Supplementary File: Brinks et al. *Age- and sex-specific incidence of systemic lupus erythematosus: an estimate from claims data of 2.3 million persons in the German statutory health insurance 2002*

1. Estimation of the age-specific incidence from prevalence data

The estimation of the incidence from the prevalence data is based on a mathematical relation between the age-specific prevalence of SLE and the age-specific incidence and mortality rates. The methodological foundations together with comprehensive simulation studies and a practical examples are given in [1,2]. There is a video tutorial about the method [3].
The age-specific prevalence of SLE for men and women in Germany in 2002 are shown in Figure S1.
The identification of cases comprises all in- and outpatient health service providers in Germany. The claims data are a representative sample of data that cover more than 80% of the German population in 2002. Although the data are checked for quality as prescribed by law, the diagnoses of SLE could not be confirmed due to data protection. Cases of drug-induced lupus could not be excluded. Further details about the underlying study are given in [4].

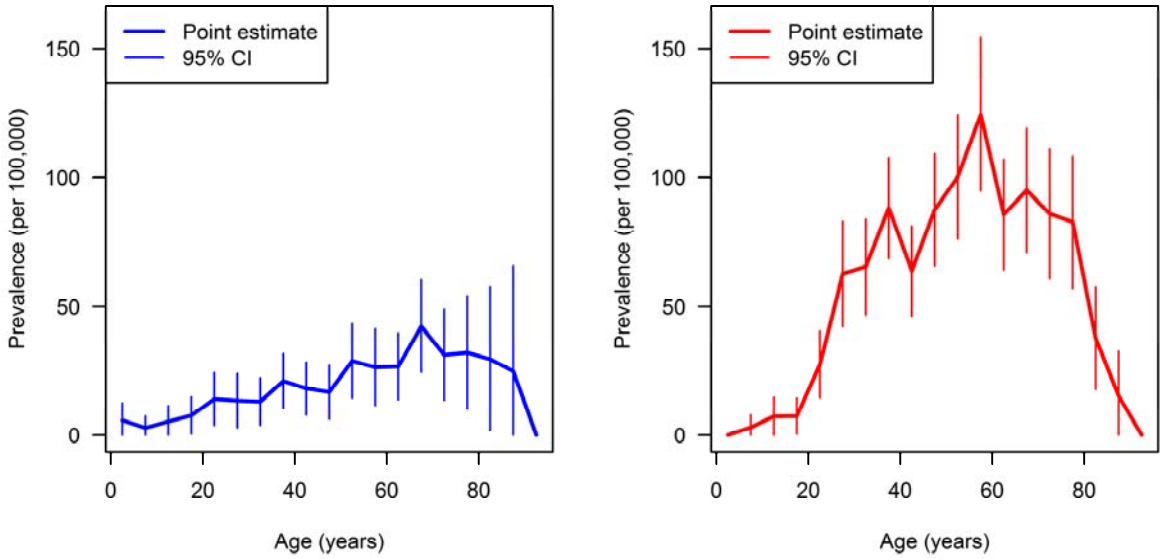

Figure S1: Age-specific prevalence of SLE in German men (left) and women (right). The vertical bars are the 95% confidence intervals (95% CI).

Apart from the prevalence data, the mortality rates of men and women from the German general population [5] and the relative mortality are required as input to the incidence estimation according to Equation (7) in [2]. Age- and sex-specific data about the relative mortality of persons with SLE compared to those without SLE stem from a meta-analysis [6]. Risk and rate ratios stable are stable measures across different populations [7, p. 59]. Confidence intervals are calculated with a bootstrap method with 5000 replicates [8].

2. Numerical values of the incidence estimates

	Men			Women		
Age group	Median	95% CI		Median	95% CI	
0 - 4	0.08	0.00	0.68	0.74	0.39	1.09
5 - 9	0.19	0.00	0.71	0.90	0.55	1.26
10 - 14	0.42	0.01	0.82	1.49	1.05	1.92
15 - 19	0.57	0.16	0.98	2.60	1.98	3.22
20 - 24	0.56	0.14	0.98	3.60	2.87	4.32
25 - 29	0.46	0.00	0.94	3.47	2.70	4.22
30 - 34	0.48	0.00	1.02	2.56	1.65	3.46
35 - 39	0.56	0.02	1.07	1.79	0.80	2.73
40 - 44	0.58	0.00	1.16	1.75	0.69	2.83
45 - 49	0.86	0.22	1.51	2.58	1.45	3.74
50 - 54	1.32	0.61	2.03	2.63	1.45	3.81
55 - 59	1.40	0.59	2.23	2.07	0.83	3.29
60 - 64	1.64	0.71	2.61	0.91	0.00	2.35
65 - 69	2.18	1.03	3.38	1.55	0.00	3.01
70 - 74	1.45	0.06	2.89	1.40	0.00	3.03
75 - 79	1.31	0.00	3.12	1.30	0.00	3.28
80 - 84	0.70	0.00	2.97	0.00	0.00	0.39

Table S1: Age-specific incidence of men and women (per 100,000 person-years) with 95% confidence intervals (95% CI). The estimates are based on 5000 bootstrap replicates (see main text).

3. Consistency of the prevalence data and the incidence estimate

To check the consistency between our incidence estimate and the underlying prevalence, we use the epidemiological rule that the prevalence equals the product of the overall incidence and the mean duration [9]. To calculate the mean duration of SLE, we subtract the mean age of onset of SLE from the mean age of death of persons with SLE [10]. We obtain the results shown in Table S2.

	Mean age of death of persons with SLE in years (A)	Mean age of onset of SLE in years (B)	Mean duration in years (C = A - B)	Estimated overall incidence per 10^5 person-years (D)	E = C × D per 10^5	Overall prevalence per 10^5 [3] (F)
Men	65.7	51.4	14.3	0.9 (0.7 to 1.1)	13 (10 to 16)	15.4 (13.1 to 17.9)
Women	68.5	39.4	29.1	1.9 (1.7 to 2.2)	55 (49 to 64)	55.4 (51.4 to 59.8)

Table S2: Comparison of the product (E) of mean duration (C) and the overall incidence (D) with the surveyed overall prevalence (F).

By comparing the last two columns (E, F) of Table S2, we see the consistency of the product of the mean duration (C) and the overall incidence (D) with the overall prevalence (F). Note that the values (D) and (F) have error bounds. Hence, the values of the last two columns E and F in Table S2 can be considered as being (statistically) equal for men and women.

References for the supplementary material

[1] Brinks R, Landwehr S, Icks A, et al. Deriving age-specific incidence from prevalence with an ordinary differential equation. *Statist Med* 2013;32(12):2070–8 doi: 10.1002/sim.5651 [published Online First: 4 October 2012].

[2] Landwehr S, Brinks, R. A comparative study of prevalence-based incidence estimation techniques with application to dementia data in Germany. *Statist Med* 2016;35(5):768–81 doi: 10.1002/sim.6736 [published Online First: 16 September 2015].

[3] Youtube-Tutorial: Swimming in the prevalence pool https://www.youtube.com/watch?v=xunIvzKRr6I (accessed July 2016)

[4] Brinks R, Fischer-Betz R, Sander O, et al. Age-specific prevalence of diagnosed systemic lupus erythematosus in Germany 2002 and projection to 2030. *Lupus* 2014;23(13):1407–11 doi: 10.1177/0961203314540352 [published Online First: 13 June 2014].

[5] Federal Statistical Office of Germany, Life-tables of Germany. URL https://www.destatis.de/DE/ZahlenFakten/GesellschaftStaat/Bevoelkerung/Sterbefaelle/Tabellen/SterbetafelFBNL.html [accessed May 2016]

[6] Bernatsky S, Boivin JF, Joseph L, et al. Mortality in Systemic Lupus Erythematosus. *Arthrit Rheumat* 2006;54(8):2550–7

[7] Breslow NE, Day NE. Statistical Methods in Cancer Research, Volume 1 – The Analysis of Case-Control Studies, Lyon, IARC Press, 1980

[8] Efron B, Tibshirani, RJ. An introduction to the bootstrap. Boca Raton, FL: CRC Press, 1994.

[9] Keiding N. Age-specific incidence and prevalence: a statistical perspective. *J Royal Statist Soc Series A* 1991; 154(3):371–412.

[10] Brinks R, Landwehr S, Waldeyer R. Age of onset in chronic diseases: new method and application to dementia in Germany. *Popul Health Metr* 2013;11:(6) doi: 10.1186/1478-7954-11-6 [published Online First: 2 May 2013].