

**AML and NHL Case-control Study
Shanghai Health Study**

Progress Report
Activities through December 2006

Prepared by:
Otto Wong, Sc.D.
Applied Health Sciences, Inc.
Post Office Box 2078
San Mateo, CA 94401
USA

Tel: 1 (650) 347-7898 USA
Fax: 1 (650) 344-6887 USA
ottowong@aol.com
ottowong@shmu.edu.cn

15 February 2007

◆◆◆◆◆◆◆◆

Status of Patient Enrollment

Currently, there are 33 hospitals in Shanghai participating in the Shanghai Health Study (SHS); consisting of 17 municipal hospitals, 13 district hospitals, 2 occupational disease hospitals and one private hospital. These hospitals contribute patients to all three research studies in the SHS. The hospitals, however, vary greatly in size. Twelve hospitals have not contributed any patients during the time period June-December 2006. Some of these hospitals are relatively small, but one of them (No. 2 Medical University Hospital) is a large hospital and should have a large population of patients. Therefore, case ascertainment at some hospitals might not be complete.

In terms of the case-control study, according to the numbers provided by Dr. Richard Irons on 13 February 2007, the cumulative numbers of patients diagnosed with AML and NHL (based on ICD-9 definition) as of 31 December 2006 were as follows:

- 423 AML patients
- 410 "NHL" patients (214 or 52% with interviews)

Using the broader WHO definition, the additional cumulative numbers of AML and lymphoid neoplasm patients as of 31 December 2006 were:

- 207 AML patients
- 230 other lymphoid neoplasm patients (213 with interviews)

The numbers of AML and NHL/lymphoid neoplasm patients reported above must be interpreted with caution. The numbers represent patients diagnosed with AML or NHL/lymphoid neoplasm at the participating hospitals. These numbers do not represent the numbers of AML or NHL/lymphoid neoplasm patients with questionnaire information. Many patients, especially those diagnosed with NHL (ICD-9), do not have questionnaire information. As indicated above, only approximately half of the NHL patients were interviewed.

I first noticed that some NHL patients did not have questionnaire information when I examined the "test datasets" of AML and NHL patients provided to us by Gail Jorgensen of ExxonMobil in July 2006, shortly before the Houston meeting on 1-2 August 2006. Of the 150 NHL patients in the test dataset, only 61 patients (41%) had questionnaire information. The other 59% NHL patients did not have any questionnaire information, and, therefore, would eventually have to be excluded from the case-control study. This problem, which would have a major impact on the case-control study, was reported to the Scientific Review Panel at the August 2006 meeting in Houston and also reported in our semi-annual progress report dated 9 August 2006, which was distributed to both panel members and sponsors. At the time of the August meeting, I did not know the reason for the missing questionnaires.

During my September trip to Shanghai, I investigated the problem of NHL patients without questionnaires. I talked to a number of people (Pei Xiaodan, Guo Xiao, Xie Junfang, Wang Yiyang, Chen Jian, and Fu Hua) and was surprised to learn that from January 2006 to early September 2006, fewer than 10 patients at the Tumor Hospital were interviewed, even though the number of patients diagnosed with NHL during the same time period was much greater. The majority of NHL patients in the SHS were patients at the Tumor Hospital. It appeared that the Clinical Coordinators at the Tumor Hospital did not make arrangement for the NHL patients to be interviewed after the patients' laboratory reports (diagnosis) were sent to the hospital.

Xie Junfang is the person at Fudan whom the participating hospitals notify with new cases and controls. Xie Junfang, in turn, notifies Guo Xiao at Fudan who actually sends out interviewers to hospitals. Xie Junfang keeps a spreadsheet of all NHL patients as well as a list of NHL patients with interviews.

The following is a summary of NHL patients with and without questionnaires by year.

Year	Total no. of NHL patients	NHL patients with questionnaires	Row percentage
2003	39	26	67%
2004	93	53	57%
2005	137	48	35%
2006 (up to mid-September)	98	9	9%
Total	367	136	37%

The distribution of NHL patients with questionnaires by hospital is as follows:

Hospital	No. of NHL patients with questionnaires	Column percentage
Tumor Hospital	91	67%
Huadong Hospital	16	12%
Tongi Hospital	9	7%
Jiangding Hospital	3	2%
Zhabei Hospital	1	1%
Shuguang Hospital	6	4%
Changning Hospital	1	1%
Changhai Hospital	1	1%
No. 5 Renmin Hospital	5	4%
Ruijin Hospital	2	1%
Huangpu Hospital	1	1%
Total	136	100%

Approximately two-thirds of the NHL patients with questionnaires are from the Tumor Hospital. It appears that many NHL patients at the Tumor Hospital were not interviewed, because the Clinical Coordinators at the Tumor Hospital did not notify Xie Junfang of cases and controls for interview.

The problem of NHL patients without an interview or questionnaire has a major impact on the case-control study and on the overall SHS as a whole. The only information we have on these NHL patients consists of their names, hospitals and approximate dates of admission. These patients, without detailed questionnaire information, will not be included in our analysis.

As many of the patients diagnosed with NHL (approximately 61%) were patients at the Tumor Hospital, it was important to take immediate actions not only to capture all newly diagnosed patients from now on but also to recapture

as many missing patients as possible at the Tumor Hospital. In October a retired nurse from the Tumor Hospital, Lu Yunhua (陆云花), was hired to deal with patient accrual at the Tumor Hospital. I met with Lu Yunhua several times in October and November. In addition to newly diagnosed NHL patients, she also tried to recapture some of the previously diagnosed patients who were not interviewed. Below is a summary of Lu’s work in November and December 2006 (based on information provided to Wang Yiyang by Lu).

(A) Newly diagnosed patients:

No. of patients registered	No. contacted by Lu	No. agreed to participate	No. refused to participated	No. actually interviewed
27	26	14	12	12

(B) Previously diagnosed patients who had not been interviewed:

No. of patients searched for contact info	No. with phone number or other info	No. of contacts attempted by Lu	No. successfully contacted	No. agreed to participate	No. actually interviewed
205	167	149	34	24	16

It appeared that only a small number of the “missing” NHL patients could be recaptured, although it may take a couple more months before a definite estimate can be made. In any event, to make up for the missing patients, Dr. Richard Irons proposed that patient accrual be extended to the end of 2007.

Power Calculation

The original target sample size was 500. For NHL, even if the end of case accrual be extended to the end of 2007, the actual sample size may still fall short of the original target. The sample size of NHL patients will likely be between 300 and 400, depending on the level of success in enrolling both newly diagnosed patients and previously diagnosed patients who were not interviewed. The following table shows the smallest risk ratios that analyses based on different sample sizes will be able to detect. This table is useful in estimating the statistical power of the overall analysis of the entire study as well as detailed analyses of subgroups (such as diagnostic subtypes). As can be seen, the smallest detectable risk ratios depend on the exposure rate in the controls. Currently, I do not have a good estimate for that, but it should be around 2% or 3%. The statistical parameters in the power calculation are as follows:

- Control-to-case ratio = 2:1,
- Alpha = 0.05,

- Power = 0.80, and
- One-sided test.

Exposure rate in controls	Minimum detectable risk ratios						
	25 cases	50 cases	100 cases	200 cases	300 cases	400 cases	500 cases*
5%	6.72	4.34	3.04	2.29	2.01	1.85	1.74
4%	7.72	4.87	3.34	2.47	2.14	1.96	1.84
3%	9.34	5.74	3.82	2.76	2.35	2.13	1.99
2%	12.54	7.42	4.74	3.27	2.73	2.44	2.25
1%	21.96	12.25	7.28	4.67	3.74	3.24	2.95

* Original target sample size.

Exposure Assessment

Under the direction of Dr. Thomas Armstrong, the exposure assessment team continues to develop exposure data/estimates using the "sector analysis" approach. The analysis for two commonly encountered industries (shoe and rubber) has been completed. A more thorough discussion of exposure assessment will be provided by Dr. Armstrong in a separate report.

Exposure data from other sources are also being collected, including published literature. According to Professor Liang, who heads the literature review task, preliminary exposure data for the following occupations/industries have been completed: shoemakers, painters, paint manufacturers, printers, dyers, and rubber workers. These literature data will supplement the Shanghai Municipal IPHS database and district IPHS data in the development of job-exposure matrix (JEM).

The report of benzene exposure in the shoe industry based on the Chinese medical literature has been completed and a shortened version has been published in a special issue of *Regulatory Toxicology and Pharmacology* entitled "The Development and Regulation of Occupational Exposure Limits in Asia," which was published in November 2006. The title of the paper is "Benzene exposure in the shoemaking industry in China, a literature survey, 1978-2004" (Wang L, Zhou Y, Liang Y, Wong O, Armstrong T, Schnatter AR, Wu Q, Fang J, Ye X, Fu H and Irons R; *Regul Toxicol Pharmacol* 2006;149-156). A copy of the paper is attached.

A similar report of benzene exposure in another two industries, paint manufacturing and painting, has been completed. The report in draft form is now being reviewed internally. After the internal review, the report will be forwarded to the Scientific Review Panel for approval for journal submission.

QA/QC

Pei Xiaodan, who is responsible for QA/QC, has been providing monthly reports on a regular basis. These monthly reports include accuracy checks of the following items: primary and secondary questionnaires, data entry and initial exposure assessment.

QA/QC of questionnaires taken between July and December 2006 indicated that 96% of primary questionnaires sampled for review were "qualified" (i.e., no problems were found), which was an improvement over the last time period. During the same time period, for patients with selected occupations of interest, 99% of their secondary questionnaires were "qualified."

Once the questionnaires are completed, the information is entered into the SHS database at Fudan. The overall data entry error from July to December 2006 was 3%, which is acceptable.

During my trips to Shanghai in September, October and November, I conducted three *ad hoc* QA checks of samples of both primary and secondary questionnaires taken in the last six months. I identified only a few (less than 10) minor mistakes in these questionnaires (with literally thousands of items/entries) per QA check. These results reinforce/validate the routine QA/QC checks by Pei Xiaodan.

Epidemiology Review of Risk Factors of AML and NHL

In anticipation of analysis planning and report writing, we have started reviewing epidemiologic studies of AML and NHL in relation to occupational and environmental risk factors. We are particularly interested in studies that provide information on subtypes of AML and NHL. We have collected the abstracts of more than 300 potentially relevant papers. We are currently reviewing the abstracts and retrieving selected papers. Wang Yiyang at Fudan is assisting me in this review.

Project Management and Timeline

Communication with the Fudan team in Shanghai remains a high priority item. In addition to email and telephone calls, I made 3 trips to Shanghai in the second half of 2006.

Professor Fu and his staff remain as my primary contact in Shanghai for the case-control study. Professor Fu's group continues to monitor and update me

the progress of the study when I am not in Shanghai. His group also provides the necessary support when I work on the study in Shanghai.

As discussed above, Dr. Irons proposed that the end of patient accrual be extended to the end of 2007. My current understanding is that the proposed extension will be accepted. Given the new end date for case accrual, final diagnoses, control selection, and interviews will most likely be completed around the end of the first quarter of 2008. Exposure assessment of newly enrolled patients (cases and controls) will be conducted in the first and second quarters of 2008. Assuming that both diagnostic and exposure data for all patients will be finalized and available by June 2008, we will be performing epidemiologic analyses of the final data in the third quarter of 2008.

To facilitate data transfer/linkage and to develop plans for statistical analyses, however, we need relatively large and representative samples of exposure estimates and diagnoses of study subjects by mid-2007 (see the timeline table below). These samples would allow us to develop plans for data transfer/linkage and to test-run analysis programs.

The timeline of the case-control study depends heavily on the timing when diagnostic and exposure data (both test data samples and final complete data sets) are provided to us. The delay of any one event in the timeline will likely affect the others. The revised timeline below for the case-control study is based on our current understanding and assumptions.

Revised timeline of the case-control study

Tasks	Date
Test data sets to AHS*	June 2007
QA/QC of test data sets	July-September 2007
Development of analytical programs/plans	October 2007-March 2008
End of patient accrual*	December 2007
Final diagnoses of all patients completed*	March 2008
Interviews of all patients completed*	March/April 2008
Questionnaire data entry completed*	May 2008
Exposure assessment completed*	June 2008
Final QA/QC and data editing	July-August 2008
Complete final data to AHS*	End of June 2008
Statistical/epidemiological analysis	August-September 2008
Preparation of draft report	September-October 2008
Presentation of study results	October 2008
Discussion with SRP	October-November 2008
Preparation of final report	October-December 2008

* Activities not under our control.