Epidemiology, survival, incidence and prevalence of idiopathic pulmonary fibrosis in the USA and Canada

To the Editor:

It was interesting to read the paper by RAGHU *et al.* [1] in the July issue of the *European Respiratory Journal* (*ERJ*). The article evaluated the epidemiology of idiopathic pulmonary fibrosis (IPF) in 18–64 year-olds in the USA.

The purpose of the article was to update the epidemiological data for IPF among younger US adults and to complement a recent study that reported the incidence and prevalence of IPF among older Americans in the Medicare programme [2]. This retrospective study utilised administrative patient claims data from Optum's Clinformatics Data Mart MultiPlan Database (Optum, Eden Prairie, MN, USA), related to the time period 2004–2010. The annual cumulative prevalence increased steadily in the first few years (from 13.4 cases per 100 000 person-years in 2005 to 18.2 cases in 2010 per 100 000 person-years), while the annual incidence of IPF decreased over time (from 7.9 cases per 100 000 person-years in 2005 to 5.8 cases in 2010 per 100 000 person-years). The overall incidence decrease was mainly driven by a decreasing trend in the younger patients (aged 18–44 years), while the incidence in older patients remained stable. The information provided in the article is very interesting. In our opinion it presents some peculiar and critical issues.

First, it analyses an age range in which IPF is less frequent and suggests that the decreased incidence observed in younger subjects may be due to advances in our diagnostic ability to distinguish IPF from similar disorders occurring more frequently in younger age groups. However, the main issues relate to the use of data from healthcare administrative databases that may potentially be affected by selection biases and not represent the general population. Table 1 of the article shows the distribution, by census region, of the population registered in the Optum database in 2010 [1]. Comparing this distribution with that of the actual American population aged 18–64 years in 2011 (US Census Bureau data), broad differences can be observed: enrolees in the "Pacific" and "East South Central" regions account for 5.2% and 2.7%, respectively, of the database analysed by RAGHU *at al.* [1] while the US population aged 18–64 years residing in these two regions is 16.4% and 5.9%; and enrolees in the New England region are 17.9% while the residents are 4.7%. Moreover, since IPF is an age- and sex-specific disease, it is essential that the actual distribution by age and sex in each area is adequately represented by the sample studied. This is why we believe that the estimates provided in the paper cannot be extended to the whole American population aged 18–64 years as stated in the discussion: "Our study reveals the following: 1) we updated the estimates of the epidemiology of IPF among young adults in the USA from recent national data", and that the effect of geographical regions as an independent predictor of the incidence of IPF should be investigated more accurately.

Secondly, when databases collect health data only from workers, as in this study, there is yet another problem as estimates may be biased by the healthy worker effect. Mortality and morbidity estimates for workers are lower than those of the general population because people with poor health are usually excluded from the job market.

Finally, a further issue is the choice of a 1-year washout period to identify IPF cases. The authors say that this choice affected both the prevalence trend (initial increase followed by a plateau) and the failure to identify incident cases before 2005. The authors should also consider that with a 1-year washout period the decrease in IPF incidence may not be an actual decrease since some of the incident cases identified in the first years of the study could be cases presenting before 2004.

A Canadian study published in the same issue of the *ERJ* relied on a 4-year washout period and used national administrative data to estimate IPF incidence and prevalence in Canada during 2011 [3], hence data from the entire population. For all ages, the narrow case prevalence was 20.0 per 100 000 person-years and incidence was 9.0 per 100 000, with both rates higher in men than in women. The study reported geographical differences in IPF prevalence (narrow case definition) in the population over 50 years of age, with lower rates in the western provinces of British Columbia, Alberta, Saskatchewan and Manitoba and the highest rate in Quebec. This comparison requires a standardisation of the rates if the populations have a different sex and age distribution.

Although the article analysed a national population and gave an overview of the epidemiology of IPF in recent years, changes in the incidence and prevalence of IPF could not be assessed given the short period considered.

In our study conducted using the healthcare administrative databases of Lombardy (a region with nearly 10 million inhabitants and the most populous region in Italy), we analysed the incidence and prevalence of IPF from 2000 to 2010 and estimated the prevalence and incidence rates for 2005–2010 only, thus allowing a 5-year washout period in order not to confuse incident and prevalent cases [4]. 5 years is an appropriate period of time for a disease with a median survival of 3–5 years [5]. We believe this is a strong point of our work and must be taken into account when analysing epidemiology studies. 100% of the population was analysed because regional administrative data was used and not data from subgroups that may not be representative of the general population. As in the study by RAGHU *et al.* [1], we examined the prevalence and incidence trends in 2005–2010 and similarly observed an increase in the prevalence while the incidence remained stable at all ages. Although the case identification method used in both studies is identical, a correct comparison of estimates from different populations (from different geographical areas or over different periods of time) requires specific statistical approaches. For example, a comparison of the estimates presented in the 2016 paper by RAGHU *et al.* [1] and those for the US population in 2000 [6] should rely on standardised rates, since in the latter study the population analysed was a sample of about 3 million people living in 20 US states: 28% in the South Atlantic region, 37% in the South Central region and 31% in the North Central region. Given the different geographical distribution of the two samples and the potential impact geographical areas can have on the incidence of IPF, the comparison cannot be made based on crude rates.

Epidemiological studies of IPF have also provided varying findings because data were collected from different sources: national registers, single centres or regions, and healthcare and administrative databases. All of these methods have their advantages and drawbacks [7]. The Canadian study [3], the study by RAGHU *et al.* [1] and our study [4] used administrative databases that permit analyses of large populations, but their limit is that correct diagnosis of IPF cannot be ascertained.

We believe that in epidemiology it is very important to examine the methodology used in each individual study, since this can impact incidence and prevalence data due to "technical" factors and data may be compared although they are not actually comparable. Moreover, methodological rigour can make a difference.

Sergio Harari[1], Antonella Caminati[1], Fabiana Madotto[2], Sara Conti[2] and Giancarlo Cesana[2]
[1]Unità Operativa di Pneumologia e Terapia Semi-Intensiva Respiratoria, Servizio di Fisiopatologia Respiratoria ed Emodinamica Polmonare, Ospedale San Giuseppe-MultiMedica IRCCS, Milan, Italy. [2]Research Centre on Public Health, Dept of Medicine and Surgery, University of Milano-Bicocca, Milan, Italy.

Correspondence: Sergio Harari, Unità Operativa di Pneumologia e Terapia Semi-Intensiva Respiratoria, Servizio di Fisiopatologia Respiratoria ed Emodinamica Polmonare, Ospedale San Giuseppe-MultiMedica IRCCS, Via San Vittore 12, Milan, 20123, Italy. E-mail: sharari@hotmail.it

References

1   Raghu G, Chen SY, Hou Q, *et al.* Incidence and prevalence of idiopathic pulmonary fibrosis in US adults 18–64 years old. *Eur Respir J* 2016; 48: 179–186.
2   Raghu G, Chen SY, Yeh WS, *et al.* Idiopathic pulmonary fibrosis in US Medicare beneficiaries aged 65 years and older: incidence, prevalence, and survival, 2001–11. *Lancet Respir Med* 2014; 2: 566–572.
3   Hopkins RB, Burke N, Fell C, *et al.* Epidemiology and survival of idiopathic pulmonary fibrosis from national data in Canada. *Eur Respir J* 2016; 48: 187–195.
4   Harari S, Madotto F, Caminati A, *et al.* Epidemiology of idiopathic pulmonary fibrosis in Northern Italy. *PLos One* 2016; 11: e0147072.
5   Raghu G, Collard HR, Egan JJ, *et al.* An official ATS/ERS/JRS/ALAT statement: idiopathic pulmonary fibrosis: evidence-based guidelines for diagnosis and management. *Am J Respir Crit Care Med* 2011; 183: 788–824.
6   Raghu G, Weycker D, Edelsberg J, *et al.* Incidence and prevalence of idiopathic pulmonary fibrosis. *Am J Respir Crit Care Med* 2006; 174: 810–816.

7   Caminati A, Madotto F, Cesana G, *et al.* Epidemiological studies in idiopathic pulmonary fibrosis: pitfalls in methodologies and data interpretation. *Eur Respir Rev* 2015; 24: 436–444.

From the authors:

Idiopathic pulmonary fibrosis (IPF) is a very complex disease and despite the diagnostic criteria published in evidence-based guidelines in 2011, there are significant pitfalls in making an accurate diagnosis of IPF [1]. This has surfaced as a reason for the high rate of screen failures in recent prospective clinical trials for IPF [2–5]. Since epidemiological studies are simply based on data entered in medical records and health claim databases, and are not subject to validation by review of actual raw data by experienced experts and multidisciplinary discussions among experts in the field of interstitial lung disease, the reports of the incidence and prevalence of IPF can only be estimates of diagnosis made on assumptions and, thus, must be interpreted with caution. Indeed, tracking the true occurrence of IPF is challenging [6].

The case definition used to identify patients with IPF in our study was not validated. Besides, false positives may well have been a problem and without correction for false positives, the algorithms used to estimate the incidence and prevalence of IPF may be an overestimate by as much as 50% [7, 8].

In this regard, the correspondence from S. Harari and colleagues is very much appreciated as it surfaces other concerns with epidemiological studies and methodological issues associated with estimating incidence and prevalence of disease in the USA, Canada and beyond. We agree that such methodological issues must be considered in appreciating and understanding the differences in reports from different studies. Indeed, we had acknowledged the limitations in our study as well.

@ERSpublications
We must be mindful of the caveats of the methods used to estimate the incidence and prevalence of IPF http://ow.ly/MAa0306XRLp

Cite this article as: Raghu G. Epidemiology, survival, incidence and prevalence of idiopathic pulmonary fibrosis in the USA and Canada. *Eur Respir J* 2017; 49: 1602384 [https://doi.org/10.1183/13993003.02384-2016].

Ganesh Raghu
Center for Interstitial Lung Disease, University of Washington, Seattle, WA, USA.

Correspondence: Ganesh Raghu, Center for Interstitial Lung Disease, University of Washington, 1959, NE Pacific Ave, Seattle, WA, USA. E-mail: graghu@uw.edu

Conflict of interest: None declared.

References

1   Raghu G, Collard HR, Egan JJ, *et al.* An official ATS/ERS/JRS/ALAT statement: idiopathic pulmonary fibrosis: evidence-based guidelines for diagnosis and management. *Am J Respir Crit Care Med* 2011; 183: 788–824.
2   Raghu G, Anstrom K, King TE Jr, *et al.* Idiopathic pulmonary fibrosis clinical research network. Prednisone, azathioprine, and *N*-acetylcysteine for pulmonary fibrosis. *N Engl J Med* 2012; 366: 1068–1977.
3   King T, Bradford W, Castro-Bernardini S, *et al.* A phase 3 trial of pirfenidone in patients with idiopathic pulmonary fibrosis. *N Engl J Med* 2014; 370: 2083–2092.
4   Richeldi L, Du Bois R, Raghu G, *et al.* Efficacy and safety of nintedanib in idiopathic pulmonary fibrosis. *N Engl J Med* 2014; 370: 2071–2082.
5   The Idiopathic Pulmonary Fibrosis Clinical Research Network. Randomized trial of acetylcysteine in idiopathic pulmonary fibrosis. *N Engl J Med* 2014; 370: 2093–2101.
6   Samet JM, Doultas D, Raghu G. Idiopathic pulmonary fibrosis: tracking the true occurrence is challenging. *Eur Respir J* 2015; 46: 604–606.
7   Raghu G, Chen S-Y, Hou Q, *et al.* Incidence and prevalence of idiopathic pulmonary fibrosis in US adults 18–64 years old. *Eur Respir J* 2016; 48: 179–186.
8   Esposito DB, Lanes S, Donneyong M, *et al.* Idiopathic pulmonary fibrosis in US automated claims: incidence, prevalence and algorithm validation. *Am J Respir Crit Care Med* 2015; 192: 1200–1207.