i. The contribution of diseases and background to the disability prevalence can be calculated as explained by Yokota *et al.*[20]

Methods to account for confounders

The contributions of diseases to disability depend on age and gender.[21] We describe how to take them into account using the AF and the AM. The methods described below can be used for other confounders, such as educational attainment.[22]

Stratification is possible with all the methods. As the contributions of diseases to disability differ according to gender,[21] it is informative to stratify the sample by gender and to perform analysis separately for men and women. The sample may also be stratified by age categories. However, the number of categories can be too large or the sample too small to be stratified. In this case, age can be considered in different ways.[14,23–25]

By definition, adjustment does not concern the unadjusted AF. Considering the adjusted and average AFs, age and gender can be included as covariates into the logistic model with confounder-specific odds ratio (OR), but no contributions are calculated. The modelling including confounding factors allows for a more unbiased estimate of the diseases contributions to disability. The AM already takes age into account, as the background cumulative disability rate differs according to the specific age groups. This is the reason why it is not relevant to include age as a covariate in the additive model. On the other hand, it makes sense to estimate the disease-specific cumulative rates of disability by age categories too, introducing interactions between age and diseases.

Interaction terms between confounders and diseases can be included into the models. By doing so, the model parameters, namely the OR for the adjusted and average AFs and the diseases-specific cumulative hazard rates for the AM, vary by confounders categories.

Adjusted and average attributable fractions were calculated using the R package *averisk*. The function ‘BinAddHaz’ in the R package *addhaz* was used to calculate the contributions of diseases to disability with the attribution method. Analyses were performed with R version 3.3.2.

Two analyses were performed. To avoid the effect of age and ensure a better comparison between the methods themselves, the first analysis was restricted to ages 55–64 years. Next, the whole population was included. As the current versions of the R packages *addhaz* and *averisk* prevent the use of continuous variables as interaction terms, age was modelled as a categorical variable (0–34, 35–44, 45–54, 55–64, 65–74, 75+ years).

Results

Contributions of diseases among the 55–65 years age group (N = 5170 individuals)

The characteristics of the population are presented in Supplementary Table S2 (available as Supplementary data at *IJE* online). Table 1 shows the contributions of diseases to disability. The unadjusted, adjusted and average AFs and the contributions of diseases with the AM summed up to 85.6%, 64.6%, 63.6% and 75.0%, respectively. The highest contributions were observed with the AFs and the lowest with the adjusted and average AFs, which were very similar, and AM was in between. Diseases ranks were nearly similar with all the methods.

Analysis of the whole sample

Supplementary Table S3 (available as Supplementary data at *IJE* online) presents the contributions of diseases without considering age and gender; it shows significant differences beyond rank 4 between the unadjusted AF and the other methods. Table 2 presents the contributions of diseases to disability stratified by gender but without considering age. Musculoskeletal and cardiovascular conditions were the top contributors. Beyond rank 2, the diseases that contributed the most to disability differ between men and women. The sum of unadjusted AFs was higher than 100% for women and the rank position was different by methods beyond rank 3. The most important differences in the rank position were observed between the unadjusted AF and the other methods, but confidence intervals were overlapping.

Table 3 presents the contributions of diseases to disability stratified by gender and accounting for age in different ways. The contributions were higher with the AM than with the adjusted and average AFs; the rank positions were almost similar with all the methods. Adjusting for age as covariate in the logistic model reduced the contributions of diseases for the adjusted AFs and average AFs in comparison with Table 2. Looking at adjusted AFs, it reversed the rank of psychiatric and cardiovascular diseases in women, but the contributions of thesetwo2 diseases and their CIs were very close. The introduction of interactions between age categories and diseases did not change the adjusted and average AFs in comparison with Table 2. Considering

Table 1. Contributions of diseases to disability using the different versions of the attributable fraction (unadjusted AF, adjusted AF and average AF) and the attribution method for the 55–64-years age group ($n = 5170$), with no other confounder than diseases: 2008–09 Disability-Health Survey, France

	Unadjusted AF[a]		Adjusted AF[b]		Average AF[c]		Attribution method[d]	
Rank	Conditions	Contribution to disability (%)	Conditions	Contribution to disability (%)	Conditions	Contribution to disability (%)	Conditions	Contribution to disability (%)
							Background	25.0 (24.1-26.0)
1	Musculoskeletal	38.8 (34.9 to 42.7)	Musculoskeletal	33.1 (29.2 to 37.0)	Musculoskeletal	32.2 (23.4 to 41.0)	Musculoskeletal	34.8 (33.3 to 36.2)
2	Cardiovascular	11.5 (7.5 to 15.4)	Cardiovascular	8.6 (4.7 to 12.4)	Cardiovascular	8.6 (5.0 to 12.2)	Cardiovascular	10.9 (9.7 to 12.0)
3	Psychiatric	10.4 (6.8 to 13.9)	Psychiatric	5.5 (1.9 to 9.0)	Psychiatric	5.3 (1.9 to 8.7)	Psychiatric	7.5 (6.7 to 8.3)
4	Respiratory	6.5 (2.5 to 10.4)	Respiratory	4.2 (0.6 to 7.8)	Respiratory	4.1 (1.4 to 6.7)	Respiratory	5.7 (4.9 to 6.5)
5	Diabetes	5.9 (2.2 to 9.6)	Diabetes	4.1 (0.9 to 7.3)	Diabetes	3.8 (0.5 to 7.2)	Diabetes	4.4 (4.0 to 5.0)
6	Accidents	4.7 (0.5 to 8.8)	Accidents	3.5 (0.0 to 7.0)	Accidents	3.6 (0.9 to 6.3)	Accidents	4.4 (3.6 to 5.3)
7	Cancer	3.8 (0.1 to 7.5)	Cancer	3.4 (0.1 to 6.7)	Cancer	3.6 (1.5 to 5.8)	Cancer	4.2 (3.5 to 5.1)
8	Urological	2.3 (-1.3 to 6.0)	Neurological	1.5 (-1.8 to 4.8)	Neurological	1.7 (0.1 to 3.2)	Neurological	1.9 (1.4 to 2.5)
9	Neurological	1.7 (-2.6 to 5.9)	Urological	0.7 (-2.9 to 4.4)	Urological	0.7 (-1.6 to 2.9)	Urological	1.1 (0.9 to 1.4)
	Sum of diseases' contributions	85.6	Sum of diseases' contributions	64.6	Sum of diseases' contributions	63.6	Sum of diseases' contributions	75.0

[a] Unadjusted AF (attributable fraction) was calculated using Levin's formula.[3] It represents the proportion of disability in the population that might be avoided if a particular disease was somehow eliminated. It is non-adjusted for multimorbidity.

[b] Adjusted AF (adjusted attributable fraction) was estimated from a logistic model where all the diseases are covariates[4,5]; no other confounder was included in the model. An adjusted AF is estimated for each specific disease by calculating predicted probabilities of disability for each individual in the dataset, under the assumption that the individual is unexposed to this specific disease, irrespective of his real status.

[c] Average AF (average attributable fraction) is based on the same methodology as the adjusted AF, but estimates the reduction in the burden of disability that would result from sequential elimination of all the diseases from the population, in any order.[8,9]

[d] Attribution method (AM) is based on competing risks analysis. It partitions disability prevalence into additive contributions of chronic conditions and 'background', that is the fact that individuals can be disabled even in the absence of any disease. The cumulative rate of background was the same among the 55-64-year age group.

Table 2. Contributions of diseases to disability stratified by gender using the different versions of the attributable fraction (unadjusted AF, adjusted AF and average AF) and the attribution method for the whole population ($n = 29\,931$ individuals, age range 0-106 years), with no other confounder than diseases: 2008-09 Disability-Health Survey, France

	Unadjusted AF[a]		Adjusted AF[b]		Average AF[c]		Attribution method[d]	
Rank	Conditions	Contribution to disability (%)	Conditions	Contribution to disability (%)	Conditions	Contribution to disability (%)	Conditions	Contribution to disability (%)
Men								
							Background	36.8 (36.0 to 37.7)
1	Musculoskeletal	34.9 (32.5 to 37.2)	Musculoskeletal	23.9 (21.4 to 26.4)	Musculoskeletal	23.2 (18.3 to 28.2)	Musculoskeletal	22.1 (21.3 to 23.0)
2	Cardiovascular	19.2 (16.7 to 21.6)	Cardiovascular	12.5 (9.8 to 15.2)	Cardiovascular	12.9 (9.2 to 16.6)	Cardiovascular	13.6 (12.7 to 14.4)
3	Respiratory	9.1 (6.6 to 11.7)	Respiratory	5.7 (3.2 to 8.1)	Respiratory	5.5 (2.8 to 8.2)	Accidents	6.4 (5.7 to 7.2)
4	Psychiatric	8.9 (6.2 to 11.5)	Accidents	5.0 (2.6 to 7.4)	Accidents	5.1(2.8 to 7.3)	Respiratory	5.8 (5.3 to 6.3)
5	Accidents	8.0 (5.3 to 10.6)	Psychiatric	4.5 (1.9 to 7.0)	Psychiatric	4.5 (2.3 to 6.7)	Psychiatric	6.0 (5.5 to 6.7)
6	Diabetes	7.0 (5.0 to 8.9)	Diabetes	3.6 (1.4 to 5.9)	Diabetes	3.5 (1.3 to 5.6)	Diabetes	2.9 (2.6 to 3.1)
7	Urological	5.7 (3.0 to 8.3)	Neurological	2.1 (-0.2 to 4.3)	Neurological	2.3 (0.7 to 3.9)	Neurological	2.5 (2.1 to 2.9)
8	Neurological	3.5 (0.7 to 6.2)	Cancer	1.8 (-0.5 to 4.2)	Cancer	1.9 (0.1 to 3.7)	Cancer	2.1 (1.8 to 2.4)
9	Cancer	3.4 (0.5 to 6.2)	Urological	1.5 (-1.0 to 4.0)	Urological	1.6 (-0.2 to 3.3)	Urological	1.8 (1.6 to 2.0)
	Sum of diseases' contributions	99.7	Sum of diseases' contributions	60.6	Sum of diseases' contributions	60.4	Sum of diseases' contributions	63.2
Women								
							Background	**36.5 (35.7 to 37.2)**
1	Musculoskeletal	47.2 (45.2 to 49.2)	Musculoskeletal	35.4 (33.2 to 37.7)	Musculoskeletal	34.6 (29.8 to 39.5)	Musculoskeletal	29.5 (28.6 to 30.4)
2	Cardiovascular	12.8 (10.4 to 15.1)	Cardiovascular	7.1 (4.8 to 9.3)	Cardiovascular	7.5 (5.5 to 9.5)	Cardiovascular	7.9 (7.3 to 8.4)
3	Psychiatric	12.2 (10.0 to 14.3)	Psychiatric	5.8 (3.6 to 7.9)	Psychiatric	5.5 (3.1 to 7.9)	Psychiatric	7.2 (6.6 to 7.7)
4	Respiratory	7.2 (4.8 to 9.6)	Respiratory	4.4 (2.2 to 6.5)	Respiratory	4.1 (1.8 to 6.4)	Respiratory	5.0 (4.6 to 5.4)
5	Diabetes	7.0 (5.0 to 8.9)	Diabetes	3.4 (1.4 to 5.3)	Diabetes	3.4 (1.7 to 5.0)	Diabetes	3.4 (3.1 to 3.7)
6	Urological	5.2 (2.9 to 7.5)	Cancer	2.6 (0.6 to 4.6)	Cancer	2.9 (1.4 to 4.4)	Cancer	3.2 (2.7 to 3.7)
7	Accidents	4.0 (1.6 to 6.3)	Neurological	2.3 (0.3 to 4.2)	Neurological	2.6 (1.4 to 3.9)	Neurological	2.6 (2.2 to 3.0)
8	Cancer	3.5 (1.2 to 5.8)	Accidents	2.2 (0.1 to 4.3)	Accidents	2.2 (0.7 to 3.6)	Accidents	2.7 (2.3 to 3.1)
9	Neurological	3.1 (0.5 to 5.7)	Urological	1.7 (-0.4 to 3.8)	Urological	1.8 (0.3 to 3.2)	Urological	2.1 (1.9 to 2.4)
	Sum of diseases' contributions	102.2	Sum of diseases' contributions	64.9	Sum of diseases' contributions	64.6	Sum of diseases' contributions	63.5

[a] Unadjusted AF (attributable fraction) was calculated using Levin's formula.[3] It represents the proportion of disability in the population that might be avoided if a particular disease was somehow eliminated. It is non-adjusted for multimorbidity.

[b] Adjusted AF (adjusted attributable fraction) was estimated from a logistic model where all the diseases are covariates.[4,5] An adjusted AF is estimated for each specific disease by calculating predicted probabilities of disability for each individual in the dataset, under the assumption that the individual is unexposed to this specific disease, irrespective of their real status.

[c] Average AF (average attributable fraction) is based on the same methodology as the adjusted AF, but it estimated the reduction in the burden of disability that would result from sequential elimination of all the diseases from the population, in any order.[8,9]

[d] Attribution method is based on competing risks analysis. It partitions disability prevalence into additive contributions of chronic conditions and 'background' (that is the fact that individuals can be disabled even in the absence of any disease). The background cumulative rate of disability varies by 10-year age group. The contribution of background corresponds to the sum of the contributions of background of each age category.

Table 3. Contributions of diseases to disability stratified by gender using different methods accounting for age: adjusting for age categories (adjusted and average AF), estimating different background cumulative rate of disability for specific age categories (AM), and introducing interactions between age categories and diseases (adjusted and average AF, AM)

	Adjusted AF				Average AF				AM			
	Adjusting for age categories[a]		Introducing interaction terms between age categories and diseases		Adjusting for age categories[a]		Introducing interaction terms between age categories and diseases		Estimating different background cumulative-rate of disability for specific age categories[b]		Introducing interaction terms between age categories and diseases too	
Rank	Conditions	Contributions to disability (%)	Conditions	Contribution to disability (%)	Conditions	Contribution to disability (%)	Conditions	Contribution to disability (%)	Conditions	Contribution to disability (%)	Conditions	Contribution to disability (%)
Men												
									Background	**36.8 (36.0-37.7)**	**Background**	**33.1 (32.4-33.9)**
1	Musculoskeletal	20.8 (15.1 to 26.5)	Musculoskeletal	23.9 (21.2 to 26.6)	Musculoskeletal	20.6 (18.0 to 23.2)	Musculoskeletal	23.9 (18.9 to 28.9)	Musculoskeletal	22.1 (21.3 to 23.0)	Musculoskeletal	23.9 (23.0 to 24.8)
2	Cardiovascular	10.7 (7.2 to 14.1)	Cardiovascular	11.4 (8.8 to 13.9)	Cardiovascular	9.7 (7.0 to 12.4)	Cardiovascular	12.0 (8.5 to 15.6)	Cardiovascular	13.6 (12.7 to 14.4)	Cardiovascular	12.9 (12.0 to 13.7)
3	Respiratory	5.7 (3.1 to 8.4)	Respiratory	5.7 (3.2 to 8.2)	Respiratory	5.5 (2.9 to 8.2)	Respiratory	5.8 (2.9 to 8.6)	Accidents	6.4 (5.7 to 7.2)	Respiratory	7.5 (5.8 to 7.4)
4	Accidents	5.0 (2.7 to 7.4)	Accidents	5.1 (2.8 to 7.3)	Accidents	5.1 (2.4 to 7.8)	Accidents	5.1 (2.8 to 7.3)	Respiratory	5.8 (5.3 to 6.3)	Accidents	6.5 (5.8 to 7.4)
5	Psychiatric	4.8 (2.6 to 7.0)	Psychiatric	4.5 (2.5 to 6.5)	Psychiatric	4.6 (2.4 to 6.8)	Psychiatric	4.6 (2.4 to 6.9)	Psychiatric	6.0 (5.5 to 6.7)	Psychiatric	6.2 (5.6 to 6.9)
6	Diabetes	2.7 (0.0 to 5.3)	Diabetes	3.6 (1.2 to 5.9)	Diabetes	2.6 (0.4 to 4.9)	Diabetes	3.5 (1.0 to 6.0)	Diabetes	2.9 (2.6 to 3.1)	Diabetes	3.2 (2.8 to 3.5)
7	Neurological	2.2 (0.7 to 3.6)	Neurological	1.9 (-0.2 to 4.0)	Neurological	1.8 (-0.6 to 4.2)	Neurological	2.2 (0.7 to 3.8)	Neurological	2.5 (2.1 to 2.9)	Neurological	2.6 (2.2 to 3.0)
8	Cancer	1.6 (0.0 to 3.0)	Cancer	1.7 (-0.6 to 4.0)	Cancer	1.4 (-0.9 to 3.7)	Cancer	1.7 (0.3 to 3.2)	Cancer	2.1 (1.8 to 2.4)	Cancer	2.1 (1.8 to 2.5)
9	Urological	0.9 (-0.8 to 2.7)	Urological	1.3 (-1.0 to 3.5)	Urological	0.8 (-1.5 to 3.1)	Urological	1.5 (-0.3 to 3.3)	Urological	1.8 (1.6 to 2.0)	Urological	2.0 (1.7 to 2.2)
	Sum of diseases contributions	54.6	Sum of diseases contributions	59.1	Sum of diseases contributions	52.1	Sum of diseases contributions	60.4	Sum of diseases contributions	63.2	Sum of diseases contributions	66.9
Women												
									Background	**36.5 (35.7 to 37.2)**	**Background**	**31.3 (29.6 to 33.2)**
1	Musculoskeletal	28.9 (26.5 to 31.3)	Musculoskeletal	35.3 (33.3 to 37.4)	Musculoskeletal	29.3 (24.2 to 34.4)	Musculoskeletal	35.5 (30.8 to 40.2)	Musculoskeletal	29.5 (28.6 to 30.4)	Musculoskeletal	32.6 (29.6 to 35.7)
2	Psychiatric	5.6 (3.6 to 7.7)	Cardiovascular	6.1 (4.2 to 8.1)	Cardiovascular	5.8 (3.5 to 8.0)	Cardiovascular	7.3 (5.3 to 9.3)	Cardiovascular	7.9 (7.3 to 8.4)	Cardiovascular	8.2 (7.1 to 9.9)
3	Cardiovascular	5.0 (2.7 to 7.3)	Psychiatric	5.9 (3.8 to 8.1)	Psychiatric	5.7 (3.4 to 8.0)	Psychiatric	5.6 (3.4 to 7.8)	Psychiatric	7.2 (6.6 to 7.7)	Psychiatric	7.5 (5.5 to 8.9)
4	Respiratory	4.4 (2.1 to 6.7)	Respiratory	4.5 (2.3 to 6.6)	Respiratory	4.4 (2.4 to 6.5)	Respiratory	4.2 (2.0 to 6.4)	Respiratory	5.0 (4.6 to 5.4)	Respiratory	5.6 (4.6 to 6.7)
5	Diabetes	2.5 (0.5 to 4.5)	Diabetes	3.1 (0.9 to 5.2)	Diabetes	2.7 (0.9 to 4.5)	Diabetes	3.2 (1.5 to 4.8)	Diabetes	3.4 (3.1 to 3.7)	Diabetes	3.5 (1.9 to 4.4)
6	Cancer	2.3 (0.3 to 4.4)	Cancer	2.5 (0.3 to 4.7)	Cancer	2.7 (1.4 to 3.9)	Cancer	2.8 (1.8 to 3.8)	Cancer	3.2 (2.7 to 3.7)	Cancer	3.1 (2.1 to 4.3)
7	Accidents	2.1 (-0.2 to 4.3)	Neurological	2.2 (0.1 to 4.2)	Neurological	2.3 (1.2 to 3.4)	Neurological	2.6 (1.5 to 3.7)	Neurological	2.6 (2.2 to 3.0)	Neurological	2.8 (2.0 to 3.8)
8	Neurological	1.9 (-0.1 to 3.9)	Accidents	2.1 (0.3 to 3.9)	Accidents	2.2 (0.8 to 3.6)	Accidents	2.1 (0.9 to 3.3)	Accidents	2.7 (2.3 to 3.1)	Accidents	2.4 (1.9 to 3.7)
9	Urological	1.1 (-0.8 to 3.0)	Urological	1.4 (-0.6 to 3.3)	Urological	1.4 (-0.3 to 3.0)	Urological	1.5 (0.1 to 3.0)	Urological	2.1 (1.9 to 2.4)	Urological	2.3 (1.8 to 3.0)
	Sum of diseases' contributions	53.8	Sum of diseases' contributions	63.1	Sum of diseases' contributions	56.5	Sum of diseases' contributions	64.8	Sum of diseases' contributions	63.5	Sum of diseases' contributions	68.7

Age categories: 0-34, 35-44, 45-54, 55-64, 65–74, 75+ years.

[a] Adjusted and average AFs are calculated as in Table 2, but here age is also included as covariate in the logistic model.

[b] This corresponds to the contributions already shown in Table 2. In fact, estimating different background cumulative rate of disability for specific age categories is inherent in the AM concept. This is the reason why adjusting for age is not meaningful with the AM.

the AM, including interactions modified slightly the contributions of diseases and their rank position.

Discussion

This is the first study comparing different methods aiming at better understanding the impact of diseases in the disability process. Despite conceptual differences summarized in Table 4, the approaches rank the diseases fairly similarly, and the estimated contribution sizes of the diseases are very close when multimorbidity is taken into account.

The results with the unadjusted AF differ quantitatively, as the contributions were higher and could sum over 100%, and to a lesser extent qualitatively, as the ranks of diseases might be different, which can lead to difference in prioritization of resource allocations. The main reason why the unadjusted AF differs from the other approaches is because it is unadjusted for multimorbidity. If the disabled population reports more than one disease, the prevalence ratio in formula (1) summarizes the risk of being disabled when having the disease of interest itself but also the other diseases reported at the same time.[19] Even if the AF is a useful metric in epidemiology, its use may be limited to investigate the contributions of diseases to disability, as multimorbidity is crucial in the disablement process. For this purpose, it may be more appropriate to use either a form of attributable fraction that is adjusted for other diseases or the attribution method.

In the present study, the adjusted AFs were lower than the unadjusted AFs and similar to the average AFs, which might be a feature of disability data and would not be expected in general. In fact, previous studies showed that adjusted AFs can add up to more than 100%.[11,24] In those studies, adjusted and average AFs were used to measure disease burden due to risk factors like smoking or unhealthy diet that were highly correlated and might all contribute to the same disease. Here, we grouped diseases by body structure, what has two main consequences: it reduces the number of comorbidities, as suggested by the low rate of multimorbidity (12% in total population, see Supplementary Table S2, available as Supplementary data at *IJE* online); diseases may be correlated in the same group, but the group of diseases may be independent. To assess the effect of different degrees of diseases overlapping, the analyses were also performed in 10-year age groups (see Supplementary Table S5, available as Supplementary data at *IJE* online): as expected, the number of comorbidities increased with age, and so did the difference between the average AF (AAF) and the AM. This suggest that AAF and AM give more similar results when there is no overlapping of chronic diseases; if the degree of overlapping increases, AAF and AM can lead to different contributions, but the ranking of diseases was very close.

To better understand the concept of average AF, we need to come back to the sequential AF, which represents the proportion of disability prevalence that can be eliminated by removing another disease from the population, over and above that which has already been eliminated by removing the first disease(s).[18,19] This approach is interesting for estimating the potential impact of the elimination of several diseases in a given order on the disability prevalence. For instance, it could be useful for policy makers to study the best way to achieve a reduction of disability prevalence, by: preventing (i) cardiovascular, (ii) musculoskeletal; or (i) musculoskeletal, (ii) cardiovascular. For these two conditions, there are only $2! = 2$ removal orders,[18] but considering the nine groups of chronic conditions included in this study, there are as much as 362 880 removal orders and sequential AFs. The ranges of sequential AFs represented by Supplementary Figure S1 (available as Supplementary data at *IJE* online) are very small, suggesting that the elimination order does not impact on the results here. The average AF of a specific disease averages all its sequential AFs. This is interesting, as it summarizes all the information in one parameter, but is difficult to interpret.

The core of the attribution method is the additive hazard model. It has its origin in survival analysis. Analogous to the analysis of competing risks, the exponential function is applied to the cumulative hazard rates of disability to obtain the cause-specific disability probabilities, assuming independence between the causes of disability. This is an interesting approach for many purposes: its additive property facilitates the interpretation of the results; this is the only method to consider that a person can be disabled even with no disease reported or included in the model and to assess this specific background rate of disability; the AM can be used for the decomposition of differences in health expectancies by causes of disability[12]; and it is adapted for a multinomial disability outcome, i.e. allows stratifying disability by severity level.[20] It is noteworthy that concepts for AFs with survival data have been introduced by Samuelsen and Eide,[25] and that the AM approach should be distinguished from the approach of McElduff *et al.*,[26] later adopted by Llorca and Delgado-Rodríguez.[10]

It is important to note that the three statistics we compare here are estimating differing population parameters, and accordingly the associated statistics are expected to differ under large sample sizes. Informally, AM estimates the proportion of disability caused by each disease, AF estimates the proportion of disability prevalence that might be avoided if a disease was eliminated on its own, and the AAF estimates the average reduction in disability

Table 4. Definition and concept, methods to manage confounders, strengths, limitations and computation of the different versions of the attributable fraction (PAF, adjusted AF and average AF) and the attribution method (AM)

	Definition and concept	Methods to manage age, gender and other confounders	Strengths	Limitations	Computation
Unadjusted AF	Represents the proportion of disability in the population that might be avoided if a particular disease was somehow eliminated Is based on the Levin's formula that depends on the prevalence of disease and on the risk of being disabled when having the disease	Stratification	Very simple	Multimorbidity is not considered Can add up to more than 100% Lead to a different ranking of diseases	Very easy (see formula 1)
Adjusted AF	Represents the proportion of the disabled population which would be prevented if the risk of being disabled in the exposed population (that is people with the disease of interest) was changed to the risk of the unexposed population (that is people that do not have the disease of interest) Is based on a multiplicative model (logistic regression)	Stratification Adjustment, including the confounding variable(s) in the logistic model as a 'fixed' variable Introducing interaction terms between the confounding variable(s) and diseases	Multimorbidity is considered More than three confounders can be taken into account	Can overestimate the impact of diseases, as it considers that the disease under study is the first to be removed	Available upon request to the authors Free software available (R package 'AF')
Average AF	Represents the expected proportion of disability preventable by the additional elimination of the disease of interest, when many disorders (including the disease of interest) are sequentially eliminated from the population in a randomly chosen order Prevents the 'first removal bias' that can be observed with the adjusted AF	Stratification Adjustment, including the confounding variable(s) in the logistic model as a 'fixed' variable Introduction of interactions terms between the confounding variable(s) and diseases	Multimorbidity is considered Do not add up over 100% More than three confounders can be taken into account	The averaging of all of the sequential attributable fractions creates a quantity that might be difficult to interpret (and may have no meaningful interpretation) It assumes that all diseases are equally modifiable	Free software available (R package 'averisk')
AM	Fits in the framework of competing risks and partitions total disability into causes due to different diseases and background	Age should be set apart from the other confounding variables. As the background cumulative rate of disability should be estimated by age category, it is partly taken into account by concept. However, this can be interesting to estimate the disease-specific cumulative rate of disability by age categories too, that is to	Multimorbidity is considered	No more than three confounders, including age, can be taken into account	Free software available (R package 'addhaz')

(Continued)

Table 4. Continued

Definition and concept	Methods to manage age, gender and other confounders	Strengths	Limitations	Computation
	introduce interaction terms between age and diseases			
Is based on an additive model (additive hazard model) Disease-specific and background-specific contributions sum to the total disability prevalence The background is estimated and can represent the age effect, other disability causes that were not included in the analysis, under-reported and undiagnosed conditions, and the disability that is not associated with any condition; its contribution is calculated as the contributions of diseases	For the other confounders, stratification and interactions can be used If there are too many interaction terms, reduced rank regression and splines have been proposed	Do not add up over 100% Allows to distinguish the contribution of the prevalence and to the disabling impact of diseases to disability Since the contribution of diseases is additive, it can be used for the decomposition of differences in health expectancies by causes of disability Is adapted for a multinomial disability outcome		Attribution tool is also available for non-R users upon request to the authors

prevalence resulting from elimination of a particular disease in a scenario where all the diseases are eliminated in an undetermined and random order. In addition, AM and AF assume different underlying models associating disability and disease, which cannot be both simultaneously true. To some degree the correct statistic to use depends on the desired interpretation of the analyst. Nevertheless, we have demonstrated that at the very least one might expect similarity in the ranks assigned to various risk factors using the three approaches, so that risk factor ordering and prioritization will be robust over the method chosen.

This paper also describes how age, gender and other confounders should be considered in the analyses. The difference of contributions of diseases to disability between men and women confirms previous findings[21,27] and highlights the interest to stratify on gender. Age should be set apart from the other confounders, particularly using the AM where it is more an informative than a confounding variable. Age is included as the background rate of disability and should vary by 5- or 10-year age category, but this is also recommended to estimate disease-specific cumulative rates by age categories, introducing interaction terms between diseases and 10- or 15-year age categories. The contributions obtained with the AM were closer to adjusted and average AFs when doing so. A third confounder, such as educational attainment, can be taken into account with the AM, creating age categories by educational attainment and introducing interactions between those new subgroups and diseases. The inclusion of interactions in the model is limited by the sample size. One alternative to have more parsimonious models is to use reduced rank regression[23] or splines, which have been described in previous applications of the AM.[12,28,29] Considering the adjusted and the average AFs, the easiest way to take confounders into account is to include them as covariates into the logistic model. Here, age is not different from the other confounders.[6] It is possible to include more than three confounders if necessary.

The AF and the AM have limitations in common, which are related to the use of cross-sectional data. A causal relationship between disease and disability is assumed in both approaches. Although this assumption is plausible,[30] causality cannot be assessed with cross-sectional data. This implies that disability is incorrectly attributed to diseases in cases where disability onset preceded disease onset. Applying the AF based on the relative risks from prospective studies could avoid this limitation, but those relative risks are currently not available. AF is based on the idea of eliminating the disease in the population, which is not realistic. In the future, it could be more relevant to estimate the effect of more plausible interventions[31]; for instance, it may be more valuable for policy makers to know which part of disability could be reduced if the prevalence of cardiovascular diseases was 2% lower or if they were 50% less disabling. The use of self-reported diseases and disability is also a limitation, as the validity of self-reported diseases is country- and disease-specific.[32] Finally, the results of this study may be specific to the French data. Therefore we also compared the methods using the Belgian Health Interview Survey data and found similar results (Supplementary Table S4, available as Supplementary data at *IJE* online), suggesting that our findings are consistent and probably independent of the data used.

In summary, adjusted versions of the attributable fraction and attribution method yielded close conclusions, as far as the contributions of diseases to disability was concerned, and similar ranking of diseases when multimorbidity was taken into account. The results were different with Levin's AF, which does not account for multimorbidity. To some degree the approach to use depends on the desired interpretation of the analyst. AM estimates the proportion of disability caused by each disease, AF estimates the proportion of disability prevalence that might be avoided if a disease was eliminated on its own, and the AAF estimates the average reduction in disability prevalence resulting from elimination of a particular disease in a scenario where all the diseases are eliminated in an undetermined and random order.

Supplementary Data

Supplementary data are available at *IJE* online.

Funding

This work was supported by the French Society of Physical Medicine and Rehabilitation.

Acknowledgements

We thank the French National Institute of Statistics and Economic Studies (INSEE), the French Head Office of Research, Studies, Evaluation and Statistics (DREES) of the Social Affairs Ministry, the Federative Research Institute on Disability (IFRH) and all the people involved in the Disability-Health Survey project.

Conflict of interest: None declared.

References

1. World Health Organization. *International Classification of Functioning, Disability and Health*. Geneva: World Health Organization, 2001.
2. Barnett K, Mercer SW, Norbury M, Watt G, Wyke S, Guthrie B. Epidemiology of multimorbidity and implications for health care, research, and medical education: a cross-sectional study. *Lancet* 2012;**380**:37–43.
3. Levin ML. The occurrence of lung cancer in man. *Acta Unio Int Contra Cancrum* 1953;**9**:531–41.
4. Bruzzi P, Green SB, Byar DP, Brinton LA, Schairer C. Estimating the population attributable risk for multiple risk factors using case-control data. *Am J Epidemiol* 1985;**122**:904–14.
5. Greenland S, Drescher K. Maximum likelihood estimation of the attributable fraction from logistic models. *Biometrics* 1993;**49**:865–72.
6. Eide GE, Heuch I. Average attributable fractions: a coherent theory for apportioning excess risk to individual risk factors and subpopulations. *Biom J* 2006;**48**:820–37.
7. Eide GE. Attributable fractions for partitioning risk and evaluating disease prevention: a practical guide. *Clin Respir J* 2008;**2**:92–103.
8. Eide GE, Heuch I. Attributable fractions: fundamental concepts and their visualization. *Stat Methods Med Res* 2001;**10**:159–93.
9. Gefeller O, Land M, Eide GE. Averaging attributable fractions in the multifactorial situation: assumptions and interpretation. *J Clin Epidemiol* 1998;**51**:437–41.
10. Llorca J, Delgado-Rodríguez M. A new way to estimate the contribution of a risk factor in populations avoided nonadditivity. *J Clin Epidemiol* 2004;**57**:479–83.
11. Rückinger S, Kries von R, Toschke AM. An illustration of and programs estimating attributable fractions in large scale surveys considering multiple risk factors. *BMC Med Res Methodol* 2009;**9**:7. BioMed Central Ltd;
12. Nusselder WJ, Looman CW. Decomposition of differences in health expectancy by cause. *Demography* 2004;**41**:315–34.
13. Neyman J (ed). Contributions to the theory of statistics. *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability, Statistical Laboratory of the University of California, 20 June-30 July 1960* University of California Press, 1961.
14. Palazzo C, Ravaud J-F, Trinquart L, Dalichampt M, Ravaud P, Poiraudeau S. Respective contribution of chronic conditions to disability in France: results from the national Disability-Health Survey. *PLoS One* 2012;**7**:e44994.
15. Palazzo C, Ravaud J-F, Papelard A, Ravaud P, Poiraudeau S. The burden of musculoskeletal conditions. *PLoS One* 2014;**9**:e90633.
16. Berger N, van Oyen H, Cambois E *et al*. Assessing the validity of the global activity limitation indicator in fourteen European countries. *BMC Med Res Methodol* 2015;**15**:1.
17. Jagger C, Gillies C, Cambois E *et al*. The Global Activity Limitation Index measured function and disability similarly across European countries. *J Clin Epidemiol* 2010;**63**:892–99.
18. Eide GE, Gefeller O. Sequential and average attributable fractions as aids in the selection of preventive strategies. *J Clin Epidemiol* 1995;**48**:645–55.
19. Ferguson J, Alvarez-Iglesias A, Newell J, Hinde J, O'Donnell M. Estimating average attributable fractions with confidence intervals for cohort and case-control studies. *Stat Methods Med Res* 2018;**27**:1141-52.
20. Yokota RTC, van Oyen H, Looman CWN *et al*. Multinomial additive hazard model to assess the disability burden using cross-sectional data. *Biom J* 2017;**384**:45.
21. van Oyen H, Nusselder W, Jagger C, Kolip P, Cambois E, Robine J-M. Gender differences in healthy life years within the EU: an exploration of the 'health-survival' paradox. *Int J Public Health* 2013;**58**:143–55.
22. Klijs B, Nusselder WJ, Looman CW, Mackenbach JP. Educational disparities in the burden of disability: contributions of disease prevalence and disabling impact. *Am J Public Health* 2014;**104**:e141–48.

23. Yee TW, Hastie TJ. Reduced-rank vector generalized linear models. *Stat Model* 2003;**3**:15–41.
24. O'Donnell MJ, Chin SL, Rangarajan S *et al*. Global and regional effects of potentially modifiable risk factors associated with acute stroke in 32 countries (INTERSTROKE): a case-control study. *Lancet* 2016;**388**:761–75.
25. Samuelsen SO, Eide GE. Attributable fractions with survival data. *Stat Med* 2008;**27**:1447–67. Apr 30;
26. McElduff P, Attia J, Ewald B, Cockburn J, Heller R. Estimating the contribution of individual risk factors to disease in a person with more than one risk factor. *J Clin Epidemiol* 2002;**55**: 588–92. Jun;
27. Yokota RT, de C, de Moura L *et al*. Contribution of chronic conditions to gender disparities in disability in the older population in Brazil, 2013. *Int J Public Health* 2016;**61**:1003–12.
28. Yokota RTC, Berger N, Nusselder WJ *et al*. Contribution of chronic diseases to the disability burden in a population 15 years and older, Belgium, 1997–2008. *BMC Public Health* 2015;**15**: 229. BioMed Central Ltd;
29. Klijs B, Nusselder WJ, Looman CW, Mackenbach JP, Contribution of chronic disease to the burden of disability. *PLoS One* 2011;**6**:e25325.
30. Verbrugge LM, Jette AM. The disablement process. *Soc Sci Med* 1994;**38**:1–14.
31. Westreich D. From patients to policy: population intervention effects in epidemiology. *Epidemiology* 2017;**28**:525–28.
32. Baumeister H, Kriston L, Bengel J, Härter M. High agreement of self-report and physician-diagnosed somatic conditions yields limited bias in examining mental-physical comorbidity. *J Clin Epidemiol* 2010;**63**:558–65.