OPEN

Hospital-Based Case-Control Study of MDS Subtypes and Benzene Exposure in Shanghai

G. Bruce Copley, PhD, A. Robert Schnatter, DrPH, Thomas W. Armstrong, PhD, Richard D. Irons, PhD, Min Chen, PhD, Xiao Qin Wang, MD, PhD, and Patrick Kerzic, PhD

Objective: Due to the sparse data on benzene exposure and myelodysplastic syndrome (MDS) subtypes, we studied this relationship in patients from 29 hospitals in Shanghai, China. **Methods:** We recruited 604 cases of MDS and 1193 controls matched on age, sex, and admission date. We interviewed subjects for information on workplace and lifestyle exposures, and developed semi-quantitative exposure estimates. **Results:** Benzene exposure showed a direct exposure–response pattern with refractory cytopenia with multilineage dysplasia, a less certain association with refractory cytopenia with unilineage dysplasia, and no association with other MDS subtypes. A different pattern was observed with farm residence and smoking, which was primarily related to refractory anemias. **Conclusions:** This research demonstrates the importance of MDS subtype specification for more robust etiologic insights. Our data suggests that subtypes with non-erythroid dysplasia are associated with benzene exposure.

Learning Objectives

- Become familiar with previous evidence on the association between benzene exposure and myelodysplastic syndrome (MDS).
- Summarize the findings of the new hospital-based case-control study, including subtype-specific MDS risks.
- Identify MDS subtypes and cell types that do and do not appear to be related to benzene exposure.

B enzene exposure is a widely attributed cause of acute myeloid leukemia (AML) although its pathogenesis remains obscure. There are few formal epidemiologic studies of myelodysplastic syndrome (MDS), largely because standardized diagnostic criteria were developed relatively recently. Previously, Hayes et al[1] reported that a combined category of MDS and AML was related to benzene concentrations of approximately 10 ppm and above. Irons et al[2]

compared highly exposed cases of MDS to unexposed cases and found that a distinct pattern of bone marrow inflammation was common to benzene-exposed cases, suggesting a potential mechanistic hypothesis. A pooled analysis of three petroleum worker case-control studies that examined five lymphohematopoietic (LH) cancers reported that relatively lower concentrations of benzene, especially peak concentrations more than 3 ppm, may result in MDS.[3] A hospital-based case-control study[4] assessed several potential risk factors among MDS subtypes, finding some suggestion of a relationship with benzene. The present hospital-based case-control study uses complete study information on benzene exposure and MDS subtypes, with updated clinical information. Previous case-control studies used this database for studies on aplastic anemia,[5] AML,[6] and non-Hodgkin lymphoma (NHL).[7] Our study is motivated by the relatively large number of MDS cases diagnosed by a single laboratory using the most recent World Health Organization (WHO)[8] criteria and the presence of workplace and lifestyle exposure information. Thus, unlike most previous epidemiologic studies of MDS, this study is able to assess subtype-specific MDS risks among a large population using uniform diagnostic protocols.

METHODS

Previous publications describe this MDS case series,[2] control selection methods,[6,7] and exposure estimating methodology[9] that form the basis for the present study. The study was approved by the Combined Institutional Review Board of the University of Colorado Health Sciences Center and the internal Review Board at Fudan University in Shanghai. Briefly, there were 649 diagnosed cases of MDS among 29 participating hospitals who were more than 18 years of age at diagnosis. Two of the 649 patients did not complete either the consent form or questionnaire, resulting in a 99.7% participation rate. Participation rates were not available for controls, but overall less than 1% of all eligible study subjects (all LH cancer cases and controls combined) refused to participate. We excluded 13 cases for which there were no eligible controls and three cases missing essential data. All MDS cases were diagnosed by a single laboratory using World Health Organization[8] criteria, using more complete follow-up information than was available for the previous studies from this database. The additional follow-up yielded a decrease in the number of myelodysplastic syndrome-unspecified (MDS-U) and refractory cytopenia with unilineage dysplasia (RCUD) cases, and a slight increase in the number of refractory cytopenia with multilineage dysplasia (RCMD) cases. Detailed results of the original versus updated follow-up are in Table S1, http://links.lww.com/JOM/A328. Of the 631 remaining cases, 27 were

TABLE 1. MDS Case World Health Organization (WHO) Subtypes and Frequencies for Assessing Statistical Relations With Benzene Exposure

MDS Subtype or Category	Definition	No. of Cases (%)	No. of Controls (%)
WHO (2008) Subtypes			
RCMD	Refractory cytopenia with multilineage dysplasia	433 (71.7)	857 (71.8)
RAEB	Refractory anemia with excess blasts	103 (17.0)*	201 (16.8)
RAEB-1	Refractory anemia with excess blasts, less advanced	38 (6.3)	73 (6.1)
RAEB-2	Refractory anemia with excess blasts, more advanced	63 (10.4)	124 (10.4)
RA	Refractory anemia	39 (6.4)	77 (6.4)
RCUD	Refractory cytopenia with unilineage dysplasia (all types)	15 (2.5)	30 (2.5)
RARS	Refractory anemia with ringed sideroblasts	7 (1.2)	14 (1.2)
MDS-U	MDS-unspecified	5 (0.8)	10 (0.8)
MDS with del (5q)	MDS associated with partial chromosome 5 missing	2 (0.3)	4 (0.3)
MDS, total	All MDS cases	604 (100)	1193 (100)

MDS, myelodysplastic syndrome.
*Two RAEB cases were not classified further.

excluded because they did not meet currently accepted clinical standards for a diagnosis of MDS, leaving 604 cases for the study.

For each case, two individually-matched controls that did not have MDS, leukemia, or other blood diseases were selected from the same hospital around the same original diagnosis date. Matching factors included age (± 5 years) and sex. Only one control was located for 15 of the 604 cases, resulting in 1193 controls matched to the 604 MDS cases. Table 1 shows the distribution of cases and controls by MDS subtypes.

Cases and controls were administered questionnaires by study coordinators, usually in the hospital setting. This instrument included demographic data, work history/tasks, and information on hobbies or other activities that could result in chemical exposure, for example, farm residence, herbicide use, exposure to specified chemicals, and smoking. We stored the clinical data separately from the exposure data to maintain investigator blinding during case ascertainment and sub-typing.

We estimated benzene exposure using (a) an historic database of benzene exposures maintained by the Shanghai Institute of Public Health Supervision, (b) surrogate factory monitoring, (c) historical Chinese literature,[10–12] and (d) specific task simulations.[9] Two clerical staff independently entered the exposure data which were then reviewed and validated by an expert panel. A subject's highest estimated shift-average exposure was initially assigned to exposure groups (EGs), as follows: EG0 (non-exposed), EG1 (less than 0.3 ppm), EG2 (0.3 to less than 3 ppm), EG3 (3 to less than 30 ppm), EG4 (30+ ppm). EG3 was further divided by the original exposure assessors before statistical analysis since it covered a large range. For this current study, we either divided or aggregated the original EGs to enhance the distribution of cases and controls for assessing exposure–response in each of the MDS subgroups. For instance, for subgroups with higher numbers of cases (eg, total MDS, RCMD), we combined the most highly exposed individuals in EG3 (12 to 30 ppm) with EG4 (31+ ppm) since EG4 contained no controls. Thus, the revised EG4 category was defined as more than 12 ppm. In some sparser subgroups (eg, RAEB) we combined EG1 (less than 0.3 ppm) and EG2 (0.3 to less than 3 ppm) as well as EG3 and EG4 to facilitate a more informative analysis based on the available data.

Statistical Analysis

We conducted the statistical analysis in two phases (see Supplemental Information, http://links.lww.com/JOM/A328). First, we conducted descriptive analyses using chi-squared tests and bivariable analyses that identified variables and potential confounders—those with P values of 0.1 or less—that were

included in the second phase multivariable conditional logistic regression analyses which examined the independent effect of benzene exposure on MDS. We forced benzene exposure parameters into the multivariable models regardless of their P value. Odds ratios (ORs) with 95% confidence intervals (95% CI) and P values were computed for each retained variable and benzene EG. The matching variables, age, and sex, were considered strata. A test for linear trend parameterized the benzene EGs as an ordinal variable.

RESULTS

Demographic and selected individual characteristics of the study subjects are shown in Table 2. Cases were significantly more likely than controls to have low educational attainment and to have worked in agriculture, transportation, and marginally more likely to have worked at some point in rubber or plastic manufacturing where benzene is more prevalent. Cases were also marginally more likely to have lower body mass index (BMI), but were less likely than controls to have worked in other manufacturing or production jobs.

All (Aggregated) MDS Subtypes

The unadjusted and the multivariable adjusted model generated a similar pattern of increasing ORs through EG3, and a slight decrease from EG3 to EG4. (See Fig. 1 and Table 3). For both models, the highest two (EG3 and EG4) groups were statistically significant, as was the test for linear trend, $P < 0.01$. Farm residence was the only covariate that materially impacted the effect estimates, as its inclusion in the multivariable adjusted model resulted in lower EG3 and EG4 ORs by 11% and 8%, respectively. Effect modification by farm residence was ruled out by heterogeneity testing. Having observed only 11 cases and three controls in EG3, an alternative multivariable analysis combining EG3 and EG4 was conducted. Again using EG0 as the reference category, the ORs and 95% CI were as follows: EG1 OR = 1.68 (95% CI: 0.94, 3.01); EG2 OR = 1.92 (95% CI: 0.88, 4.20); EG3&4 OR = 3.88 (95% CI: 2.04, 7.38).

RCMD

RCMD is the most common type of MDS in China,[11] and involves cytopenia and dysplasia of multiple cell lineages. It comprises 71% of all MDS cases in this series. The ORs for EG2 to EG4 were each statistically significant in both crude and adjusted models, with adjusted ORs of 3.19 (95% CI: 1.25, 8.12), 5.38 (95% CI: 1.36, 21.3), and 3.75 (95% CI: 1.53, 9.21), respectively (see Fig. 1 and Table 4). The linear trend was significant ($P < 0.01$). The confounding by farm residence observed in the all-MDS analysis persisted in

TABLE 2. Comparison of Case and Control Characteristics

Variable	Measure	Cases (%)	Controls (%)	*P*
Sex	Male	319 (52.8)	631 (52.9)	0.98
Age (yrs)	Mean (std. dev.)	53.4 (17.4)	53.3 (17.0)	0.90
Body mass index (BMI)	Mean (std. dev.)	22.7 (3.2)	23.0 (4.0)	0.06
Marital status	Married	525 (86.9)	1014 (85.0)	0.27
	Divorced	6 (0.1)	11 (0.9)	0.88
	Widowed	49 (8.1)	107 (9.0)	0.54
	Never married	49 (8.1)	122 (10.2)	0.15
Educational attainment	None to primary	192 (31.9)	258 (21.7)	0.00
	Middle to high	314 (52.2)	645 (54.2)	0.38
	University to post-graduate	96 (16.0)	286 (24.0)	0.01
Smoking	Number (%) ever smoked	191 (31.6)	375 (31.4)	0.94
Occupational work history (Note: an individual can have multiple work histories)	Agriculture	160 (26.5)	239 (20.0)	0.00
	Textiles	61 (9.6)	109 (9.1)	0.75
	Education	57 (9.4)	149 (12.5)	0.08
	Machinery manufacturing	56 (9.3)	127 (10.6)	0.19
	Electrical equipment manufacturing	51 (8.4)	96 (8.0)	0.77
	Transportation	43 (7.1)	58 (4.9)	0.05
	Government	33 (5.5)	71 (6.0)	0.67
	Metal production	30 (5.0)	45 (3.8)	0.23
	Rubber/plastic manufacturing	27 (4.5)	33 (2.8)	0.06
	Non-metal mineral production	23 (3.8)	34 (2.8)	0.28
	Sewing industry	23 (3.8)	47 (3.9)	0.89
	Stationary Arts	22 (3.6)	33 (2.8)	0.31
	Other manufacturing or production	213 (35.2)	550 (46.1)	0.02
	Other non-manufacturing or production	259 (42.9)	564 (47.3)	0.08

the RCMD analysis to an even larger extent, resulting in a 15% to 16% lowering of ORs in EG3 and EG4. An alternative multivariable analysis combining EG3 and EG4 generated the following results: EG1 OR = 1.72 (95% CI: 0.84, 3.50); EG2 OR = 3.15 (95% CI: 1.24, 7.99); EG3&4 OR = 4.20 (95% CI: 1.98, 8.90).

RCMD was also a driver for elevated benzene-related risks observed for all multilineage subtypes combined (Table S4, http://links.lww.com/JOM/A328) which were 92% RCMD. Cases in the EG3 and EG4 exposure groups were over-represented by rubber/plastic manufacturing (6 cases/1 control) and agriculture

FIGURE 1. Multivariable-adjusted[a] odds ratios and 95% confidence intervals for benzene exposure group statistical associations with (A) all MDS cases, (B) RCMD, (C) RAEB, and (D) total RCUD. [a]Odds ratios adjusted for all other covariates which met model inclusion, that is, those with *P* values of less than or equal to 0.1 in bivariable analyses. MDS, myelodysplastic syndrome; RAEB, refractory anemia with excess blasts; RCUD, refractory cytopenia with unilineage dysplasia.

Benzene Exposures (ppm)

Endpoints	Case/Control No.	Ptrend	OR(95% CI)
all-MDS			
EG0: Unexposed	531/1126		1.00 (reference)
EG1: >0-<0.3 ppm	22/37		1.68 (0.93 - 3.01)
EG2: 0.3-<3ppm	14/16		1.94 (0.88 - 4.26)
EG3: 3 -<12ppm	11/3		5.35 (1.44 - 19.9)
EG4: 12ppm-	26/11		3.47 (1.65 - 7.28)
A		<0.001	
RCMD			
EG0: Unexposed	380/809		1.00 (reference)
EG1: >0-<0.3 ppm	14/27		1.71 (0.84 - 3.49)
EG2: 0.3-<3ppm	12/10		3.19 (1.25 - 8.12)
EG3: 3 -<12ppm	9/3		5.38 (1.36 - 21.3)
EG4: 12ppm-	18/8		3.75 (1.53 - 9.21)
B		<0.001	
RAEB			
EG0: Unexposed	93/185		1.00 (reference)
EG1: >0-<0.3 ppm	7/8		1.73 (0.55 - 5.41)
EG2: 0.3-<3ppm	1/5		0.31 (0.03 - 3.36)
EG3-4: 3ppm-	2/3		0.81 (0.11 - 5.78)
C		0.789	
Total RCUD			
EG0: Unexposed	45/104		1.00 (reference)
EG1: >0-<0.3 ppm	1/2		1.34 (0.09 - 20.1)
EG2-4: 0.3ppm-	8/1		15.7 (1.40 - 178.)
D		0.047	

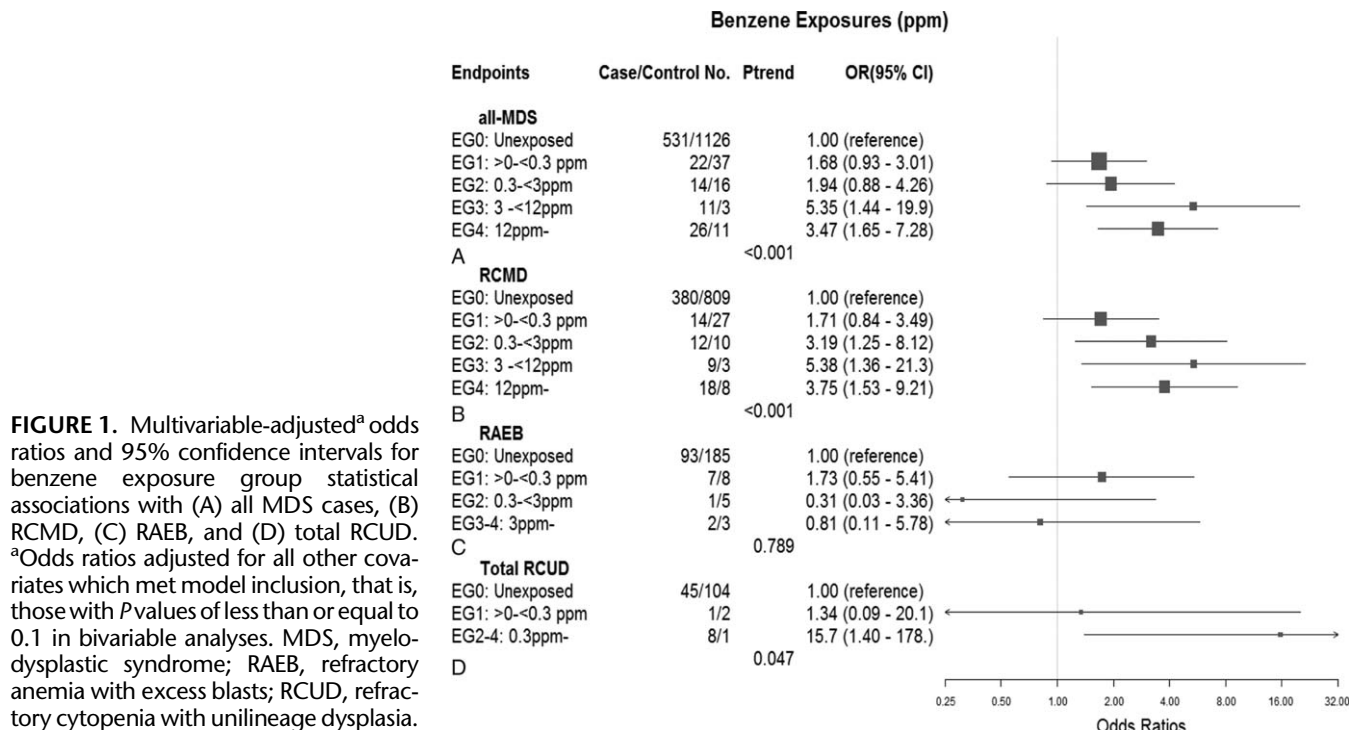Odds Ratios: 0.25   0.50   1.00   2.00   4.00   8.00   16.00   32.00

TABLE 3. Results of Bivariable and Multivariable Analyses for all MDS, With Odds Ratio (OR) and 95% Confidence Intervals (CI)

Exposures/Variables	No. Cases	No. Controls	Odds Ratio (95% CI)	
			Unadjusted	Adjusted
Benzene average exposure, ppm:				
EG0: Unexposed	531	1126	[reference]	[reference]
EG1: <0.3	22	37	1.23 (0.71, 2.11)	1.68 (0.93, 3.00)
EG2: 0.3–2.9	14	16	1.99 (0.96, 4.13)	1.94 (0.88, 4.26)
EG3: 3.0–11.9	11	3	7.81 (2.17, 28.1)	5.35 (1.44, 19.9)
EG4: 12.0–	26	11	4.81 (2.38, 9.74)	3.47 (1.65, 7.28)
BMI (continuous)*	298	606	0.97 (0.95, 1.00)	0.97 (0.95–1.00)
				$P_{trend} < 0.01$
Education:				
College/post-grad	96	286	[Reference]	[Reference]
Middle/high school	314	645	1.51 (1.14, 2.00)	1.48 (1.09, 2.00)
None or primary	192	258	2.70 (1.94, 3.77)	2.21 (1.52, 3.21)
				$P_{trend} < 0.01$
Farm residence	290	405	1.98 (1.59, 2.47)	1.66 (1.30, 2.11)
Diesel fuel	7	23	0.56 (0.23, 1.33)	0.33 (0.12, 0.95)
Metal machining	25	76	0.61 (0.38, 0.98)	0.59 (0.35, 0.97)
Anti-TB medication	29	44	1.32 (0.82, 2.12)	1.61 (0.95, 2.71)
Diabetes	39	142	0.49 (0.33, 0.72)	0.48 (0.32, 0.73)
Welding	11	34	0.64 (0.32, 1.27)	0.53 (0.25, 1.13)

BMI, body mass index.

*Number of cases and controls represent number of patients at or above the BMI median value (22.49), presented for informational purposes only; BMI was parameterized as a continuous variable in the analysis as indicated.

(7 cases/5 controls) occupations. Of the 53 RCMD cases exposed to benzene, 26 (49%) were last exposed more than 10 years before diagnosis, while only 30% of other MDS diagnoses were exposed more than 10 years prediagnosis.

RAEB and RAEB Stages

RAEB represents the most advanced stage of MDS, and poses a relatively higher risk of advancing to acute myeloid leukemia (AML). Benzene exposure by EG showed no clear relationship with RAEB, although the precision of the effect estimates was less than that for RCMD (see Fig. 1 and Table 5). An ever/never benzene parameterization resulted in an essentially null finding (OR = 1.10; 95% CI

0.46, 2.65). There were too few cases to analyze RAEB I and RAEB II separately. A history of ever smoking was observed to be a statistically significant and independent risk factor (OR = 2.96; 95% CI: 1.37, 6.39) for RAEB. Other adjustment covariates were: exposure to fertilizer, BMI, and diabetes; the latter was negatively associated with RAEB (OR = 0.28; 95% CI 0.10, 0.77).

RCUD

RCUD is a combined category comprised of refractory anemia (RA, 39 cases) and refractory cytopenia with unilineage dysplasia (RCUD) other than RA (15 cases). For the 54 combined cases, nine were exposed to benzene (see Fig. 1 and Table S5,

TABLE 4. Results of Bivariable and Multivariable Analyses for RCMD, with Odds Ratios (OR) and 95% Confidence Intervals (CI)

Exposures/Variables	No. Cases	No. Controls	Odds Ratio (95% CI)	
			Unadjusted	Adjusted
Benzene average exposure, ppm:				
EG0: unexposed	380	809	[reference]	[reference]
EG1: <0.3	14	27	1.12 (0.58, 2.15)	1.71 (0.84, 3.49)
EG2: 0.3–2.9	12	10	2.85 (1.20, 6.74)	3.19 (1.25, 8.12)
EG3: 3.0–11.9	9	3	6.79 (1.81, 25.5)	5.38 (1.36, 21.3)
EG4: 12.0–	18	8	4.65 (2.10, 10.7)	3.75 (1.53, 9.21)
				$P_{trend} < 0.01$
Education:				
College/post-grad	130	169	[reference]	[reference]
Middle/high school	231	477	1.48 (1.07, 2.01)	1.52 (1.06, 2.17)
None or primary	71	208	2.76 (1.86, 4.11)	2.44 (1.54, 3.86)
				$P_{trend} < 0.01$
Farm residence	211	291	1.97 (1.53, 2.54)	1.53 (1.15, 2.04)
Herbicides	23	15	3.34 (1.68, 6.62)	1.99 (0.94, 4.23)
Metals	15	51	0.56 (0.31, 1.02)	0.50 (0.27, 0.94)
Anti-TB medication	20	26	1.53 (0.85, 2.76)	2.07 (1.08, 3.98)
Diabetes	29	96	0.54 (0.35, 0.85)	0.56 (0.35, 0.90)
Ever smoke	131	279	0.83 (0.59, 1.16)	0.69 (0.48, 0.99)

RCMD, refractory cytopenia with multilineage dysplasia.

TABLE 5. Results of Bivariable and Multivariable Analyses for RAEB, with Odds Ratios (OR) and 95% Confidence Intervals (CI)

Exposures/Variables	No. Cases	No. Controls	Odds Ratio (95% CI)	
			Crude	Adjusted
Benzene average exposure, ppm:				
EG0: unexposed	93	185	[reference]	[reference]
EG1: <0.3	7	8	1.62 (0.56, 4.70)	1.73 (0.55, 5.41)
EG2: 0.3–2.9	1	5	0.42 (0.05, 3.58)	0.31 (0.03, 3.36)
EG3/4: 3.0+	2	3	1.33 (0.22, 7.98)	0.81 (0.11, 5.78)
				$P_{trend} = 0.95$
Ever smoked	45	68	2.01 (1.03, 3.93)	2.96 (1.37, 6.39)
Fertilizer exposure	15	19	1.84 (0.81, 4.18)	2.63 (1.02, 6.75)
Metal machining	7	20	0.57 (0.21, 1.55)	0.36 (0.12, 1.09)
Diabetes	5	29	0.31 (0.12, 0.82)	0.28 (0.10, 0.77)
BMI (continuous)*	43	110	0.92 (0.86, 0.99)	0.90 (0.83, 0.98)

BMI, body mass index; RCMD, refractory cytopenia with multilineage dysplasia.

*BMI was parameterized as a continuous variable in the analysis; number of cases and controls represent number of patients at or above the BMI median value (22.49), presented for informational purposes only.

http://links.lww.com/JOM/A328) although seven of these nine were from the smaller subgroup—that is, RCUD other than RA. Furthermore, seven of the nine exposed cases were exposed to more than 3 ppm (EG3 and EG4). Since there were no matched controls exposed above 3 ppm, risks for benzene exposure are not estimable (eg, infinite) for total RCUD via logistic regression. Approximate ORs can be calculated based on Gart and Zweifel[13]; for exposures above 3 ppm, RA shows an OR of 9.9 (95% CI: 0.46 to 212), while RCUD other than RA shows an OR of 39.4 (95% CI: 1.97 to 787). Thus, the statistical association between benzene and all types of RCUD was driven by a relatively high proportion (7/15, or 46%) of non-RA cases of RCUD exposed to benzene. However, these findings are imprecise based on the wide confidence intervals.

DISCUSSION

This hospital-based case-control study of benzene exposure and major subtypes of MDS indicate a consistent relationship between benzene exposure and aggregate MDS which is more pronounced for RCMD. Notably, there was no suggestive relationship with benzene exposure for refractory anemias, a grouping which includes RAEB, the subtype with the highest propensity to progress to AML. Despite the relatively low number of RCUD cases and the accompanying statistical imprecision, evidence for benzene effects above 3 ppm was also suggestive for the non-RA fraction. Some of the rarest MDS subtypes (eg, RARS, MDS with 5q-, MDS-U) precluded informative statistical analyses.

One common feature of RCMD and non-RA RCUD is the presence of non-erythroid dysplasia (not necessarily at the exclusion of erythroid dysplasia). While the specific mechanism in which benzene interacts with myeloid cells in the bone marrow to produce MDS is not known, these statistical associations show a preference for benzene to ultimately result in MDS involving lineages that are not predominantly erythroid.

Our study suggests that benzene (as well as other agents) may affect different WHO subtypes differently. Most notably, besides the observed effect of benzene on subtypes which are not predominantly erythroid, the elevated risk for smoking shows the opposite effect, as it appears to be restricted to all refractory anemias, including RAEB.

Farm Residence and Agricultural Exposures

Although the primary focus of this study was on benzene exposure, our results are in concert with several previous studies which report that MDS is likely influenced by several endogenous and environmental factors. Farm residence, past or present, was an independent risk factor for most of the MDS subtypes and groupings, and was the only substantial confounder among the covariates. However, it had limited *inferential* impact on the benzene-MDS relation given the strength of the ORs for benzene exposure. In a previous US study of similar design, Strom et al[4] observed a positive exposure–response gradient between total MDS and agricultural chemical exposures in men in their unadjusted results for RA/RARS and for RAEB in multivariable analyses. Exposure to farm/agrichemical exposures has also been reported to decrease survival in some MDS patients.[14]

Other Covariates

Smoking was not positively associated with total MDS or RCMD; however, it showed a positive association with RAEB, especially advanced RAEB (RAEB II, OR = 3.2) and all RAs, OR = 2.6, both of which were statistically significant. Du et al[15] conducted a meta-analysis of 10 case-control studies that examined the relation between smoking and MDS overall, finding an overall risk of 1.45 (95% CI 1.21 to 1.74); all studies except one reported a positive association. One of those studies[4] examined this relation by MDS subtype, observing independent positive associations between smoking and RA/RARS and with RAEB which were consistent with the present study. Smoking is also a well-known risk factor for AML.[16] The specific association of cigarette smoking with RA and RAEB in our study has a biologic basis since these subtypes have a greater propensity for clonal cytogenetic abnormalities versus other MDS subtypes, and cigarette smoke contains genotoxic substances such as PAH's and formaldehyde.

Diabetes was a remarkably persistent protective factor, as fewer cases had had diabetes versus controls. No clear associations between diabetes and MDS have been previously reported.[17–19] Since diabetes was not excluded from control diagnoses, recurring hospitalizations among diabetic controls due to a myriad of potential complications could likely have produced a spurious protective association. Whether the reduced risk of MDS for those with a history of diabetes has a biologic basis or is related to the use of a hospitalized control population cannot be resolved further.

MDS risk is seemingly reduced 3% to 10% for every increased unit of BMI, depending on subtype. Since obesity is associated with diabetes and other chronic diseases, the negative BMI association could, like diabetes, be biased low due to the use of hospitalized controls. Hainer and Aldhoon-Hainerova[20] suggest that a reduced BMI risk for serious diseases (such as MDS) may be due to controls

being more robust (ie, having higher BMIs), since illnesses for which control individuals were hospitalized were likely, on average, to be of lesser clinical severity than MDS. This explanation is further supported considering that BMI showed the most protective relationship with RAEB, the subtype with the worst prognosis.

Educational attainment was associated with all MDS subtypes and with RCMD alone. The educational gradient showed evidence of positive linear trend ($P < 0.01$) from the lowest to highest attainment levels, and also showed a non-confounded independent relation with MDS. Education has frequently been used as a surrogate for socioeconomic position which is often a reliable predictor of many health outcomes.

MDS and Benzene

While benzene has historically been linked to acute leukemia,[21] and acute myeloid leukemias,[22] its link to MDS has been investigated only recently. MDS was formally classified as a disease in 1982, and became reportable in cancer registries as late as 1999. One of the first studies reporting a potential link between benzene and MDS reported excess risk for the combined diagnoses of AML and MDS.[1] More recently, an international pooled analysis in petroleum distribution workers reported significant OR's for MDS in workers exposed to peak benzene exposures more than 3 ppm; these same workers were predominantly exposed to average concentrations less than 1 ppm.[3] More recently, two cohort studies in similar workers have neither confirmed nor refuted these observations.[23,24]

Irons et al[2] using the same database, compared benzene "signal" MDS cases with high (more than 20 ppm) average benzene exposure to other MDS cases with only background benzene exposure. Signal cases were 11.1 times more likely to have MDS-U (based on the 2001 WHO nomenclature) versus non-signal cases. However, this study likely over-represented MDS-U cases that were provisionally diagnosed based on the limited diagnostic information available at case presentation. Based on the more complete follow-up information used in the present study, MDS-U diagnoses were markedly less prevalent: only five of 604 cases in the present study, versus 29 of 649 previously.

Previous studies have revealed that MDS cases with a history of benzene exposure exhibited distinct clinical, immunologic and cytogenetic features.[2,25] Bone marrow findings included evidence of inflammation and eosinophilic dysplasia in addition to multilineage dysplasia. This was accompanied by increased $CD8^+$ T cells with clonal and oligoclonal TCR gene rearrangements in T cell populations.[2,25,26] These findings suggest a role for an autoimmune pathogenesis for typically low-risk MDS cases that is frequently associated with RA and a subset of RCMD with immunologic features. Several authors have demonstrated that a proximate cause of progressive bone marrow cytotoxity in some cases of MDS is the presence of activated $CD8^+$ lymphocytes which produce persistent inflammation, apoptosis, and cytotoxicity via production of $TNF-\alpha$.[27–31] These findings are also consistent with those of Kordasti et al[32] in which cytotoxic $+CD8^+$ T cells predominate in low risk MDS, while increased expression of $Foxp3^+$ regulatory $CD4^+$ T cells correlate with progression of high risk MDS, that is, RAEB.

Benzene has been shown to be non-discriminatory on specific cell line lineages in terms of its non-neoplastic hematotoxic effects. It has long been known to produce pancytopenia,[33] while more recent studies[34,35] show that all prominent cell types (leukocytes, erythrocytes, platelets) are reduced in workers when there is sufficient benzene exposure. This is due to benzene's effect on bone marrow function rather than specific circulating cells per se. The association reported here with RCMD, which also involves multiple lineages, is consistent with previous hematotoxicity observations.

RAEB is the MDS subtype that has the highest rate of progression to AML, and benzene has been linked to AML more consistently than any other hematopoietic cancer.[36] Accordingly, one might have predicted that RAEB would be the subtype most likely to show a relationship with benzene. The fact that our data do not suggest this relationship may point to distinct etiologies for benzene-induced AML and MDS, but this is not necessarily so. In fact, recent data suggest a similar pattern of cytogenetic abnormalities in benzene-induced AML, MDS and de novo cases.[26] Given the limited follow-up data, we could not ascertain whether RCMD cases progressed to either RAEB and/or AML. It is possible that benzene-induced MDS is the stronger response when compared with benzene-induced AML, and spontaneous transformation of benzene-induced MDS to AML (if it occurs at all) may be governed more by stochastic mechanisms rather than continued benzene exposure. Addressing this hypothesis would require sustained follow-up of populations exposed to appreciable levels of benzene.

Study Strengths and Limitations

Strengths include the large number of cases, standardized diagnosis by a single reference laboratory, well-matched controls, and collection of extensive information on occupational exposure, lifestyle, and other individual-level characteristics. Robust retrospective benzene exposure assessment methods were employed in a blinded fashion, with estimates validated by an expert panel. The ability to examine MDS subtypes using the most current WHO diagnostic criteria made concurrent with treatment is another distinctive aspect of this study. Likewise, the high level of post-collection quality checks on the research dataset enhances the validity of the findings.

The primary study limitation is our use of hospital-based controls, a practice that could result in control group that does not represent the base population that gave rise to the cases. However, the controls were matched on age and sex, were admitted close to the case's admission date, and were patients from the same hospitals as the cases, in toto reducing the opportunity for selection bias.[37] QA/QC procedures documented the absence of blood and lymphoid diseases among controls and confirmed that the most frequent hospital departments for controls were cardiovascular, respiratory, internal medicine, and endocrine. Benzene risk estimates would be unaffected unless the control diagnoses were somehow related to a higher or lower opportunity for benzene exposure. Controls were significantly less likely than cases to have ever lived on a farm (34% vs. 48%, $P < 0.01$), which was an independent MDS risk factor in these data. Nevertheless, the potential impact from such a selection bias was lessened by control for farm residence (as a confounder) in the calculation of benzene risk estimates, and by the fact that the study covered a demographically fluid period in China. Specifically, between 2000 and 2010, the rural population of China decreased from 64% to 50%,[38] while the Shanghai population increased from 13.3 million to 16.4 million.[39] While we cannot totally rule out selection bias, the consistency with previous findings on benzene and key covariates provides some degree of certainty and credence to our findings.

Another limitation is the use of questionnaires, subject to recall bias or socially acceptable responses, to collect data on several co-exposures, and lifestyle factors. In addition, we limited this assessment to long-term benzene concentrations and did not study the duration of such exposures, although most workers were long-term workers with very few job changes. While the exposure estimating process was rigorous and subject to validation,[9] some misclassification still likely occurred although the ordinal exposure groupings likely reduced the impact of such errors.

CONCLUSIONS

In this study population, we observed statistically significant associations between benzene and MDS, which were largely driven by RCMD, in which statistically significant associations were observed for concentration estimates of 0.3 to 3 ppm, 3 to

12 ppm, and more than 12 ppm. No associations between benzene exposure and RAEB, the MDS subtype most likely to progress to AML, were observed. Another relationship with higher uncertainty was found for RCUD not classified as RA; together these findings suggest that benzene affects MDS cell types that are not predominantly of erythroid cell lineages. While selection bias is a potential threat to the magnitude of our benzene-specific findings, it would not be expected to differ by MDS cell-type. The findings suggesting benzene's effect on non-erythroid MDS may provide important clues toward uncovering a more precise mechanistic model for benzene-induced MDS.

ACKNOWLEDGMENTS

REFERENCES

1. Hayes RB, Yin S-N, Dosemeci M, et al. Benzene and the dose-related incidence of hematologic neoplasms in China. *J Natl Cancer Inst.* 1997;89:1065–1071.

2. Irons RD, Gross SA, Le A, et al. Integrating WHO 2001–2008 criteria for the diagnosis of myelodysplastic syndrome (MDS): a case-case analysis of benzene exposure. *Chem Biol Interact.* 2010;184:30–38.

3. Schnatter AR, Glass DC, Tang G, Irons RD, Rushton L. Myelodysplastic syndrome and benzene exposure among petroleum workers: an international pooled analysis. *J Natl Cancer Inst.* 2012;104:1724–1737.

4. Strom SS, Gu Y, Gruschkus SK, Pierce SA, Estey EH. Risk factors of myelodysplastic syndromes: a case-control study. *Leukemia.* 2005;19:1912–1918.

5. Gross SA, Irons RD, Schnatter AR, et al. A hospital based case-control study of aplastic anemia in Shanghai, China. *Chem Biol Interact.* 2010;184:165–173.

6. Wong O, Harris F, Armstrong TW, Hu F. A hospital-based case-control study of acute myeloid leukemia in Shanghai: analysis of environmental and occupational risk factors by subtypes of the WHO classification. *Chem Biol Interact.* 2010;184:112–128.

7. Wong O, Harris F, Armstrong TW, Hua F. A hospital-based case-control study of non-Hodgkin lymphoid neoplasms in Shanghai: analysis of environmental and occupational risk factors by subtypes of the WHO classification. *Chem Biol Interact.* 2010;184:129–146.

8. World Health Organization (WHO). In: Swerdlow SH, Campo E, Harris NL, Jaffe ES, PIleri SA, Stein H, Thiele J, Vardiman JW, editors. *Classification of Tumours of Haematopoietic and Lymphoid Tissues.* Lyon: IARC; 2008.

9. Armstrong TW, Liang Y, Hetherington Y, et al. Retrospective occupational exposure assessment for case-control and case-series epidemiology studies based in Shanghai, China. *J Occup Environ Hyg.* 2011;8:561–572.

10. Liang YX, Wong O, Armstrong TW, et al. An overview of published benzene exposure data by industry in China, 1960–2003. *Chem Biol Interact.* 2005;153–154:55–64.

11. Wang SA, Galili N, Cerny J, et al. Chronic myelomonocytic leukemia evolving from preexisting myelodysplasia shares many features with de novo disease. *Am J Clin Pathol.* 2006;126:789–797.

12. Liu H, Liang Y, Bowes SM, et al. Benzene exposure in industries using or manufacturing paint in China-a literature review, 1956-2005. *J Occup Environ Health.* 2009;6:1–12.

13. Gart JJ, Zweifel JR. On the bias of various estimators of the logit and its variance with application to quantal bioassay. *Biometrika.* 1967;54:181–187.

14. Kpola Y, Gbito E, Garcia-Manero G, Strom SS. Evaluation of epidemiologic factors in survival of patients with de novo myelodysplastic syndromes. *Cancer Causes Control.* 2014;25:425–435.

15. Du Y, Fryzek J, Sekeres MA, Taioli E. Smoking and alcohol intake as risk factors for myelodysplastic syndromes (MDS). *Leuk Res.* 2010;34:1–5.

16. Colamesta V, D'Aguanno S, Breccia M, Bruffa S, Cartoni C, La Torre G. Do the smoking intensity and duration, the years since quitting, the methodological quality and the year of publication of the studies affect the results of the meta-analysis on cigarette smoking and acute myeloid leukemia (AML) in adults? *Crit Rev Oncol Hematol.* 2016;99:376–388.

17. Goldberg SL, Chen E, Corral M, et al. Incidence and clinical complications of myelodysplastic syndromes among United States Medicare beneficiaries. *J Clin Oncol.* 2010;28:2847–2852.

18. Hamoudeh E, Zeidan AM, Barbarotta L, Rosano N. The interactions between diabetes mellitus and myelodysplastic syndromes: current state of the evidence and future directions. *Curr Diabetes Rev.* 2016;12:231–239.

19. Mobini M, Shekarriz R, Ali Mohammad Pour R, Zakeri S. Inflammatory rheumatologic disorders in patients with myelodysplastic syndromes: a cross-sectional study. *Int J Hematol Oncol Stem Cell Res.* 2015;9:22–25.

20. Hainer V, Aldhoon-Hainerova I. Obesity paradox does exist. *Diabetes Care.* 2013;36(suppl 2):S276–S281.

21. Vigliani EC, Saita G. Benzene and leukemia. *N Engl J Med.* 1964;271:872–876.

22. Aksoy M. Different types of malignancies due to occupational exposure to benzene: a review of recent observations in Turkey. *Environ Res.* 1980;23:181–190.

23. Collins JJ, Anteau SE, Swaen GMH, Bodner KM, Bodnar CM. Lymphatic and hematopoietic cancers among benzene-exposed workers. *J Occup Environ Med.* 2015;57:159–163.

24. Sorahan T, Mohammed N. Incidence of myelodysplastic syndrome in UK petroleum distribution and refinery workers, 1995–2011. *Int J Environ Res Public Health.* 2016;13:474.

25. Irons RD, Lv L, Gross SA, et al. Chronic exposure to benzene results in a unique form of dysplasia. *Leuk Res.* 2005;29:1371–1380.

26. Irons RD, Kerzic P. Cytogenetics in benzene-associated myelodysplastic syndromes and acute myeloid leukemia: new insights into a disease continuum. *Ann NY Acad Sci.* 2014;1310:84–88.

27. Maciejewski JP, Selleri C, Anderson S, Young NS. Fas antigen expression on CD34+ human marrow cells is induced by interferon gamma and tumor necrosis factor alpha and potentiates cytokine-mediated hematopoietic suppression in vitro. *Blood.* 1995;85:3183–3190.

28. Molldrem JJ, Jiang YZ, Stetler-Stevenson M, Mavroudis D, Hensel N, Barrett AJ. Haematologic response of patients with MDS to antihymocyte globulin associated with a loss of lymphocyte-mediated inhibition of CFU-GM and alterations in T-cell receptor Vbeta profiles. *Br J Haematol.* 1998;102:1314–1322.

29. Voulgarelis M, Giannouli S, Ritis K, Tzioufas AG. Myelodysplasia-associated autoimmunity: clinical and pathophysiologic concepts. *Euro J Clin Invest.* 2004;34:690–700.

30. Braun T, Fenaux P. Myelodysplastic syndromes (MDS) and autoimmune disorders: cause or consequence:. *Best Pract Res Clin Haematol.* 2013;26:327–336.

31. Calado RT. Immunologic aspects of hypoplastic myelodysplastic syndrome. *Semin Oncol.* 2011;38:667–672.

32. Kordasti SY, Ingram W, Hayden J, et al. CD4+CD25highFoxp3+ regulatory T cells in myelodysplastic syndrome (MDS). *Blood.* 2007;110:847–850.

33. Aksoy M, DinCol K, Akgun T, Erdem S, DinCol G. Haematological effects of chronic benzene poisoning in 217 workers. *Br J Ind Med.* 1971;28:296–302.

34. Schnatter AR, Kerzic PJ, Zhou Y, et al. Peripheral blood effects in benzene-exposed workers. *Chem Biol Interact.* 2010;184:174–181.

35. Lan Q, Zhang L, Li G, et al. Hematotoxicity in workers exposed to low levels of benzene. *Science.* 2004;306:1774–1776.

36. Schnatter AR, Rosamilia K, Wojcik NC. Review of the literature on benzene exposure and leukemia subtypes. *Chem Biol Interact.* 2005;153–154:9–21.

37. Rothman KJ, Greenland S, Lash TL, editors. *Modern Epidemiology.* 3rd ed. Philadelphia, PA: Lippincott, Williams & Wilkins; 2008.

38. China National Bureau of Statistics. Statistical Year Book of China; 2012.

39. Shanghai Bureau of Statistics. Basic Statistics on National Population Census. Available at: http://www.stats-sh.gov.cn/tjnj/nje11.htm?d1=2011tjnje/E0226.htm. Accessed June 9, 2015.