BMJ Open

Trajectories of cigarette smoking and exposure to welding fumes and their impact on lung cancer risks: a latent class modelling approach

Benjamin Kendzia [ID],[1] Dirk Taeger [ID],[1] Hermann Pohabeln,[2] Wolfgang Ahrens,[2] Heinz-Erich Wichmann,[3] Karl-Heinz Jöckel,[4] Thomas Brüning,[1] Thomas Behrens [ID] [1]

ABSTRACT

Objectives Traditional epidemiological approaches usually assume a constant relationship between cumulative exposure and disease, which implies that exposure duration and intensity contribute equally to the studied outcome. But individuals with the same cumulative exposure but different temporal exposure patterns may show different risks. Trajectory classification is a good way to assess exposure–risk associations and leads to a better understanding of lifetime variability in exposure levels. Therefore, this study aimed to estimate lung cancer risk according to the exposure trajectory classes on welding fumes and cigarette smoking.

Design Two population-based German case–control studies.

Participants 3498 male lung cancer cases and 3539 male control subjects.

Methods Separate latent class mixed models (LCMM) were determined to identify profiles of exposure trajectories of cigarette smoking and occupational exposure to welding fumes. To investigate the risk of lung cancer by class membership, ORs with 95% CI were estimated via multiple logistic regression analyses.

Results LCMM each identified four latent classes of smoking and welding-fume exposure. Classes of smokers showed much higher risk of lung cancer compared with never smokers or subjects exposed to welding fumes. Smokers in one class characterised with the highest exposure over the past 10 years had the highest adjusted lung cancer risk (OR=39; 95% CI 29 to 53). For welding, the highest lung cancer risks were found for the class in which exposure to welding fumes in the past 10 years prior to the diagnosis of lung cancer was highest and the duration of welding was also quite high (OR=1.71; 95% CI 0.92 to 3.15).

Conclusions In summary, LCMM opens a new perspective on dose–effect relationships and could be employed to complement established epidemiological methods.

STRENGTHS AND LIMITATIONS OF THIS STUDY

⇒ Latent class mixed models were used to capture the time-varying intensities of welding fumes and cigarette smoking as well-known causes of lung cancer.

⇒ The critical time windows during which these exposures have the strongest influence on the lung cancer risk were also analysed.

⇒ This study can be limited by the fact that this modelling approach assumes that the calculated latent classes exist, which, of course, is not certain.

information on the extent of exposure at different time points. This interplay has a dynamic relationship during the life course of an exposed individual. However, traditional epidemiological approaches usually assume a constant cumulative exposure (the product of intensity and duration of exposure) over time.[1] The use of this metric implicitly assumes that the effects of exposures are additive and that the impact of an exposure on disease risk is the same regardless of exposure dose and its temporal pattern. However, individuals with the same cumulative exposure, but different temporal exposure patterns, may show different risks. In terms of prevention, it is important to know whether long-term, low-intensity exposure leads to a different risk than short-term, high-intensity exposure. The general assumption that incidence rates or risk estimates are proportional to cumulative exposure has been questioned.[2]

The association between cigarette smoking or occupational exposure to welding fumes and lung cancer has been extensively investigated in the literature.[3–6] There is still a need for research to adequately capture the time-varying intensity of exposure and to identify critical time windows during which exposure has the strongest impact on lung cancer risk.

INTRODUCTION

Exposures often span a lifetime with varying exposure intensity over time. Epidemiological studies therefore attempt to collect

For this reason, latent class mixed models (LCMM) to identify profiles of exposure histories were proposed to estimate associations with disease risk.[7–10] LCMM simplifies heterogeneous lifetime exposures into more homogeneous classes and identifies distinct subgroups of exposed individuals following a similar exposure pattern over the lifetime.[11]

The purpose of this study was to determine latent classes for smoking habits and welding-fume exposure in two case–control studies and to use these classes to estimate lung cancer risks via logistic regression. This updated analysis could provide a more complete picture of the exposure history, open a new perspective on dose–effect relationships between exposures to cigarette smoke and welding fume on lung cancer.

MATERIALS AND METHODS
Study population
Data from 3498 male lung cancer cases and 3539 male control subjects of two population-based German case–control studies were used (Humanisierung des Arbeitslebens (HdA) and Arbeit und Technik (AUT)). Because only a few women were ever exposed to welding fumes, we excluded women for this analysis. The study designs were previously described in detail.[12 13] HdA study participants were interviewed between 1988 and 1993, with a response proportion of 69% for cases and 68% for controls. Controls were randomly selected from mandatory registries and matched 1:1 to cases by sex, age and region of residence. Participants of the AUT study were interviewed with similar matching criteria between 1990 and 1996. The response was 77% among cases and 41% among controls. Institutional ethics approval was not required at the time of this study under German regulations, as the study involved only non-invasive questionnaire data.

Lifelong work and smoking histories, including start and end years of each exposure period, assuming constant exposure patterns, were collected by trained interviewers. Smoking intensity for each exposure period was calculated as the number of average cigarettes smoked per day. Detailed descriptions of the exposure assessment were previously published.[5] For each welding-exposed subject and a given job period, exposure levels for each welding activity were determined using a measurement-based job-exposure matrix (WEM) with estimates from 15 473 air measurements collected at welding workplaces in Germany.[5 14 15] Average exposure intensity levels were calculated as the product of welding-fume concentration, factor of the base material and frequency factors for welding activities, and are expressed in $\mu g/m^3$.[5] Subjects who reported that they simultaneously performed several welding processes during a given calendar year were additionally weighted with respect to welding time.

Patient and public involvement
Patients or the public were not involved in the design or conduct of this research.

Statistical analysis
All statistical analyses were conducted with R software, V.4.2.2. LCMMs are advanced statistical methods employed to identify distinct latent subgroups within a population based on longitudinal data. These models integrate latent class analysis, which classifies individuals into unobserved groups with similar developmental patterns, and mixed-effects modelling, which accounts for individual-level variability across time.[16] Thus, the purpose of this analysis was to estimate the number and size of trajectory classes and to assign latent class membership to participants based on the posterior probability of each subject's growth trajectory fitting with one of the sets of estimated classes. Classes of individual lifetime trajectories of smoking and welding-fume exposure were identified using separate analyses. The modelling of trajectories was restricted to men who had ever smoked or were exposed to welding fumes. If ever exposed subjects were not smoking cigarettes or were not exposed to welding fumes during a certain year, a zero-exposure value was coded. The years from first smoking cigarettes or first exposure to welding fumes to the year of interview (controls) or lung cancer diagnosis (cases, hereafter referred to as index dates) were used in the modelling as the metric of time (from 0 to 65 years). The LCMM function fitted smoking intensity or welding-fume exposure on a continuous scale as a function of this time metric. To consider non-normally distributed outcomes and to allow a more flexible assessment of the dose shape, a nonlinear function as a basis of a quadratic I-spline with three knots—whose number we determined—was used[17]: For smoking cigarettes, knots were placed at the quantiles (25%, 50%, 75%) 12, 24, 35 cigarettes/day, and for welding-fume exposure at 29, 93, and 588 $\mu g/m^3$, respectively. The sensitivity of the spline functions to alternative knot locations (10th, 50th and 90th percentiles and 5th, 50th and 95th percentiles) was assessed in a separate analysis. Because we observed only a marginal effect on model fit and prediction, all analyses were performed using the a priori knot locations. The LCMM function from the package *lcmm* was used to estimate the latent classes.[9 18]

To select the optimal number of latent classes, LCMM analyses with different numbers of classes (1–8) were estimated and compared in terms of fit criteria (Bayesian Information Criteria (BIC), sample-size-adjusted BIC (SABIC), Akaike Information Criteria (AIC)) with the lowest values indicating the best model.[19] In addition, diagnostic capacity criteria of the models based on entropy power and the average latent class posterior probability (ALCPP) can also help specify the number of classes, where values close to 1 are desirable.[19 20] Further, the aim was to include classes with at least 1% capture of the participants to assess the explanatory power of the class patterns. Membership of an individual to a specific class
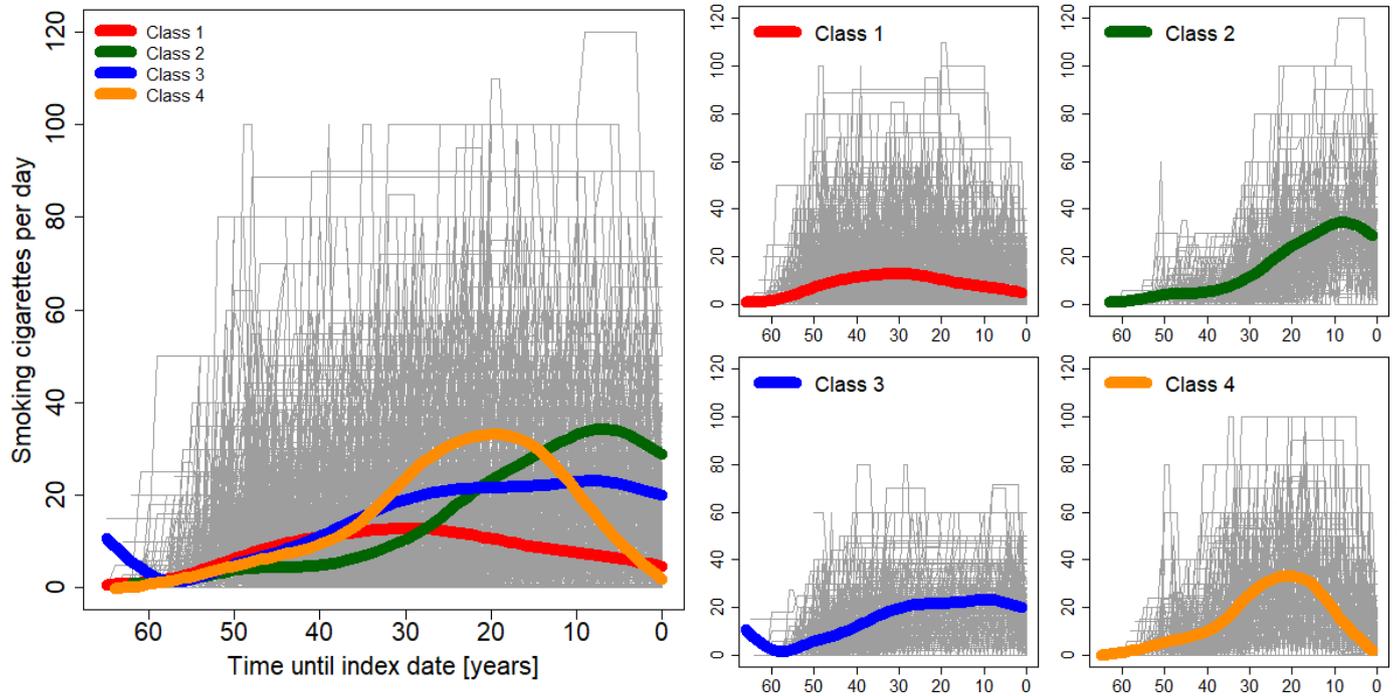
Figure 1 Spaghetti plots of smoking cigarettes per day and time until index date to illustrate the mean exposure history (grey lines). Index date means for cases the date of cancer diagnosis and for controls the date of interview. Estimates of average exposure per latent class using generalised additive models are shown with coloured lines. The right panels show the results for each class individually.

was determined by calculating the posterior probability of belonging to one class and by assigning the participant to the class with the highest probability. Also, BIC, SABIC and AIC were used to identify the best number of classes,

combined with an assessment of whether the data distinguished distinctive features in a parsimonious way.[19 20]

Spaghetti plots were performed with package *CorrMixed* to illustrate the mean exposure history as a function of
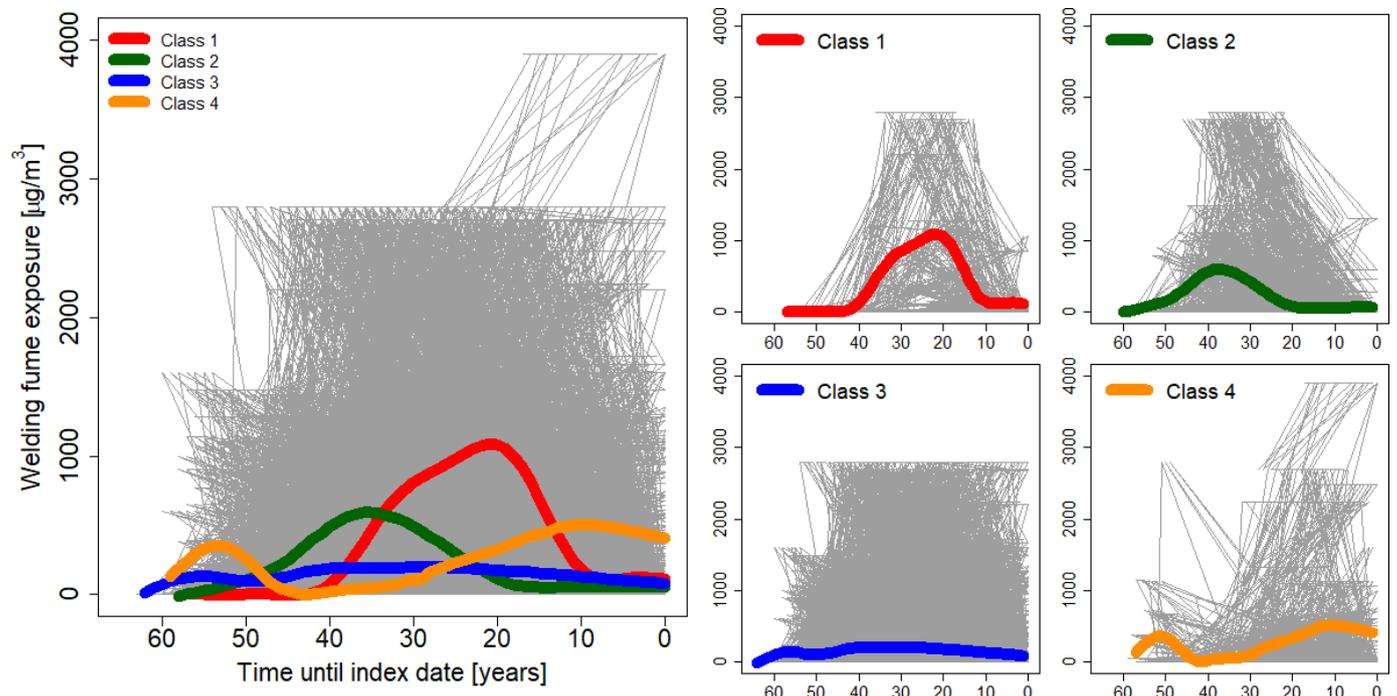

Figure 2 Spaghetti plots of welding fume exposure and time until index date to illustrate the mean exposure history (grey lines). Index date means for cases the date of cancer diagnosis and for controls the date of interview. Estimates of average exposure per latent class using generalised additive models are shown with coloured lines. The right panels show the results for each class individually.

Table 1 Characteristics and results from logistic regression analysis with risk of lung cancer of the four trajectory classes of smoking cigarettes, Germany, 1988–1996

Trajectory of smoking cigarettes		Never smokers	Class 1	Class 2	Class 3	Class 4
Cases	N (%)	67 (8)	1577 (47)	570 (17)	905 (27)	327 (10)
Controls	N (%)	786 (92)	1886 (72)	181 (7)	327 (12)	235 (9)
OR1* (95% CI)		1.00	9.37 (7.2 to 12.1)	40.40 (29.8 to 54.8)	31.85 (24.1 to 42.1)	15.88 (11.8 to 21.5)
OR2† (95% CI)		1.00	9.15 (7.1 to 11.9)	39.09 (28.8 to 53.0)	31.08 (23.5 to 41.1)	15.50 (11.5 to 21.0)
Age at index date (years)	Median (Q1–Q3)	59 (53–64)	64 (58–68)	53 (47–59)	61 (55–66)	62 (56–66)
Smoking intensity‡ (cigarettes/day)						
Lifetime	Median (Q1–Q3)	0	15 (10–15)	20 (12–35)	20 (13–23)	20 (12–40)
At index date	Median (Q1–Q3)	0	14 (8–20)	30 (20–40)	20 (15–25)	10 (4–15)
1–9 years before index date	Median (Q1–Q3)	0	18 (10–20)	30 (20–40)	20 (20–25)	20 (12–30)
10–19 years before index date	Median (Q1–Q3)	0	18 (10–20)	25 (20–40)	20 (20–25)	30 (20–40)
20–29 years before index date	Median (Q1–Q3)	0	18 (10–20)	20 (15.25)	20 (15–25)	25 (20–40)
30–39 years before index date	Median (Q1–Q3)	0	15 (10–20)	10 (5–15)	16 (10–20)	15 (10–20)
40+years before index date	Median (Q1–Q3)	0	10 (6–20)	7 (5–10)	10 (5–15)	10 (5–15)
Smoking duration§ (years)	Median (Q1–Q3)	0	29 (27–39)	33 (28–39)	42 (37–47)	34 (29–40)
Former smoking¶	N (%)	0	2373 (69)	49 (7)	102 (8)	420 (75)
Time since smoking cessation** (years)	Median (Q1–Q3)	0	19 (11–28)	3 (2–4)	3 (2–5)	9 (6–13)
Age at starting smoking (years)	Median (Q1–Q3)	0	18 (17–20)	18 (16–20)	17 (16–19)	18 (16–20)

Index date means for cases the date of cancer diagnosis and for controls the date of interview.
*OR1 is adjusted for log(age at index date) and conditional on study.
†OR2 are additionally adjusted for log(cumulative dose welding fume at the index date+1) and ever working at-risk jobs.
‡Annual intensities of smoking cigarettes.
§Duration of smoking cigarettes at the index date over all periods of smoking, excluding periods of interruptions.
¶Status at the index date.
**Calculated among former smokers only at the index date.

time (figures 1 and 2). Generalised additive models were used to estimate the average exposure per latent class, fitted by penalised iteratively reweighted least squares and smoothing selection with constrained maximum likelihood (bam function from package *mgcv*).[21]

After identifying the optimal number of latent classes to investigate the association between latent class membership and lung cancer, conditional logistic regression models were constructed to estimate ORs and 95% CIs. Subjects who had never smoked cigarettes or never worked in a welding-related occupation comprised the reference groups. All estimates were adjusted for age at index date (log transformed) and conditional on study (OR1). To control for confounding, risk estimates were additionally adjusted for several risk factors for lung cancer (OR2, see footnotes of tables 1 and 2). Ever having worked in at-risk occupations, defined by 'list A'—a list of occupations and industries associated with recognised increased risk of lung cancer risks—was used as an additional adjustment variable to account for exposure to other occupational lung carcinogens.[22 23]

The robustness of results was examined using the following sensitivity analyses:
- ► Restricting the subjects to regular welders and current smokers to limit potential uncertainties of the exposure assessment.
- ► To assess the influence of the latent classes on the risk estimates of the adjustment variables, regression models were also conducted with and without the classes.
- ► Regarding smoking, stratified analyses by different histologic types of lung cancer (squamous cell carcinoma, small cell lung carcinoma, adenocarcinoma) were also performed.
- ► Additional LCMM analyses were conducted with no consideration of smoking in the last 5 years prior to diagnosis or interview.

RESULTS

The characteristics of 3498 lung cancer cases and 3539 control subjects are shown in table 3. Lung cancer cases had a higher proportion of smokers (62% vs 26%),

Table 2 Characteristics and results from logistic regression analysis with risk of lung cancer of the four trajectory classes of exposure to welding fumes, Germany, 1988–1996

Trajectory of exposure to welding fumes		Never welding	Class 1	Class 2	Class 3	Class 4
Cases	N (%)	2698 (48)	12 (2)	37 (5)	721 (90)	30 (4)
Controls	N (%)	2894 (52)	7 (1)	38 (6)	576 (89)	24 (4)
OR1* (95% CI)		1.00	1.89 (0.74 to 4.83)	1.05 (0.67 to 1.66)	1.35 (1.19 to 1.52)	1.36 (0.79 to 2.33)
OR2† (95% CI)		1.00	1.26 (0.46 to 3.49)	1.00 (0.59 to 1.70)	1.26 (1.10 to 1.45)	1.71 (0.92 to 3.15)
Age at index date (years)	Median (Q1–Q3)	62 (56–67)	51 (45–57)	58 (53–62)	60 (53–66)	56 (48–66)
Welding exposure‡ (µg/m³)						
Lifetime	Median (Q1–Q3)	0	373 (21–1082)	56 (18–429)	93 (33–574)	59 (23–500)
At index date	Median (Q1–Q3)	0	21 (16–70)	25 (16–66)	124 (34–520)	450 (175–748)
1–9 years before index date	Median (Q1–Q3)	0	19 (16–45)	68 (27–34)	107 (34–547)	450 (121–880)
10–19 years before index date	Median (Q1–Q3)	0	374 (45–1111)	24 (16–56)	106 (35–575)	121 (24–675)
20–29 years before index date	Median (Q1–Q3)	0	976 (374–2200)	41 (18–331)	107 (35–654)	35 (17–108)
30–39 years before index date	Median (Q1–Q3)	0	900 (373–2200)	270 (72–1000)	93 (33–605)	24 (12–55)
40+years before index date	Median (Q1–Q3)	0	3 (3–9)	369 (90–975)	86 (28–440)	42 (19–125)
Welding duration§ (years)	Median (Q1–Q3)	0	28 (21–35)	31 (20–38)	12 (5–28)	30 (22–36)
Former welding¶	N (%)	0	6 (32)	50 (67)	1121 (86)	33 (61)
Time since last exposure** (years)	Median (Q1–Q3)	0	11 (5–13)	12 (6–20)	26 (11–37)	7 (3–14)
Time since first exposure (years)	Median (Q1–Q3)	0	31 (24–36)	40 (34–44)	40 (32–47)	35 (29–48)
Smoking status¶						
Never smoker	N (%)	711 (13)	1 (5)	11 (15)	123 (9)	7 (13)
Former smoker	N (%)	2386 (43)	4 (21)	25 (33)	506 (39)	23 (43)
Current smoker	N (%)	2349 (42)	14 (74)	38 (51)	640 (49)	23 (43)
Smoking other type of tobacco	N (%)	146 (3)	0	1 (1)	28 (2)	1 (2)
Smoking pack-years¶	Median (Q1–Q3)	28 (14–41)	30 (22–42)	28 (17–38)	29 (17–40)	18 (12–35)

Index date means for cases the date of cancer diagnosis and for controls the date of interview.
*OR1 is adjusted for log(age at the index date) and conditional on study.
†OR2 are additionally adjusted for log(pack-years at the index date+1) and time-since-quitting smoking cigarettes (current smokers, stopped smoking 2–7, 8–15, 16–25 or ≥26 years before the index date, smoking other type of tobacco, never smokers).
‡Annual intensities of welding-fume exposure.
§Duration of work in welding-related occupations at the index date over all periods of job phases, excluding periods of interruptions.
¶Status at the index date.
**Calculated among former workers in welding-related occupations only at the index date.

smoked more cigarettes (median: 18 vs 11 cigarettes/day) and smoked cigarettes for a longer duration compared with controls (median: 37 vs 21 years). 800 cases and 645 controls reported ever being exposed to welding fumes. The proportion of men ever exposed to welding fumes was somewhat higher among cases than controls (23% vs 18%). The median of average intensity exposure to welding fumes was higher for lung cancer cases than controls (96 vs 90 µg/m³).

For both, smoking and welding exposure, several latent class models with the number of classes ranging from 1 to 8 were considered (online supplemental table 1). The models became unstable with the eight-class models and were therefore not considered further. The four-class models for smoking and for welding fumes were identified as the best solution according to fit and diagnostic criteria. AIC, BIC and SABIC values were lowest, and the lowest value was on the off diagonal of the ALCPP, indicating good discrimination of classes. Further, classes were well differentiated with entropy indexes>0.90. Therefore, the four-class models for both smoking and welding fume were used in the following analyses in relation to the lung cancer risk.

Table 1 describes the characteristics and risk estimates of smoking behaviour considering the four classes. As expected, the association of lung cancer with classes of smoking trajectories was much stronger than for never smokers: All ORs were above 9. Adjustment for exposure to welding fumes or working in occupations with known lung cancer risk had no effect on relative risks from smoking. More than 40 years before the index date, participants in all classes had similar smoking habits

Table 3 Characteristics of lung cancer cases and controls at the index date, Germany, 1988–1996

Characteristics		Cases (n=3498)	Controls (n=3539)
Study			
HdA	N (%)	839 (24)	837 (24)
AUT	N (%)	2659 (76)	2702 (76)
Age* (years)	Median (Q1–Q3)	62 (55–67)	61 (55–67)
Smoking status*			
Never smoker	N (%)	67 (2)	786 (22)
Former smoker	N (%)	1220 (35)	1724 (49)
Current smoker	N (%)	2159 (62)	905 (26)
Smoking other type of tobacco	N (%)	52 (1)	124 (4)
Time since quitting* (years)			
<8	N (%)	444 (36)	258 (15)
8–15	N (%)	320 (26)	379 (22)
16–25	N (%)	275 (23)	532 (31)
>25	N (%)	181 (15)	555 (32)
Average smoking intensity*	Median (Q1–Q3)	18 (14–23)	11 (2–19)
Smoking duration* (years)	Median (Q1–Q3)	37 (29–44)	21 (3–34)
Smoking pack-years*	Median (Q1–Q3)	33 (22–45)	12 (1–28)
Age at starting smoking (years)	Median (Q1–Q3)	18 (16–20)	18 (17–20)
Occupational exposure to welding fume*			
Never exposed	N (%)	2698 (77)	2894 (82)
Former welding	N (%)	652 (19)	558 (16)
Current welding	N (%)	148 (4)	87 (2)
Time since last exposure* (years)	Median (Q1–Q3)	23 (10–35)	26 (11–38)
Average intensity of exposure* (µg/m^3)	Median (Q1–Q3)	96 (30–675)	90 (28–519)
Welding duration* (years)	Median (Q1–Q3)	15 (6–31)	14 (5–27)
Cumulative exposure* (µg/m^3 × years)	Median (Q1–Q3)	2732 (595–8099)	1838 (373–7425)
Age at starting welding (years)	Median (Q1–Q3)	17 (15–24)	17 (15–23)

Index date means for cases the date of cancer diagnosis and for controls the date of interview.
*Status at the index date.
AUT, Arbeit und Technik; HdA, Humanisierung des Arbeitslebens.

(figure 1). Smokers in class 2 were characterised by the highest proportion of current smokers and younger age compared with men in all other classes. Class 2 smokers showed the highest adjusted lung cancer risk compared

with never smokers (OR2=39.09; 95% CI 28.8 to 53.0). These subjects, on average, had smoked most over the past 10 years compared with the other classes (figure 1). Class 3 included a similar proportion of current smokers to class 2, showed the longest smoking duration, but a constant average intensity during the previous 30 years. Compared with never smokers, the second strongest adjusted OR (31.08; 95% CI 23.5 to 41.1) was observed. Most smokers were assigned to class 1 with the lowest average lifetime intensity and lowest duration, with a high frequency of former smokers and the highest average time since cessation. As expected, these lifetime smoking habits were associated with lower lung cancer risks compared with the other classes, although ORs were still strongly elevated (OR2=9.15; 95% CI 7.1 to 11.9). Restricting the analysis to current smokers resulted in similar risk estimates for subjects in classes 2 and 3 (online supplemental table 2). However, in the other two classes, risks were estimated as two times as high compared with the analysis presented in table 1, because these classes included many former smokers (class 1: OR2=18.9; 95% CI 14.3 to 25.0; class 4: OR2=34.4; 95% CI 21.8 to 54.6). Individual major histological subtypes were also strongly associated with smoking, with a similar pattern between classes, although the ORs for adenocarcinoma were substantially lower than for squamous cell carcinoma and small cell lung carcinoma (online supplemental tables 3-5). Lagging cigarette-smoking time by 5 years did not obviously alter the ORs: the highest risk estimates were observed for class memberships with the highest recent smoking intensity (online supplemental table 6). Table 2 shows the description and risk estimates associated with the trajectory classes of exposure to welding fumes. The classes contained roughly the same proportion of cases and controls; however, all risks of lung cancer were above 1. The classes of welding-fume exposure trajectories showed weaker associations with lung cancer than the classes of smoking trajectories. The highest smoking-adjusted lung cancer risk was found in class 4 (OR2=1.71; 95% CI 0.92 to 3.15). During the past 10 years, these welders were exposed to the highest levels of welding fumes compared with the other classes (figure 2). Class 4 subjects worked in welding-related occupations with a median of 30 years and were those with the lowest average time since last exposure and those with the lowest average smoking pack-years. Most welding-fume exposed men were assigned to class 3 with an increased lung cancer risk (OR2=1.26; 95% CI 1.10 to 1.45). compared with the other classes, they showed the shortest average duration of welding-fume exposure, with the highest proportion of formerly exposed subjects and the highest average time since last working in welding-related occupations. Their welding fume exposure was rather constant over time, averaging around 100 µg/m^3 (figure 2). Restricting the analysis to regular welders revealed higher exposure levels to welding fumes, especially for participants in class 2 (online supplemental table 7). Here, the average exposure was more than 520 µg/m^3 in each working decade,

resulting in slightly higher risk estimates (OR2=1.39; 95% CI 1.14 to 1.70).

Of note, the risk estimates of the adjustment variables were not relevantly changed when the latent classes were omitted from the model.

DISCUSSION

In this extended study, four latent classes for each, smoking cigarettes and occupational welding-fume exposure, respectively, were identified using LCMM. The selection of the final number of trajectories was based on classification accuracy values, distinctiveness and representativeness of the classes. Using these class memberships, risks of lung cancer were estimated. The four classes associated with different smoking trajectories showed a much higher risk than classes estimating trajectories of welding-fume exposure. Participants in all classes showed similar smoking habits in the 40 years in the past, but smoking patterns subsequently changed, revealing widely different lung cancer risks, with ORs ranging from 9 to 41. The class of participants characterised by the highest proportion of current smokers who, on average, had smoked the most during the past 10 years presented with the highest risk estimates for lung cancer. A similar pattern was observed for the analysis of welding-fume exposure: For welding, the highest lung cancer risks were found for class memberships with the highest exposure to welding fumes in the past 10 years and an average duration of employment of 30 years. These welders had accumulated the smallest number of smoking pack-years as compared with the other three welding classes.

Our results are consistent with previous results using traditional epidemiological approaches for the same study population or portions thereof.[5 24] Our findings highlight the strong influence of recent smoking intensity on lung cancer risk. The highest lung cancer risk was found in a class with current smokers, despite their younger age and shorter smoking duration compared with other classes. Pesch and co-workers observed a decreasing trend with time since last welding-fume exposure, ranging from OR of 1.72 for current welding to an OR of 0.87 after ≥40 years of exposure. These trends are confirmed by our data. Recent high exposure to welding fumes (median 560 μg/m³ in the past 10 years) in regular welders was associated with the highest lung cancer risk in the range of about 2, although point estimates failed to reach the level of formal statistical significance. In contrast, much higher lifetime cumulative welding-fume concentrations were found for excess risks in industrial-based cohorts.[25] As discussed in detail before,[5] different approaches to assess exposure patterns may result in different estimates that are not directly comparable. Here, we also refrained from further adjusting for socioeconomic factors, as such adjustment may result in overadjustment and consequently underestimate the true risk of smoking or welding fumes on lung cancer, particularly given the strong association between smoking and these factors. Overall, in particular, our findings regarding recent welding-fume exposure could not be estimated using a traditional analytical approach that we reported before.[5] It cannot be assumed that the classes we have identified represent the exposure classes in the general population.

A particular strength of this analysis is the use of each subject's complete and extensive smoking and welding histories. Information was available if subjects discontinued smoking habits through lifetime or varied in smoking intensity. If a participant reported having welded, a supplemental welding-specific questionnaire was applied, soliciting detailed information on the welding process, materials used and the work environment. Exposure to welding fumes was assessed on the basis of a WEM combining these detailed occupational data with a large number of personal measurements of inhalable welding fumes at various welding workplaces in Germany.[14] For each welding period, the exposure probability, frequency, intensity and type of welding process were considered for estimating the level of welding-fume exposure.[5 15]

There are several potential limitations of this study. Exposures were assessed retrospectively. Although they were reported using a standardised questionnaire which was completed face to face by trained interviewers.[12 26] With respect to smoking, it can be assumed that the reported average number of cigarettes smoked per day was approximately constant during each reported smoking period. Other limitations include that individual personal measurements of welding fumes were not available for study participants. Thus, all subjects performing the same welding tasks were assigned the same average exposure value, as documented in the WEM.

Interest in LCMM has increased in recent years and could be an important analytical approach for answering questions in occupational epidemiology. Classifying trajectories is a good way to summarise the population under study and leads to a better understanding of the life-course variability in exposure levels. Lung cancer risks may depend not only on the number of cigarette pack-years or cumulative welding-fume exposure, but also on dynamic aspects in exposure history. These dynamics include features of the exposure history itself (such as duration of cigarette smoking or time-varying intensities of exposure due to changes in the workday of a welder) and aspects of age-related susceptibility. The latter has been formalised in the concept of life-course epidemiology, according to which risk of lung cancer may vary with exposure at critical times in life. The use of cumulative exposure metrics implicitly assumes that the effects of exposures are additive and that the effect of an exposure metric on lung cancer risk is the same regardless of time of occurrence or age. Recent exposures may impose higher risks than past exposures, and histories of higher smoking intensity may involve different risks than lower ones, even for the same cumulative dose. Consequently, the indiscriminate use of the cumulative exposure metric could bias the reported associations between exposure to hazardous substances and cancer risk,[27] which may induce

a loss of information. Although it is known that the effects of intensity and duration of exposure differ, the simple cumulative exposure does not explicitly allow for modification by other time-related variables. In fact, the duration of cigarette consumption has a greater influence on lung cancer risks than the intensity of smoking.[28] The cumulative exposure metric has the advantage of simplicity and has been shown to be one of the best predictors of many dose–effect relationships.[29] Another way to assess the independent effects of duration and intensity of exposure on lung cancer risk is to examine the effect modification of the association between cumulative exposure and lung cancer by intensity.[28] Excess relative risk models, which are linear for total exposure and exponential for exposure intensity, have already been successfully applied to study the effect modification of cumulative exposure by intensity.[27 30] The simple cumulative exposure metric should not be dismissed, but LCMM may help to improve the understanding of the dose–effect relationship.

LCMM assumes that the calculated latent classes exist, which, of course, is not certain. Moreover, the modelling assigns subjects to classes based on the probability that they belong to a certain class based on the pattern of their exposure.[20] For this reason, subjects may be misclassified into individual classes and a correct class assignment is not guaranteed. Since class assignment is based on probabilities, the precise number or percentage of sample members within each class cannot be determined a priori.[31] Further, there were potential analysis problems with small sample sizes, including poor function fit indices, convergence errors and failure to detect classes with low memberships.[32] The modelling may not correctly detect small trajectories or not detect them at all, because there may be insufficient substantive information to identify these classes accurately.[33]

This study showed that it is useful to consider already known risk factors for certain diseases over a lifetime. A recent high smoking intensity among current smokers was associated with a strong impact on lung cancer risks. Regarding welding, the highest relative lung cancer risks were observed after a recent high exposure to welding fumes in regular welders. The construction of latent class trajectories, here for smoking and occupational exposure to welding fumes, opens a new perspective in analysing dose–effect relationships and could be additive to established epidemiological methods.

Author affiliations
[1]Institute of the Ruhr-Universität Bochum, Institute for Prevention and Occupational Medicine of the German Social Accident Insurance (IPA), Bochum, Germany
[2]Epidemiological Methods and Etiological Research, Leibniz-Institute for Prevention Research and Epidemiology-BIPS GmbH and Institute for Statistics, University of Bremen, Bremen, Germany
[3]Institute of Epidemiology I, Helmholtz Centre Munich, Neuherberg, Germany
[4]Institute for Medical Informatics, Biometry and Epidemiology, University Hospital Essen, Essen, Germany

ORCID iDs
Benjamin Kendzia https://orcid.org/0000-0002-8662-795X
Dirk Taeger https://orcid.org/0000-0002-5340-5584
Thomas Behrens https://orcid.org/0000-0002-4583-5234

REFERENCES

1 Kriebel D, Checkoway H, Pearce N. Exposure and dose modelling in occupational epidemiology. *Occup Environ Med* 2007;64:492–8.
2 Peto J. That the effects of smoking should be measured in pack-years: misconceptions 4. *Br J Cancer* 2012;107:406–7.
3 Kendzia B, Behrens T, Jöckel K-H, *et al*. Welding and lung cancer in a pooled analysis of case-control studies. *Am J Epidemiol* 2013;178:1513–25.
4 International Agency for Research on Cancer. Welding, molybdenum trioxide and indium tin oxide: iarc working group on the evaluation of carcinogenic risks to humans. Lyon; 2019.
5 Pesch B, Kendzia B, Pohlabeln H, *et al*. Exposure to Welding Fumes, Hexavalent Chromium, or Nickel and Risk of Lung Cancer. *Am J Epidemiol* 2019;188:1984–93.
6 International Agency for Research on Cancer. Personal habits and indoor combustion: iarc monographs on the evaluation of carcinogenic risks to humans. Lyon; 2012.

7 McCulloch CE, Lin H, Slate EH, *et al*. Discovering subpopulation structure with latent class mixed models. *Stat Med* 2002;21:417–29.

8 Henderson R, Diggle P, Dobson A. Joint modelling of longitudinal measurements and event time data. *Biostatistics* 2000;1:465–80.

9 Proust-Lima C, Séne M, Taylor JMG, *et al*. Joint latent class models for longitudinal and time-to-event data: a review. *Stat Methods Med Res* 2014;23:74–90.

10 Lévêque E, Lacourt A, Philipps V, *et al*. A new trajectory approach for investigating the association between an environmental or occupational exposure over lifetime and the risk of chronic disease: Application to smoking, asbestos, and lung cancer. *PLoS One* 2020;15:e0236736.

11 Nagin DS. Group-Based Modeling of Development. Cambridge: Harvard University Press, 2009.

12 Jöckel KH, Ahrens W, Pohlabeln H, *et al*. Lung cancer risk and welding: results from a case-control study in Germany. *Am J Ind Med* 1998;33:313–20.

13 Brüske-Hohlfeld I, Möhner M, Pohlabeln H, *et al*. Occupational lung cancer risk for men in Germany: results from a pooled case-control study. *Am J Epidemiol* 2000;151:384–95.

14 Kendzia B, Koppisch D, Van Gelder R, *et al*. Modelling of exposure to respirable and inhalable welding fumes at German workplaces. *J Occup Environ Hyg* 2019;16:400–9.

15 Kendzia B, Kaerlev L, Ahrens W, *et al*. Lifetime Exposure to Welding Fumes and Risk of Some Rare Cancers. *Am J Epidemiol* 2022;191:1753–65.

16 Meyer JD, Mutambudzi M. Construction of life-course occupational trajectories: evidence for work as a mediator of racial disparities in hypertension. *J Occup Environ Med* 2012;54:1201–7.

17 Ramsay JO. Monotone Regression Splines in Action. *Statist Sci* 1988;3.

18 Proust-Lima C, Amieva H, Jacqmin-Gadda H. Analysis of multivariate mixed longitudinal data: a flexible latent process approach. *Br J Math Stat Psychol* 2013;66:470–87.

19 Nylund KL, Asparouhov T, Muthén BO. Deciding on the Number of Classes in Latent Class Analysis and Growth Mixture Modeling: A Monte Carlo Simulation Study. *Structural Equation Modeling: A Multidisciplinary Journal* 2007;14:535–69.

20 Muthén B, Muthén LK. Integrating Person-Centered and Variable-Centered Analyses: Growth Mixture Modeling With Latent Trajectory Classes. *Alcoholism Clin & Exp Res* 2000;24:882–91.

21 Wood SN, Goude Y, Shaw S. Generalized Additive Models for Large Data Sets. *Journal of the Royal Statistical Society Series C* 2015;64:139–55.

22 Ahrens W, Merletti F. A standard tool for the analysis of occupational lung cancer in epidemiologic studies. *Int J Occup Environ Health* 1998;4:236–40.

23 Mirabelli D, Chiusolo M, Calisti R, *et al*. Database di occupazioni e attività industriali che comportano rischio di tumore del polmone. *Epidemiol Prev* 2001;25:215–21.

24 Pesch B, Kendzia B, Gustavsson P, *et al*. Cigarette smoking and lung cancer--relative risk estimates for the major histological types from a pooled analysis of case-control studies. *Int J Cancer* 2012;131:1210–9.

25 Sørensen AR, Thulstrup AM, Hansen J, *et al*. Risk of lung cancer according to mild steel and stainless steel welding. *Scand J Work Environ Health* 2007;33:379–86.

26 Brüske-Hohlfeld I, Möhner M, Ahrens W, *et al*. Lung cancer risk in male workers occupationally exposed to diesel motor emissions in Germany. *Am J Ind Med* 1999;36:405–14.

27 Vlaanderen J, Portengen L, Schüz J, *et al*. Effect modification of the association of cumulative exposure and cancer risk by intensity of exposure and time since exposure cessation: a flexible method applied to cigarette smoking and lung cancer in the SYNERGY Study. *Am J Epidemiol* 2014;179:290–8.

28 Lubin JH, Caporaso N, Wichmann HE, *et al*. Cigarette smoking and lung cancer: modeling effect modification of total exposure and intensity. *Epidemiology* 2007;18:639–48.

29 Thomas DC. Invited commentary: is it time to retire the "pack-years" variable? Maybe not! *Am J Epidemiol* 2014;179:299–302.

30 Richardson DB, Cole SR, Langholz B. Regression Models for the Effects of Exposure Rate and Cumulative Exposure. *Epidemiology (Sunnyvale)* 2012;23:892–9.

31 Weller BE, Bowen NK, Faubert SJ. Latent Class Analysis: A Guide to Best Practice. *Journal of Black Psychology* 2020;46:287–311.

32 Nylund-Gibson K, Choi AY. Ten frequently asked questions about latent class analysis. *Transl Issues Psychol Sci* 2018;4:440–61.

33 van de Schoot R, Sijbrandij M, Winter SD, *et al*. The GRoLTS-Checklist: Guidelines for Reporting on Latent Trajectory Studies. *Structural Equation Modeling: A Multidisciplinary Journal* 2017;24:451–67.