Chemicals and childhood leukemia

Claire Infante-Rivard MD, PhD
McGill University, Montréal, Canada

Currently at Inserm UMR-S 754, Paris, France
Supported by a UICC Yamagiwa-Yoshida Memorial International Cancer Study Grant

Plan

- Included and excluded "chemicals"
- Brief review of published results on
  - the association between these chemicals and leukemia
  - effect of these chemicals on leukemia as modified by gene variants (gene-environment interactions)
- Brief comments on these studies
- Suggestions for the next steps (focus on additional use of available studies but applicable to new studies)

Included and *excluded* chemicals

- *Excluded* elements (because covered elsewhere today): pesticides, parental (pre-conceptional, prenatal) occupational exposure

Included and excluded chemicals

- Environmental exposures were most often studied as:
  - mixtures from sources (e.g., "hydrocarbons" from proximity to traffic & garages, "solvents" from paint, environmental tobacco smoke as marker for many chemicals)
  - occasionally, specific contaminants or families of contaminants were studied (e.g., trihalomethanes in chlorinated water, or benzene outside the residence)
  - The best classification of pollutants we could achieve with the leukemia studies is probably: organic and inorganic (not helpful!)
- Classifications of chemicals are very rare in the relevant epidemiological papers
  - See one example Infante-Rivard et al. EHP 2005

Table 1. Matrix of specific chemicals, complex mixtures of chemicals, and chemical families used in the analysis.

	Code[a]	Chemical families[b]						
		1	2	3	4	5	6	7
Specific chemicals								
Methanol	232		XX[c]					
Ethanol	233		XX					
Isopropanol	234		XX					
Ethylene glycol	235		XX					
Carbon tetrachloride	237			XX				
Chloroform	238			XX				
Methylene chloride	239			XX				
1,1,1-Trichloroethane	240			XX				
Trichloroethylene	242				XX			
Perchloroethylene	243				XX			
Ethylene dichloride	300				XX			
Acetone	248					XX		
Methyl ethyl ketone	304					XX		
Benzene	252						XX	
Toluene	253						XX	
Xylene	254						XX	
Ethyl acetate	302							XX
Diethyl ether	250							
Turpentine	280							
Carbon disulfide	266							
Butyl cellosolve	306							
Mixtures								
Mineral spirits post-1970[d]	202	X[e]					X	
Mineral spirits pre-1970[d]	203	X					X	
Leaded gasoline	191	X					X	
Unleaded gasoline	299	X					X	
Aviation gasoline	190	X					X	
Kerosene	195	X					X	

[a]These codes were used by Siemiatycki (1991) to catalogue and define the various substances, and they can thus be used to easily find additional information on these chemicals in that reference. [b]Chemical families: 1, alkanes (C5–C17); 2, aliphatic alcohols; 3, chlorinated alkanes; 4, chlorinated alkenes; 5, aliphatic ketones; 6, mononuclear aromatic hydrocarbons; 7, aliphatic esters. [c]XX signifies that the agent listed to the left is a member of the chemical family indicated at the top. [d]Before 1970, mineral spirits contained relatively higher amounts of benzene, toluene, and xylene due to ignorance of their toxic effects. [e]X signifies that the agent listed to the left contains components that are members of the chemical family indicated at the top.

Included and *excluded* "chemicals"

- Included, are studies published between *1998-2008* on:
  - Parental smoking
  - Outdoor pollution (traffic, industries)
  - Indoor contaminants
  - Water contaminants
- Focus on acute lymphoblastic leukemia (ALL)
- Time window for exposure: preconception, prenatal, postnatal
- Excluded (but could have been included!) are studies on:
  - Medication, recreational drugs, and diet
  - Parental alcohol consumption
  - Cluster-related studies that could involve chemicals

Review of studies-Parental smoking

- 12 case-control reports (3 from same French study) and 1 registry-based cohort study
- Exposure assessment
  - parental reporting at time of diagnosis or ascertainment, or smoking information at first prenatal care visit as entered on birth registry (1 study)
- Results are quite consistently negative

Review of studies-Parental smoking

Study author, year, country	N cases (ALL)	Father Preconception (a,b,c)	Mother pregnancy	Post-natal mother	Post-natal father
Brondum '99 US	1842	NS (a)	NS		
Schuz '99 Germany	1037	NS (b)	NS		
Infante-Rivard '00 Quebec, Canada	491		NS	NS	NS
Alexander '01 UK	49		NS		
Sorahan '01 UK	148	**S* (c)**			
Ocku '02 Texas, US	83		NS		
Pang '03 UK	1375	NS (b)	NS		
Mucci '04 Sweden	505		**S*** (negative)		
Chang '06 Northen Calif.	281	NS (b)	NS	NS	
Clavel '05 Menegaux '05, '07 France	407	NS (b)	NS S ('07)	NS	NS
MacArthur '08 Canada	351	NS (c)	NS (father NS)		

Review of studies-outdoor pollution

- 10 case-control reports from 9 studies
- 4 ecological studies (area exposure associated with geographical area incidence data
- Exposure assessment
  - varied, some using sophisticated modeling
  - most exposure estimates apply to an area near the home (and at diagnosis only), but a few are specific measures (and occasionally of specific contaminants) at the home
- Results mostly negative but not as consistently so as with parental smoking

Review of studies-indoor pollution: case-control studies (1)

Author, year, Country	N cases of leukemia (ALL) Study period	Elements of the Methods	Main results S=statistically significant
Feychting' 98 Sweden	39 1976-83 *Living near power lines	-when? -peak NO_2 concentrations for 1-hour averages over 1 year	NS
Pearson '01 Denver US	97 (78 ALL) 1976-1983	-exposure estimates apply to address at diagnosis -distance-weighted traffic density based '79, '90 -expressed in vehicles per day (VPD)	S Increased risk at highest VPD
Raaschou-Nielsen '01 Denmark	900 (731 ALL) 1968-1991	-exposure estimates apply to period from pregnancy to diagnosis -extensive modeling -traffic density -NO_2 and benzene at front door	NS
Langholz '02 Los Angeles US	212 1978-1984	-exposure estimates apply to home at which child resided longest -distance weighted traffic density	NS
Reynolds '04 California US	90 1988-97	-exposure estimates apply to birth period -Road and traffic density	NS

Review of studies-indoor pollution: case-control studies (2)

Author, year, Country	N cases of leukemia (ALL) Study period	Elements of the Methods	Main results S=statistically significant
Steffen' 04 France	280 (240 ALL) 1995-99	-Y/N exposure applies to period between and diagnosis -property adjoining neighboring business -traffic in vicinity	S property adjoining garage or petrol station)
Crosignani '04 Northern Italy	120 1978-1997	-exposure estimates apply to address at diagnosis -mean annual concentration of benzene outside home	S increased risk at highest level
Yu '06 Taiwan	94 (ALL) 1997-2003	-exposure estimates apply to period from two years before birth to diagnosis -score based on proximity to petrochemical complexes	NS
Weng '08 Taiwan	308 dead cases 1995-2005	-exposure estimates apply to home at death -NO_2 yearly average from monitoring stations	S increased at highest levels
Weng '08 Taiwan	404 dead cases 1995-2005	-exposure estimates apply to home at death -proportion of workers in municipality employed in petrochemical industry	S at highest percentile

Review of studies-outdoor pollution: ecological studies

Author, year, Country	N cases of leukemia (ALL) Study period	Elements of the Methods	Main results S=statistically significant
Harrison '99 UK	Regional incidence Leukemias 1990-1994	- exposure applies to address at diagnosis -living <=100 m road of high traffic density or petrol station or both	NS
Reynolds '02 California	Regional incidence Leukemias and ALL 1988-1994	-exposure estimates apply to address at diagnosis -vehicle density, road density, traffic counts correlated with ambient air monitoring data	NS
Reynolds '03 California	Regional incidence Leukemias and ALL 1988-1994	-exposure estimates apply to address at diagnosis - score for 25 hazardous air pollutants based on their cancer potency and modeled outdoor air concentrations	S at highest level for combined mobile, area, and point sources
Visser '04 Amsterdam, NL	Regional incidence Leukemias 1989-1997	-exposure estimates apply to address at diagnosis -traffic intensity score along main roads	NS

Review of studies: indoor air contaminants

- Only two studies found
- Both are large and used the same questionnaire
- Control selection was different
- Results are apparently contradictory

Review of studies: indoor air contaminants

Author, year, Country	N cases of leukemia (ALL) Study period	Elements of the Methods	Main results S=statistically significant
Freedman '99 US	640 ALL 1989-1993	-controls selected by RRD (date not given) -potential exposure to solvents by questionnaire -through hobbies and household maintenance activities (applies to year before diagnosis) -through house painting (applies to preconception, pregnancy, and postnatal periods) -emphasis on frequency of activities	S Frequent art work using solvents Mothers living in homes with extensive painting prior to birth
Infante-Rivard '05 Quebec, Canada	790 ALL 1980-2000	-Controls selected from census-like sampling frame at time of case diagnosis -potential exposure to solvents by questionnaire for preconception, pregnancy, and postnatal periods -same activities assessed as in Freedman with less emphasis on frequencies	NS

Review of studies: water contaminants

- Two studies: one case-control and one ecological
- Exposure assessment
  - Specific contaminants measured
  - detailed in the C-C study including ecological and individual measurements
  - vague and limited in the ecological study
- Both studies show negative results

Review of studies: water contaminants

Author, year, Country	N cases of leukemia (ALL) Study period	Elements of the Methods	Main results S=statistically significant
Infante-Rivard '01 Quebec, Canada	490 ALL 1980-1993	-Controls selected from census-like sampling frame at time of case diagnosis -exposure applies to prenatal and postnatal periods -exposure to specific and total concentrations of trihalomethanes, and to specific metals -based on municipal and provinvial regulatory data and on tap water samples (the latter collected after diagnosis)	NS
Moore '02 Nevada, US	Regional incidence Leukemia 1979-1999	-exposure estimates apply to address at diagnosis -arsenic in water based on health department data (limited information)	NS

Review of studies-gene-environment

- Four reports (for the chemicals reviewed here) from only two studies
- Analysis of interaction from case-control, case-only, and case trios designs
- Some positive signals from these limited data

Review of studies-gene-environment

Author, year, Country	Type of analysis for GxE N subjects	Genes Environmental factors	Main results S=statistically significant
Infante-Rivard '00 Québec, Canada	Case-only 158 ALL cases	CYP1A1*2A, *2B, *4 Maternal smoking during pregnancy; paternal smoking from birth to dagnosis	NS increased risk at highest evel during pregnancy with *2A and *4 S decreased risk at intermediate level during pregnancy with *2B
Clavel '05 France	Case-control Case-only 219 (195 ALL) cases 105 controls	-CIP1A1*2A, GST (M1, T1, P1), NQO1*2, EPHX1*3, *4 -Maternal smoking during pregnancy	S ever vs never for CYP1A1*2A and GSTM1 in case-only analysis
Infante-Rivard '07 Québec, Canada	Case-trios (M, F, C) 654 families (ALL)	-NQO1*2 -Maternal smoking during pregnancy by trimester	NS
Infante-Rivard '02 Québec, Canada	Case-only 161 ALL cases	-GSTT1, CYP2E1*5 -Total trihalomethanes from pregnancy to date of diagnosis, bromoform, chloroform	S for GSTT1 and total THM NS but increased for CYP2E1*5 and highest level of chloroform

Brief comments on the reviewed studies

- Study design:
  - Often, all types of leukemias analyzed together instead of ALL only
  - small study size
  - non-concurrent selection of controls by RRD at unspecified time but likely at time of study rather than at time of case incidence (i.e., not incidence density sampling)
  - reporting of participation incomplete

Brief comments on the reviewed studies

- Study for G x E
  - C-C inefficient and vulnerable to population structure bias (PSB)
  - C-O very efficient but requires independence assumption of G-E in controls, and is also vulnerable to PSB
  - The case-parent trio approach (to study departure from expected transmission of the alleles) not vulnerable to PBS; relaxed independence assumption; more costly
  - Sample size issue

Sample size: alpha=.05, power=.8, allele freq=.05, prevalence of E=.10, Re=1.5, Rg=1.5 Rge=3
Dominant model
Goal: study G x E

G x E	955 C-C pairs	830 sib pairs	728 case-parent trios	251 cases only

Brief comments on the reviewed studies

- Exposure assessment: parental smoking
  - Assuming that parental smoking is measured adequately, results for maternal smoking during pregnancy are very disappointing and do not point to strong environmental effects
  - However, G x E analyses may change this perspective

Brief comments on the reviewed studies

- Exposure assessment: outdoor pollutants
  - major challenge
  - most studies measured exposure near the address at diagnosis, and only for that one
  - one C-C study (Raaschou-Nielsen) developed sophisticated measures for appropriate time window, with contaminant measures at the home, and included a validation sub-study (results were negative)

Brief comments on the reviewed studies

- The rate of published studies on environmental factors and childhood leukemia seems to have gone down in recent years
- The more recent studies using dead cases are not ideal
- Overall, no environmental risk factor (among the reviewed ones) comes out as a strong determinant for childhood leukemia
- Is that the end of the story?

Some suggestions for next steps

- Assuming that:
  - new large C-C studies are not going to be available in the immediate future
  - little can be done in existing studies to markedly improve exposure assessment methods (another chapter!)
  - most studies from which the results were presented here have collected a set of risk factor data that are relatively similar
  - many have or are collecting genetic material
  - most PIs have published basic analyses on this set of risk factors
- What are some possible strategies to maximize use of collected data?

Some suggestions for next steps

- 1. Pooling of data for E and G x E effects
  - Advantages:
    - Increase in study power
    - Possibly revealing meaningful differences (if any) leading to further hypotheses
  - Disadvantages:
    - Requires non trivial amount of work to harmonize the data; and may not always be possible
    - Some loss of ownership

Some suggestions for next steps

- 2. Pooling of analyses (within and across studies) using case-control and family-based data together (a number of studies have collected genetic material from case (and occasionally control) parents)
  - Such statistical genetics methods are available and some are developing ("hybrid models")
  - Advantages:
    - Increased efficiency; combined estimate; formal testing for PSB before combining estimates
  - Disadvantages:
    - Not trivial or mainstream yet

Some suggestions for next steps

- 3. A more in-depth review of published results (not to criticize but maybe to better understand the results)
  - List of detailed methodological criteria:
    - particularly related to selection and participation of controls (selection bias)
    - also as related to interviewing
  - Quantitative probabilistic assessment of sensitivity to misclassification, selection and confounder bias in published studies

Some suggestions for next steps

- 4. Combining ecological and individual study designs/data
  - e.g., studying the relation between a water contaminant in distribution systems and rates of cancer in a region
  - Inference limitations if there is within-area variability for the main and related contaminants as well as confounding (e.g., age and gender distribution)
  - Based on levels of contaminant under study, develop strategy for phase 2 sampling of individuals (case-control status, confounders, home samples, etc)

Some suggestions for next steps

- Maybe with some counter-matching (cases and controls selected for future finer exposure assessment based on being in opposite categories of the proxy E measure
- Use both phase I and phase II data in the analysis
- Power/efficiency of individual exposure data may be increased while ecological bias (unmeasured area-level variables or factors that vary between individuals) reduced
- Methodological refinements are in development and related analyses (and software codes) as well

Conclusions

- So far none of the environmental contaminants covered in this talk have emerged as a strong RF for leukemia
- Fewer studies being published recently
- Turning point:
  - Wait for new and better studies! or revisit the considerable potential of existing ones with feasible strategies
  - Urgently: “Go genetics!”