DEVELOPMENT OF A HEALTH EFFECTS-BASED PRIORITY RANKING SYSTEM FOR AIR EMISSIONS REDUCTIONS FROM OIL REFINERIES IN CANADA

Prepared for the Health Prioritization Sub-group of the National Framework for Petroleum Refinery Emission Reductions (NFPRER), a multi-stakeholder initiative of the National Air Issues Coordination Committee – Other Air Issues (NAICC-A)

Final Report

April 20, 2004

Network for Environmental Risk Assessment and Management (NERAM) and the Institute for Risk Research

Stephen McColl (Principal Investigator)
Stephanie Gower (Research Associate and Chief Program Author)
John Hicks, John Shortreed, Lorraine Craig (Co-investigators)

TABLE OF CONTENTS

ACRONYMS AND ABBREVIATIONS

ADI	acceptable daily intake
CalEPA	California Environmental Protection Agency
CAC	criteria air contaminants
CCME	Canadian Council of Ministers of the Environment
CEPA	Canadian Environmental Protection Act
CNS	central nervous system
COPD	chronic obstructive pulmonary disease
CR	concentration response
CWS	Canada-wide standards
DALY	disability adjusted life year
DC	degraded concentration
DCbg	degraded concentration plus background
DW	disability weight
EP	exposed population
EPA	US Environmental Protection Agency (USEPA)
GIS	geographic information system
HAP	hazardous air pollutants
HEAST	Health Effects Assessment Summary Tables (USEPA)
HEC	human equivalent concentration
HEIDI	Health Effects Indicators Decision Index
HPSG	Health Prioritization Sub-group
IRIS	Integrated Risk Information System (USEPA)
ITER	International Toxicity Estimates for Risk
MB	Mantel-Bryan extrapolation model
MOE	Ontario Ministry of the Environment
MTBE	methyl-t-butyl ether
NERAM	Network for Environmental Risk Assessment and Management
NFPRER	National Framework for Petroleum Refinery Emission Reductions
NOx	nitrogen oxides
NOEL	No Observed Effect Level

NOAEL	No Observed Adverse Effect Level
NPRI	National Pollutant Release Inventory
NTP	U.S. National Toxicology Program
PAH	polyaromatic hydrocarbon
PBT criteria	persistence, bioaccumulation and toxicity
PM	particulate matter (including both PM_{10} and $PM_{2.5}$)
$PM_{2.5}$	particulate matter less than 2.5 micrometers in diameter
PM_{10}	particulate matter less than 10 micrometers in diameter
POPs	persistent organic pollutants
PSL1	Priority Substances List 1
QRA	quantitative risk analysis
QALY	quality adjusted life year
REL	reference exposure level
RfC	reference concentration
RfD	reference dose (USEPA)
RP	response parameter
RSEI	Risk-Screening Environmental Indicators Model (USEPA)
SOx	sulphur oxides (including SO_2 and sulphates)
$T_{1/2}$	degradation half life
TC	Tolerable Concentration (Health Canada)
TCDD	dioxin (Tetrachlorodibenzo-p-dioxin)
TDI	Tolerable Daily Intake (Health Canada)
TERA	Toxicology Excellence in Risk Assessment
TRI	Toxic Release Inventory (US)
TW	toxicity weight (USEPA)
UF	uncertainty factors
VOCs	volatile organic compounds
WHO	World Health Association

GLOSSARY OF TERMS

Abatement	The reduction in degree or intensity of pollutant emissions.
Air toxics	Toxic air pollutants, also known as hazardous air pollutants, are those pollutants that cause or may cause cancer or other serious health effects, such as reproductive effects or birth defects, or adverse environmental and ecological effects.
Cost benefit analysis	An economic technique applied to public decision-making that attempts to quantify in dollar terms, the advantages (benefits) and disadvantages (costs) association with a particular policy option.
Criteria air pollutants	An air pollutant for which acceptable levels of exposure can be determined and for which an ambient air quality standard has been set. Examples include: ozone, carbon monoxide, nitrogen dioxide, sulfur dioxide, and particulate matter.
Degraded concentration	The residual air concentration of each substance calculated as Emission Concentration x f(T1/2) where f(T1/2) is a function of the degradation half-life of the toxic in air.
Disability Adjusted Life Year	DALY is a measure of the burden of disease that reflects the total amount of healthy life lost including time lived with a disability and the time lost due to premature death. The DALY strives to tally the complete health burden associated with a particular disease. Key elements in the calculation of the DALY include i) duration of time lost at each age due to death, ii) disability weights or degrees of incapacity or suffering associated with different non-fatal conditions, iii) age-weights, which indicate the relative importance of healthy life at different ages and iv) time preference, which is the value of health gains today compared to the value attached to health gains in the future.
Effective concentration	Concentration of a substance that causes a defined magnitude of response in a given system: EC50 is the median concentration that causes 50 % of maximal response.
Effective dose	Dose of a substance that causes a defined magnitude of response in a given system: ED05 is the median dose that causes 5% of maximal response.
Half life	Time in which the concentration of a substance will be reduced by half, assuming a first order elimination process or radioactive decay.
Human equivalent concentration	Exposure concentration for humans that has been adjusted for dosimetric differences between experimental animal species and humans to be equivalent to the exposure concentration associated

	humans to be equivalent to the exposure concentration associated with observed effects in the experimental animal species. If occupational human exposures are used for extrapolation, the human equivalent concentration represents the equivalent human exposure concentration adjusted to a continuous basis.
Inhalation unit risk	The Inhalation Unit Risk is the upper-bound excess lifetime cancer risk estimated to result from continuous exposure to an agent at a concentration of 1 ug/m^3 in air.
Intake fraction	The intake fraction *(iF)* is the fraction of chemical mass emitted into the environment that eventually passes into a member of the population through inhalation, ingestion, or dermal exposure *iF* provides a simple, transparent and potentially comprehensive measure of the relationship between emissions and human exposure that incorporates fate, transport, exposure and toxicity.
Linearity	The simplest of toxicological dose-response relationships in which a doubling of the original dose would be expected to results in a doubling of the response frequency, and a halving of the original dose would produce a halving of the response frequency and so on down the dose ladder to zero dose. Used primarily for carcinogen or mutagenic environmental contaminants.
Log(dose) probit distribution function	A dose-response model which assumes that each animal has its own threshold dose, below which no response occurs and above which a tumor [or other effect] is produced by exposure to a chemical.
Mantel-Bryan (MB) extrapolation model	The Mantel-Bryan extrapolation is a special case of the conventional log(dose):probit function that describes the dose-response relationship for threshold-acting agents in a population of exposed individuals. It is a means of predicting the probability of a incident health effect for any exposure level (dose) at or below the notional threshold for a given threshold-acting substance. As the notional threshold is typically close to the experimental ED05 level (the exposure level at which no more than 5% of the exposed population is affected), the Mantel-Bryan extrapolation is anchored on the observed ED05 level, with the corresponding slope of the log(dose):probit function assumed conservatively to be equal to one. The actual slope may assume values other than one. Assuming that sufficient dose-response data is available for a given substance, the actual slope may be used in place of the default slope of one. In the HEIDI package, the inclusion of a slope-modifying factor other than one would thus transform the level of analysis from subgroup 4c to subgroup 4d.
Mixing height	The expanse in which the air rises from the earth and mixes with the air above it until it meets air that is equal or warmer in temperature.

Non-threshold toxicity	A class of toxicity mechanisms where the damaging biological processes are thought to occur at any exposure level about zero dose, often in a linear dose-response relationship.
Physicochemical characteristics	Parameters such as atmospheric and non-atmospheric half-life and intake fraction used to estimate the atmospheric degradation rate of air toxics in Group 3 analyses.
Quantitative risk assessment	The use of science-based risk information and analytical methods to characterize the nature and extent of environmental health risks. Risk assessment employs techniques for measuring and estimating the likely health impacts, and other adverse results of releasing or discharging specified amounts of pollutants. Risk assessment normally includes the risk identification and risk estimation steps, and may in some risk frameworks also include the risk evaluation step.
Reference concentration	An estimate (with uncertainty spanning perhaps an order of magnitude) of a continuous inhalation exposure to the human population (including sensitive subgroups) that is likely to be without an appreciable risk of deleterious effects during a lifetime. It can be derived from a NOAEL, LOAEL, or benchmark concentration, with uncertainty factors generally applied to reflect limitations of the data used. Generally used in EPA's noncancer health assessments.
Reference dose (RfD)	An estimate (with uncertainty spanning perhaps an order of magnitude) of a daily oral exposure to the human population (including sensitive subgroups) that is likely to be without an appreciable risk of deleterious effects during a lifetime. It can be derived from a NOAEL, LOAEL, or benchmark dose, with uncertainty factors generally applied to reflect limitations of the data used. Generally used in EPA's noncancer health assessments.
Slope modifying factor	A numerical factor that modifies the default parameter value of the dose-response slope, either for the unit risk function for nonthreshold agents or for the log(dose):response function (Mantel-Bryan extrapolation) for threshold-acting agents. Whenever the parameter value for the Slope Modifying (SP) factor is set at 1 by default, it has no effect on the dose-response slope. When SP values greater than or less than 1 are introduced, this changes the dose-response slope to produce a steeper slope (narrower range of population responses) or a shallower slope (wider

	range of population responses). By definition, subgroup 4d is the analysis used when the SP has been modified to a value greater than or less than 1. Such SP values should be introduced only on the basis of reliable experimental data obtained from suitable dose-response studies.
Tolerable Daily Intake	The total daily intake of a substance occurring over a person's lifetime that should not cause appreciable risk to health on the basis of all known facts. It is usually expressed in milligrams of chemical per kilogram of body weight per day (mg/kg/day). The RfD is calculated in a manner analogous to the TDI.
Threshold toxicity	A class of toxicity mechanisms where disease occurs when underlying biological perturbations exceed a critical level of cell damage or physiological malfunction. Usually applied to environmental substances that are thought not to act by carcinogenic or mutagenic mechanisms.
Toxicity weights	Health effect benchmark under the REIS methodology based on calculation of the USEPA Reference Dose (RfC) value for threshold-acting substances, where TW = 1/RfC. Also used indirectly (for comparative purposes) to derive a corresponding Health Canada toxicity benchmark based on the Health Canada Tolerable Concentration (TC) value, where TW = 1/TC.
Uncertainty factors (replaces the older term Safety factors)	One of several, generally 10-fold factors, used in operationally deriving the RfD and RfC from experimental data. UFs are intended to account for (1) the variation in sensitivity among the members of the human population, i.e., interhuman or intraspecies variability; (2) the uncertainty in extrapolating animal data to humans, i.e., interspecies variability; (3) the uncertainty in extrapolating from data obtained in a study with less-than-lifetime exposure to lifetime exposure, i.e., extrapolating from subchronic to chronic exposure; (4) the uncertainty in extrapolating from a LOAEL rather than from a NOAEL; and (5) the uncertainty associated with extrapolation from animal data when the data base is incomplete.
Unit risk	A measure of the health risk association with a continuous daily exposure to a pre-defined dose of a toxic substance, usually a carcinogenic agent. For example, for a hypothetical carcinogen, the Unit Risk for continuous exposure to 1 milligram/kilogram body weight per day might result in a lifetime cancer risk of 5×10^{-5} (i.e. 5 in 10,000; 10 in 20,000).

Weight of evidence	Considerations involved in assessing the reliability of available information about hazard; and the quality of testing methods, the size and power of the study design, the consistency of results across studies, and the biological plausibility of exposure-response relationships and statistical associations.

Sources

McColl, S., Hicks, J., Craig, L., and Shortreed, J. 2000. Environmental Health Risk Management. A Primer for Canadians. NERAM Report No 4. http://www.neram.ca/Pages/research/primer.htm

Risk Assessment Information System. Glossary of Useful Terms Found In Risk Assessment, EMBAM, Health Physics, and Waste Management Reports. http://risk.lsd.ornl.gov/homepage/glossary.shtml#U

Royal Society of Canada. 2001. Report of an Expert Panel to Review the Socio-Economic Models and Related Components Supporting the Development of Canada-wide Standards for PM and Ozone. http://www.rsc.ca/english/

Society for Risk Analysis. Glossary of Risk Analysis terms. http://www.sra.org/glossary.htm#index

US EPA. Glossary of IRIS terms http://www.epa.gov/iris/gloss8.htm

EXECUTIVE SUMMARY

This NERAM report documents the conceptual and methodological approach to the development of a health effects-based priority ranking system for air emissions reductions from oil refineries in Canada. The *Health Effects Indicators Decision Index - Version 2* (HEIDI II) is a MS Excel spreadsheet screening-level tool designed to assist policymakers in prioritizing reductions of air emissions from Canadian petroleum refineries on the basis of estimated risk to human health. The HEIDI II model is an expansion of the HEIDI 1 prototype (Model 4c) previously developed by NERAM in 2002-03.

The tool produces facility-level rankings of the potential health impacts associated with three classes of air emissions: (1) carcinogenic air toxics (2) non-carcinogenic air toxics, and (3) criteria air contaminants (CACs) for each of the 20 refineries in Canada.

HEIDI II provides relative rankings of the estimated health impacts associated within the three classes of substances emitted from each facility based on predicted incidence of health effects, as well as using a summary measure of health impacts that allows for a comparative ranking of the incidence and severity of health effects across the three classes of air emissions, if desired by the user. As inputs to its calculations, HEIDI II considers the site-specific annual pollutant emission data, ambient air concentrations associated with these releases, concentration-response functions for various types of health effects, location-specific background air concentrations, site-specific population densities, and the baseline incidence of different health effects endpoints, such as cancer, non-cancer illnesses, and cardiorespiratory illnesses and death.

What substances are included and how were they selected?

HEIDI II considers selected air pollutants that are reported annually in Environment Canada's National Pollutant Release Inventory (NPRI) database. HEIDI II includes 29 air toxics including all polycyclic aromatic hydrocarbons (PAHs) as a mixture class and benzene, toluene, ethylene, and xylene (BTEX) substances as another mixture class. The air toxics were selected in consultation with NAICC-A's NFPRER Health Prioritization Subgroup based on the following criteria – quantity of emissions reported in NPRI, CEPA-toxic substances, substances included on Health Canada Priority Substance List (PSL2), and PSL scores for toxicity, persistence and bioaccumulation.

What outputs does HEIDI provide?

HEIDI provides the following three health impact ranking outputs for each facility:

1) Ranking of pollutants based on predicted number of annual cases of health effects. The predicted number of health effects is useful only for purposes of making risk-related comparisons between chemicals and do not represent actual risk. This ranking does not take into consideration differences between types of health effects i.e. temporary, chronic, and fatal conditions.
2) Ranking of pollutants based on simplified Disability Adjusted Life Years (DALYs) that provide a common measure for comparing the severity of different health endpoints (e.g. cancer, non-cancer illnesses, and cardiopulmonary illness and death) across the three classes of air emissions. The DALY calculation is based on years of life lost due to death and loss of quality of life due to illness. DALYS for each pollutant are shown as a percentage of the total DALYS within each category.
3) Ranking of pollutants based on more complex Disability Adjusted Life Years (DALYs) that consider type of cancer, type of systemic disease, or type of cardiopulmonary health effects.

How can the results be applied?

The purpose of HEIDI II is to provide a screening-level risk-based ranking of refinery NPRI emissions, to help inform users in prioritizing reductions in petroleum refinery emissions. There are considerable uncertainties in the data inputs and modeling assumptions within each of the three modules, and care is advised when comparing health impacts across chemical classes, particularly between cancer, non-cancer effects, and the criteria air contaminants. The rankings rely on rough statistical estimates of predicted incidence rates for a variety of health endpoints of widely differing severity. The statistical models used to calculate priority rankings can provide useful guidance in relative terms by comparing estimated health impacts associated with annual emissions at the facility level, but they cannot adequately represent absolute estimates of health risk in the exposed populations.

What data is used to provide the health impact rankings?

The HEIDI II tool is comprised of three modules:

(1) the **Air Exposure Model** uses a USEPA air dispersion computer model (AERMOD) to estimate ambient concentrations of carcinogenic and non-carcinogenic air toxics and particulate matter (PM) in the airshed impacted by each refinery. Refinery emissions data are from Environment Canada's NPRI database (2001) for the air toxics. The module uses 2001 criteria air contaminant emissions data provided by Environment Canada in 2003 for the HEIDI I project. This data was collected from CPPI member refineries and from publicly available information for non-CPPI refineries. HEIDI II also estimates in a simplified manner the formation of secondary particulate matter from PM precursors (NO_2 and SO_2) using conversion factors found in the research literature. The air pollutants are assumed to be emitted from a single stack in the centre of the refinery property. It is assumed that each substance is emitted at a standard stack height (30 m) at a constant rate over the period of one year. A generic meteorological profile representing southwestern Ontario is used as the default scenario.

(2) the **Health Effects Module** estimates, for each refinery location, the predicted cancer incidence, systemic disease incidence, and cardiopulmonary disease incidence associated with the refinery's contribution to the ambient air concentration of each substance. Health effects are estimated within 5 radial zones, each with 4 geographical quadrants, within a 25 km boundary. Physical air distribution patterns are generic and not site-specific.

This module uses Geographical Information System (GIS) software ArcMap to determine the exposed population at risk -- incorporating site specific population density profiles and generic Canadian age/sex distribution profiles derived from 2001 Statistics Canada Census Data. This module also considers Environment Canada data on background air levels of pollutants from anthropogenic and natural sources collected in the vicinity of each of the refineries, to estimate the facilities' attributable contribution to ambient air concentrations above background levels at each location.

For estimating population health effects of air toxics, HEIDI II uses concentration-response parameter values based on standardized measures of concentration-response derived primarily from Health Canada source materials, or where Health Canada values are not available, for USEPA or CalEPA sources. HEIDI II estimates chronic health effects associated with exposure to particulate matter (PM) based on the extensively peer-reviewed American Cancer Society and Harvard Six-City chronic epidemiology studies. The population health impacts associated with chronic exposure to PM are estimated to be as large as or greater than those from acute exposure. It is recognized however, that HEIDI II will likely underestimate the health effects associated with acute (daily) PM exposure to some extent.

(3) The **Health Impacts Module** aggregates diverse health effects of varying severity using a common metric. A series of simplified Disability Adjusted Life Years (DALYs) are calculated based on the approach developed by the International Life Sciences Institute (ILSI) which accounts for three basic levels of severity.

The more complex form of DALYs, based on the World Health Organization 'global burden of disease' approach, uses 140 illness categories representing fatal and non-fatal outcomes according to age, sex and other demographic factors. The final output of the HEIDI II package is a priority ranking of those NPRI substances deemed most suitable for emissions reduction, according to the predicted health effects case-incidence rates (which do not consider impact) or the predicted health impact DALY statistics (which attempt to take the impact of the health effect into account).

The HEIDI II priority ranking tool has successfully demonstrated that it is possible to develop a consistent and objective methodological approach for ranking priority reductions of air emissions within the oil refinery sector in Canada.

As HEIDI II is a fully functional prototype computer program, it can be used by decision-makers and other concerned parties to help inform the process whereby emissions reductions decisions are achieved. It can support decision-making with several user-configurable features that enable informed judgment about the interpretation of the ranking results -- these included program transparency, detailed descriptive information regarding health effects, alternate modes of output rankings within and across classes of substances, and sensitivity analysis of critical input parameters (stack heights, photodegradation time, imputed values for 'zero' reported emissions).

The NERAM project group therefore suggests that the HEIDI II should be considered for adoption by NFPRER and the Canadian Council of Ministers of the Environment (CCME) as one of the recommended decision tools to help inform the priority ranking of air emissions from oil refineries in Canada.

I. BACKGROUND

In 2002-2003 the Network for Environmental Risk Assessment and Management (NERAM) completed an evaluation of issues and approaches for human health risk-based prioritization with respect to their applicability to setting priorities for refinery emission reductions (McColl et al. 2003). The project was carried out for the NAICC-A National Framework for Petroleum Refinery Emissions (NFPRER) Health Prioritization Sub-group. The NAICC-A NFPRER is a multi-stakeholder initiative co-chaired by Environment Canada and Alberta Environment. It was originally proposed by Canadian Petroleum Products Institute (CPPI) in 2001, and was accepted by CCME's National Air Issues Coordinating Committee - Other Air Issues (NAICC-A), in 2002. The NFPRER will provide a Framework containing principles and methods for jurisdictions to establish performance-based facility emissions caps for criteria air pollutants and air toxics from the petroleum refining industry. The role of the NFPRER Health Prioritization Sub-group is to gather information on health implications of refinery emissions and make recommendations for prioritizing and phasing in emission reductions. To assist in this task, NERAM completed the proof of concept and prototype development of a risk analysis tool called the Health Effects Indicator Decision Index (HEIDI). HEIDI performs a spreadsheet analysis to determine priority rankings for air emission reduction within a refinery site using Environment Canada National Pollutant Release Inventory (NPRI) emission data and various toxicity, fate and exposure parameters. The tool was applied to six air toxics in three refinery locations across Canada and one hypothetical worst-case refinery. CCME contracted NERAM to undertake a further phase of development of the HEIDI software tool (HEIDI II) with the following objectives:

1. Extend the model to enable its application to Common Air Pollutants (SOx, NOx, VOC, $PM_{2.5}$, PM_{10}) and Air Toxics (including Benzene) in consultation with the HPSG (Health Prioritization Sub-group), or a sub-group of it, if appropriate.
2. Conduct a study of the feasibility of a common risk metric for air toxics (both carcinogenic and non-carcinogenic) and common air pollutants, including a critical analysis of the strengths and limitations of such an approach;
3. Expand the list of substances to be modeled to about 20-25 substances of the approximate 100 substances in the refinery emissions data for Canada. These would be selected on the basis of toxicity and quantity, with technical decision support from the technical sub-committee of the HPSG;
4. Apply HEIDI (and associated methods) to about 20-25 substances of the approximate 100 substances in the refinery emissions data for Canada, to incorporate: a full set of refineries in Canada; real world population distribution profiles for refineries; and inclusion of background concentrations as an illustrative example only (for a clean Canadian refinery location and a dirty air-shed).

The principal objective of the project was to develop a health effects-based priority ranking system for air emissions reductions from oil refineries in Canada, with the intention that such a tool could be applied as a potential resource for supporting decision-making, constituting one of several means of informing the decision process for air emissions reductions.

Phase-1 study

From October 2002 to May 2003, a preliminary assessment of comparative human health risk-based prioritization schemes for Canadian petroleum refinery emission reductions was undertaken by the Network for Environmental Risk Assessment and Management (NERAM) for the NAICC-A National Framework for Petroleum Refinery Emission Reductions (NFPRER) health prioritization sub-group. The overall objective of the project was to *"carry out an assessment of human health risk-based prioritization schemes that may be useful for priority-setting for refinery emission reductions"*. Based on a review of literature, 14 issues were identified for consideration in the evaluation of alternative priority screening methods. These issues addressed uncertainties in relating emissions (from various pathways) to human exposures; limited scientific understanding of short and long term health effects associated with acute and chronic exposures; limited monitoring data to characterize background concentrations, the need for approaches which consider threshold and non-threshold acting substances, air toxics and criteria air pollutants; and the data and resource requirements both to carry out and validate the prioritization approach.

The study approach was based primarily on the typology for priority setting approaches developed by Pennington and Bare (2001). Five levels of prioritization ranking schemes were identified according to comprehensiveness of model inputs and increasing complexity:

Analysis Group 1: ranking by total emissions mass only (direct data summation)
Analysis Group 2: ranking by emissions mass with toxicity weightings (effect normalization)
Analysis Group 3: ranking by emissions mass with toxicity weightings, and physicochemical characteristics (criteria-based score and ranking)
Analysis Group 4: ranking by emission mass, toxicity weighting, physicochemical characteristics, and exposed population (model-based approaches)
Analysis Group 5: model-based exposure assessment and quantitative dose-response assessment (full risk assessment)

According to conventional risk assessment frameworks, the first three levels of analysis should be considered a type of hazard assessment, since the input parameters focus exclusively on the inherent characteristics of the chemical agent and lack site-specific information such as exposed population distributions. The fourth and fifth levels represent quantitative risk assessment models because they

estimate the probable incidence of health effects in exposed human populations, based on a defined dose-response relationship.

Following the review of literature and assessment of existing prioritization methods it was determined that to carry out the study, it was necessary for NERAM to develop a prototype analysis tool (HEIDI – Health Effects Indicators Decision Index) outside of the contractual arrangements. HEIDI offered a number of capabilities not available in existing prioritization tools: i) the capability to incorporate physicochemical parameters, toxicological dose-response parameters, population density functions, and background air concentrations for a variety of air toxics; ii) the capability to assess priority setting methods of varying complexity and assess the sensitivity of various input parameters; iii) the capability to extend the analysis of emission reduction priorities to all Canadian refineries and all NPRI emissions, and; iv) the capability for model validation and groundtruthing.

The basic concepts of the Analysis Groups 1 to 4 (Group 5 cannot be modeled using generic ranking formulas) were operationalized using algebraic formulas within a prototype computer spreadsheet package called HEIDI for the purpose of comparing the priority rankings produced by each of the four Analysis Groups, including several sub-analyses. The spreadsheet package includes a series of standardized datasets that supply the ranking formulas with the required physicochemical and toxicological parameters for each air toxic substance, and the NPRI annual emission inventory data for various petroleum refineries in Canada.

To establish proof of concept by evaluating the relative strengths and weaknesses of each of the Analysis Groups six substances were selected for ranking analysis based on their importance and their representativeness as various classes of chemical toxicants (carcinogens, metals, VOCs etc). The six air toxics assessed were benzene, MTBE, mercury, n-hexane, toluene and ethylbenzene. The NPRI emissions datasets were assessed for the Shell Scotford Refinery in Fort Saskatchewan Alberta, the Chevron Refinery in Burnaby BC, and the Irving Oil refinery in Saint John NB, plus a hypothetical worst case refinery.

The analysis of the four prioritization approaches indicated that the preferred base model for determining the rank-order for prioritization of NPRI refinery emissions based on health effects was Pennington analysis group 4. This approach accounts for the expected background concentration for each substance, and relies on two types of dose-response formula which are both a continuous function of exposure (dose) applicable across any possible range of exposure concentrations – at all exposure levels for substances not exhibiting a threshold; and below, near, or above the threshold levels for threshold-acting agents. The use of continuous linear functions for both non-threshold and threshold-acting agents ensures that the estimated population incidence of health effects is founded on sound toxicological theory.

NERAM recommended a number of additional tasks and further methodological work to improve the capability of HEIDI to inform decision-makers on air pollutant emission reduction strategies for refineries in Canada.

The following next steps were proposed for a possible Phase 2 development of HEIDI:

1. **Undertake expansion of the existing (HEIDI 1) ranking system to include:**
    a. all relevant NPRI substances; and
    b. all NPRI refinery facilities
2. **Address fundamental methodological issues:**
    a. assessment of multiple exposures to agents with similar health endpoints (class assessment) initially using Toxic Equivalencies (TEQ) for PAHs mixtures and PCDD/PCDF mixtures; other types of class analysis possible later for various chemical classes (specified mixtures of aromatics, aliphatics, etc.)
    b. development of rank-ordering system for precursors of key secondary criteria pollutants (ozone, $PM_{2.5}$)—these include VOCs (various classes), NOx, SOx, primary $PM_{2.5}$
    c. constructing a common risk metric for air toxics and criteria air pollutants, using weighted health effects measures (e.g. PYLL, QALYs) or monetization techniques such as WTP or contingent valuation
    d. refinement of ED05 calibration based on primary dose-response data for threshold-acting substances, instead of indirect derivation from RfC or TC values
3. **Provide data-driven refinement of the ranking model:**
    a. systematic examination of relevant Canadian air modeling datasets to assess quality, reliability, and relevance to required inputs to various levels of the HEIDI model
    b. inclusion of real-world population distribution profiles for all refinery facilities based on StatsCan data and GIS analysis
    c. inclusion of site-specific background air concentrations of air toxics and criteria air pollutants
    d. inclusion of wind rose and seasonal climatic effects for prototype model
    e. consider routine inclusion of ISC air model for transport and fate modeling, if suitable data available—e.g. stack data, other site-specific emissions data, meteorological and topographical data
    f. inclusion of a slope-modifying factor (analysis subgroup 4d) for selected substances with data demonstrating a non-linear dose-response function (e.g. PB-PK analysis)

4. **Conduct calibration and validation studies of the ranking model**

    a. comparison of HEIDI model predictions with available quantitative risk estimation studies on refinery facilities (e.g. Saint John, U.S. risk assessments)

5. **Consider additional computational features for the ranking model:**

    a. provision for computation of the estimated risk reduction for a given emission reduction, by calculation of the ratio of [risk-before]/[risk-after] for any specified substance

    b. uncertainty analysis on key variables using Monte Carlo simulation techniques

Following completion of the HEIDI 1 prototype, CCME contracted NERAM for a Phase 2 study to further develop the capability of the software tool to provide relative rankings of 20-25 NPRI substances (including air toxics and criteria air contaminants) for all refineries in Canada using a common health impact metric. The development of HEIDI II incorporates the tasks identified in recommendations 1, 2 and 3a-d. Recommendations 4 and 5 were seen as outside the scope and timeframe of the Phase 2 study and may be considered at a later date.

II. HEIDI II DEVELOPMENT PROCESS

In October 2004, the NERAM team began the second phase of development of the air emissions priority-ranking system on behalf of NFPRER, with the development of the improved HEIDI II model serving as the means of organizing the conceptual and methodological refinement of the previous Phase-1 work. The principal objective was to produce a working HEIDI II model that would demonstrate the capacity of this type of ranking system to help inform decision-making for air emissions reductions from oil refineries in Canada. Throughout the inception and development of the framework ranking system, the NERAM project team reported to the Health Prioritization Sub-group of NFPRER, with ongoing liaison maintained by Environment Canada. The Health Prioritization Sub-group was constituted as a representative multistakeholder working group comprised of representatives from federal, provincial and local government agencies (e.g. Environment Canada, Health Canada, Alberta Environment), from the private sector petroleum refinery industry, and from non-governmental organizations (NGOs) with interests in the environmental and health aspects of air emissions.

It was agreed by all parties that development of the HEIDI II model should be accompanied by ongoing input and approval of the ranking methodology by the Health Prioritization Sub-group. This collaboration was important because it was deemed essential that the Sub-group be able to provide informed support and commentary for the completed HEIDI II model when the final product was to be forwarded to NFPRER for possible inclusion in the Framework. It was also held essential that the HEIDI II model should be transparent with respect to the methods of computation of potential health effects and health impacts, and that its conceptual foundations and methodological limitations be well documented in the final report and User Guide. Table Table 7 in Section V summarizes the underlying assumptions and limitations of the HEIDI model with respect to NPRI emissions, air modeling, health effects modeling and health impacts assessment.

Accordingly, throughout the Phase-2 development cycle of the HEIDI-II model, all significant conceptual and methodological issues were reviewed on a periodic basis by the entire Health Prioritization Sub-group, or by delegated members of the Health Prioritization Sub-group. Ongoing communications and exchange of materials between NERAM and the Health Prioritization Sub-group were maintained via frequent emails and periodic telephone conference calls coordinated through the offices of Environment Canada. All methodological decisions and assumptions in HEIDI II were documented in NERAM and NFPRER Sub-group minutes. Documentation of decisions on key parameters and assumptions in HEIDI II is provided in Table 9 of Appendix A.

Whenever a predictive risk-based model is intended to inform multi-stakeholder decision-making in environmental health, it is important that the decision tool be made available for examination and critical appraisal by scientists, stakeholders, and interested members of the general public. Therefore, after a limited period of internal review and comment by members of the NFPRER stakeholder groups, the

NERAM organization will make the HEIDI-II model and its attendant documentation publicly available for the use of any person or organization.

It is recommended that CCME or NFPRER should consider the potential implications of unsupervised public use of the HEIDI II tool for addressing environmental health issues, and should consider developing means by which the tool can be better used and understood by third parties within or outside of the NFPRER framework.

III. METHODS DEVELOPMENT

The HEIDI II prototype model was structured conceptually according to the three main analytical components required to produce a health effects based priority ranking system. The HEIDI tool is comprised of three modules:

(1) the **Air Exposure Module**;

(2) the **Health Prioritization Module**; and

(3) the **Health Impacts Module**.

The analytical structure of each module and the interrelationship of the three modules within the HEIDI II model is shown in detail in Figure 1.

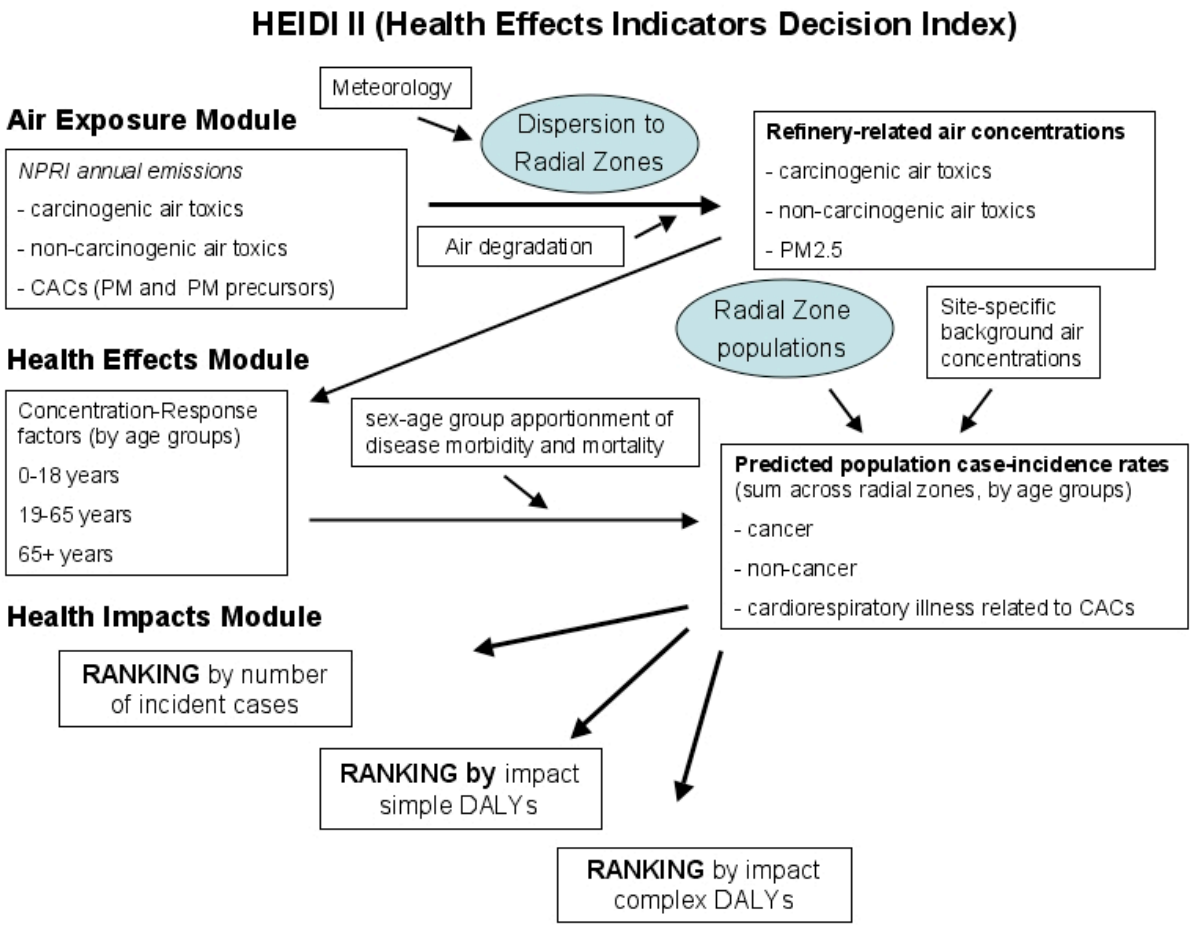
1. Air Exposure Module

The Air Exposure Module uses AERMOD, a recent derivative of the US EPA air dispersion computer model (ISCST3), to estimate ambient concentrations of carcinogen and non-carcinogen air toxics and PM in the airshed impacted by each refinery.

For the air toxics, refinery emissions data are from Environment Canada's NPRI database (2001). The module uses 2001 criteria air contaminant emissions data provided by Environment Canada in 2003 for the HEIDI I project. This data was collected from CPPI member refineries and from publicly available information for non-CPPI refineries.

HEIDI also estimates in a simplified manner the formation of secondary particulate matter from PM precursors (NO_2 and SO_2) using conversion factors found in the research literature. As a simplifying assumption, the air pollutants are assumed to be emitted from a single stack in the centre of the refinery property. It is assumed that each substance is emitted at a default stack height (30 m) at a constant daily rate over the period of one year. A generic meteorological profile representing southwestern Ontario is used as the default scenario.

Figure 1 Major Components of the HEIDI II Model



A. Air Dispersion Modeling

HEIDI II incorporates air dispersion modeling to provide an estimate of the downwind contaminant concentrations for each refinery location in Canada. The HEIDI II model employs the concentration estimates at each Canadian site to estimate health effects in the exposed population.

Dispersion Model: HEIDI II uses the US EPA regulatory model AERMOD to estimate the concentrations of emitted substances in the atmosphere surrounding the refinery locations. The model has a wide application for modeling point, area, and volume sources and is applicable to most refinery locations in Canada if appropriate meteorological and terrain conditions are provided.

Emission Inventory: Environment Canada's National Pollutant Release Inventory (NPRI) was used to obtain refinery-specific annual air pollutant emissions data. Emission inventories are estimated losses of a material from a facility. The estimate for each pollutant is a composite number of point source (i.e. stack), area, and volume emissions from a variety of emission sources within a refinery location, or losses through spills evaporation and escape from containment. The emissions may occur periodically or continuously over the 12 month period. It is important to note that the emission inventory is largely based on model estimates of losses, and not on measured or controlled experimentation of actual facility emission rates. There are considerable differences in the emissions from site to site in Canada, and it is likely that many of the differences are attributable to variation in procedures for calculating and estimating material losses.

If a substance is emitted uniformly and continuously over the annual period as recorded in the inventory, it can be modeled using a dispersion program with representative conditions. However, if the release is periodic, or perhaps as accidental releases of short duration but high concentration, the air modeling will not describe the maximal concentration or location with the same degree of accuracy. In fact, for these types of releases that may happen only a few times per year, the AERMOD dispersion model may not reflect exposure conditions adequately, and other models are recommended if the release conditions are known. In the development of the HEIDI II model there has been an explicit assumption that the emissions are uniform and continuous with time.

Terrain Conditions: HEIDI II assumes that the refineries are located in an area of flat terrain. This assumption is considered generally applicable to all Canadian refineries with the exception of the British Columbia locations (Chevron-Burnaby refinery and Husky Oil Prince George refinery). Despite this, it is assumed that the primary population and terrain impacts are within a generally regular geographical terrain within a few kilometres of these facilities, and that the rough terrain occurs at the outer reaches of the exposed area.

Meteorology: HEIDI II uses one generic location to represent the meteorological conditions that exist at all refinery locations. This assumption was agreed to by the NFPRER health prioritization Sub-group in order to meet the terms of reference for the study within the project scope and budget. HEIDI II uses four years of meteorological data (1996-2000) for the southwestern Ontario region. This appeared to represent the most number of Canadian refineries in any one regional area. This included four of the 20 locations, including Sarnia (three sites) and Nanticoke locations. Upper atmospheric soundings for the same representative region was employed in the AERMET meteorological preprocessor module. A wind rose plot is provided in
Figure 1.

One meteorological profile cannot represent conditions at all locations. Differences occur in several parameters such as wind direction, wind speed, temperatures, cloud coverage, and precipitation. A closer

approximation of individual refinery contributions to local airsheds requires site-specific modeling. However, HEIDI II is designed to assess the applicability of the HEIDI model on a broad national basis, and one "generic" meteorology is used.

The dispersion of emitted pollutants was estimated for each of the Canadian refineries. HEIDI II includes all of the substances that are reported in the 2001 National Pollutant Release Inventory database for each refinery.

Emission Site Conditions: HEIDI II assumes that emissions are released from one stack in the centrepoint of the refinery. The original stack height was assumed to be 30 metres in height. A sensitivity analysis was performed to assess stack height at 15 metres and 5 metres to estimate differences in the concentration pattern around the refinery site. The results are provided in Tables 3a-3c of Appendix B.

Output Conditions: The emissions are assumed to originate from a point source location in the centre of the refinery property. It is assumed that each material is emitted at a constant rate over the period of one year on a continuous basis. Table 1 summarizes the emission conditions used in the dispersion model.

Table 1 Summary of Modeling Conditions

Parameter	Modeled Estimate
Stack Height	30 metres from the ground sensitivity at 15 and 5 metres
Stack Exit Velocity	1.0 metre/second
Stack Diameter	0.2 metres
Emission Temperature	323 ^{o}K
Terrain Type	Flat, Urban
Emission Rate	1 gram/second (note: scaled to the real emission for each material)
Loss/Removal	Depends on material emitted. In draft model pollutants are conserved

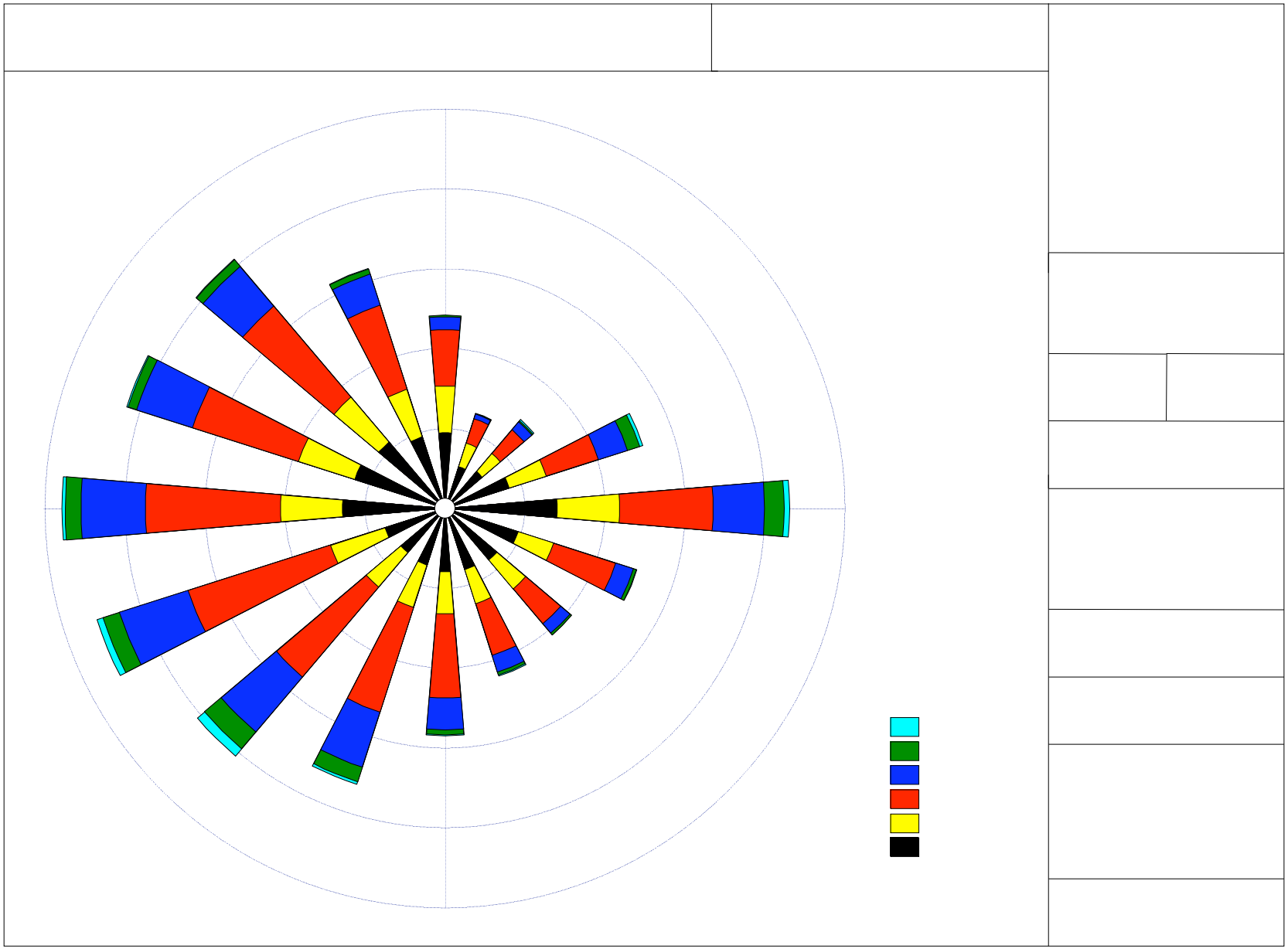


Figure 1 Wind rose describing meteorological data used in HEIDI II

Receptor Coordinate System: The receptor locations were identified using a polar coordinate system. Concentric rings were located around the refinery at radii of 1.0, 2.5, 5, 10, and 25 kilometres. Each concentric ring was divided into four quadrants associated with the four directions of northeast, southeast, southwest, and northwest from the refinery centre, resulting in a total of 20 sectors. The population residing in each of the 20 sectors was determined using 2001 Census Canada data and ArcMap software.

The concentrations of each primary pollutant emission are given as an annual average for each of the 20 sectors. This is a direct function of the estimations of the AERMOD dispersion model. Other secondary pollutants are estimated outside of the model and described in the following section.

Averaging Time: Canada-wide standards for PM and ozone are expressed with specific "averaging times". An air quality standard that is designed to protect individuals over the exposure period of one day (i.e. 24 hours) may also have a second standard that is designed to protect over short-term periods of one hour. The dispersion model in HEIDI II was set to provide data for annual average pollutant conditions.

Model Estimates of Dispersion: The AERMOD model of each facility provides estimates of the ground-level atmospheric concentration of an emission of 1 gram/ second. All of the emission inventory substances are emitted at levels that are different from 1 gram/second. Therefore the dispersion estimates are used as a scaling factor to estimate the concentration of the inventory pollutants at specific distances from the refinery centre. The annual average scaling factors are given in the Table 3 below for each radial exposure area.

Table 2 Summary of "Probable Worst Case" Conditions Used in this Assessment

Variable	Condition
Emission Duration	not worst case emission – emission is uniform over 12 months
Averaging Time	representative but not worst-case – annual averages are estimated
Meteorology	representative of Southwestern Ontario
Removal Processes	No dry, wet, or photo-oxidation processes included in modeling set, photo-decomposition and secondary PM estimated separately from dispersion model

Table 3 Dispersion Modeling for a 30 Metre Stack Emission Annual Averaging Factor
(to convert from source emission in g/s to receptor)

Direction Degrees	500 metres	1,000 metres	2,500 metres	5,000 metres	10,000 metres	25,000 metres
45	0.688	0.334	0.104	0.048	0.029	0.013
135	0.692	0.270	0.083	0.045	0.033	0.016
225	0.219	0.102	0.037	0.020	0.015	0.007
315	0.349	0.149	0.058	0.031	0.023	0.013

B. Modeling of Secondary Pollutants

The atmospheric concentrations of primary pollutants that are emitted from the refineries are estimated using the AERMOD model. It is assumed that these primary pollutants undergo dispersion to areas of lower concentration. Many of these pollutants undergo losses in the air due to precipitation scavenging, photolysis, and chemical reactions.

Several primary pollutants such as sulphur dioxide, nitrogen oxides, and volatile organic compounds may react in the atmosphere to form secondary pollutants, such as particulate matter and photochemical smog. The same loss mechanisms that reduce the primary pollutant concentrations in the atmosphere may fuel the generation of secondary pollutants. The estimation of these secondary pollutants is necessary as part of the health effects module.

The physical and chemical conditions that allow the formation of secondary pollutants may occur on one day but not the next. Consequently, our estimation procedure must be able to estimate the potential health impacts if primary pollutants are conserved, or if they are consumed to form secondary pollutants. This dichotomy may lead to "double counting" the effects of a species as a primary and a secondary pollution constituent, when in fact it may depend on the weather conditions in that location and at that time of year, month, or day.

Secondary Particulate Estimation: In the first NAICC-A study (HEIDI 1) secondary particulate was estimated from the consumption of sulphur dioxide and nitrogen dioxide to sulphate and nitrate anions. These anions were assumed to be combined with atmospheric ammonia to form ammonium sulphate and ammonium nitrate, respectively. These materials are classified as secondary solid particulate species in the $PM_{2.5}$ size regime.

The formation of ammonium sulphate and ammonium nitrate operates under the assumption that ammonia is available in the atmosphere to allow this formation to occur. There are natural sources of ammonia that provide this reactant to the atmosphere. The amount of available ammonia is largely dependent on the terrain and climate conditions of the region. Although in

reality the amount of ammonia may be a limiting factor in the formation of ammonium salts, this study assumes that it is in ample amount in the atmosphere in general.

We believe this is a suitable approximation for HEIDI II as it is consistent with most summertime conditions in the eastern portions of the country.

Modeling of the Formation of Secondary Particulate Matter

Reaction rates: SO_2 is assumed to convert to $(NH_4)_2SO_4$ at a rate of 5% available per hour, NOx is assumed to convert to NH_4NO_3 at a similar rate. These values were used in HEIDI II and are representative of conversions in atmospheric conditions where other pollutants exist. In pristine conditions, the rate of conversion is normally assumed to be 4% per hour.

Chemical and Physical Conditions: The converted gases change with the addition of oxygen and ammonia and both of these reactants are found in the lower tropospheric atmosphere. In this assessment it has been assumed that all refinery airshed environments have sufficient ammonia in the atmosphere for this reaction to occur fully.

In addition, it is important for appropriate physical conditions to exist at the time of the conversion reactions. This includes sunlight of sufficient spectral intensity to generate the formation of chemical radicals such as OH, and the presence of reactive surfaces such as other particles and droplets in the atmosphere. However, it is likely that these chemical and physical atmospheric conditions exist only some of the time when maximum PM conversion is able to occur. Ammonia is more likely to be available in warmer seasonal conditions, and sunlight is available in daytime hours with varying intensity.

Concurrent Dispersion Processes: At the same time as this conversion is occurring, the plume is dispersing from areas of high concentrations to regions of low concentration. There are two processes being considered:

- a particulate formation process going from low to higher masses of PM in the hours after precursor gases are emitted, and

- a dispersion of gases and particles to greater volume resulting in lower concentrations of PM in the air.

The relative effect on the regional PM concentration is estimated from these two process calculations.

Total $PM_{2.5}$ is the sum of primary $PM_{2.5}$ from the original source, and secondary $PM_{2.5}$ formed from SO_2 and NOx in the hours after emission. Each component is calculated separately at

regular time periods after the initial emission from the refinery. The three components, primary, sulphate, and nitrate particulate matter are summed as a total concentration at each time interval.

The potential formation of $PM_{2.5}$ for each facility is estimated and listed in Table 10 of Appendix B. The estimates of particulate concentration (after dispersal) are given as an annual average for each refinery location in this appendix table.

2. Health Effects Module

The Health Effects Module estimates, for each refinery location, cancer incidence and mortality, systemic disease incidence and mortality, irritation, and cardiopulmonary disease incidence and mortality associated with the refinery's contribution to the ambient air concentration of each substance.

Health effects are estimated within 5 discrete radial zones in a 25 km boundary surrounding each facility by dividing each radial zone into four 90 degree quadrants that are oriented NE, SE, NW, SW. Thus there is a total are 20 geographical segments whose exposed population size and estimated level of exposure to each pollutant must be calculated.

Exposed Populations:

This module uses Geographical Information System (GIS) software ArcMap to determine the exposed population at risk in each of the 20 geographical segments -- incorporating population density profiles, Statistics Canada Census Data, baseline mortality and morbidity data from Statistics Canada and the Canadian Cancer Society. For each of the 20 geographical segments, the size of the exposed population was calculated by converting census tract population counts into the corresponding geographical segments around each refinery.

Background Air Pollutants:

This module also considers Environment Canada data on background air levels of pollutants from anthropogenic and natural sources collected in the vicinity of each of the refineries, to estimate the facilities' attributable contribution to ambient air concentrations above background levels at each location. This data was used to calculate the predicted number of incident case of health effects attributable to background pollutant levels (unrelated to refinery emissions) for each of the noncarcinogenic air toxics. The net health effect of refinery-related air emissions was then calculated as:

Incident cases (refinery)
= Incident cases (background+refinery) - incident cases (background)

As the carcinogenic air toxics and CACs are thought to produce health effects independent of background concentration, these classes of pollutants were not subject to background discounting.

Concentration-Response (C-R) Functions:

To estimate population health effects of air toxics, HEIDI uses concentration-response parameter values based on standardized measures of concentration-response derived from Health Canada source materials, as well as a set of alternative values obtained from USEPA or CalEPA sources where Health Canada values were not available.

a) epidemiological C-R functions for air toxics - carcinogens

For those air toxics classified by Health Canada as 'human carcinogens' or as 'reasonably anticipated to be carcinogenic in humans', it was assumed that the concentration-response relationship would follow a linear non-threshold function. Air toxics in lower categories of carcinogenic evidence were assessed as 'air toxics - noncarcinogens'. Therefore, the predicted case incidence of cancer health effects (including both solid tumours and leukemia) in the exposed population was calculated in HEIDI II according to the following equation:

Incident Cases = Inhalation Unit Risk * Conc * exposed population

The inhalation unit risk was derived from the tumorigenic dose producing a 5% response rate (TD05) in the most sensitive test species according to Health Canada criteria.

b) C-R functions for air toxics - noncarcinogens

For air toxics - noncarcinogens, the concentration-response function was assumed to follow a threshold-acting behaviour. In the HEIDI 1 prototype, the nonlinear function selected to model the threshold-like behaviour was the Mantel-Bryan function, which is a special case of the classical $\log_e$(dose)-probit C-R function, where the response curve is anchored at the toxicological ED05 value and the default slope is set equal to one.

However, in early HEIDI II development, it was found that the C-R function for a Mantel-Bryan function with a slope = 1 was too shallow, so that it tends toward asymptotic zero response very slowly at lower doses below the ED05 (i.e. it is very conservative at low doses), while it tends to underestimate the responses at high doses above the ED05. After trying out a modified slope of 2 (which was too steep), the NERAM team settled on a slope of 1.5 for the revised $\log_e$(dose)-probit function. This 'steep' Mantel-Bryan function produces very good dose-response behavior across the entire range of possible air concentration values -- at low doses it asymptotes to zero reasonably quickly (producing a threshold-like response), while at doses above the ED05 it closely parallels the Unit Risk function for carcinogens. The latter feature is important, to help ensure a 'level playing field' in the dose-response functions for carcinogens (Inhalation Unit Risk)

and for threshold-acting air toxics ($\log_e$(dose)-probit) in the air concentration range above the ED05.

In addition, we found that the value for the Tolerable Concentration (TC, RfC) corresponding to the 'de minimis' risk level (10-5) used by Health Canada and USEPA is virtually identical using the 'steep' Mantel-Bryan function (slope=1.5) as it would be when calculated by the conventional threshold formula TC = ED05/100. This means that the conventional 100x uncertainty factors are automatically 'built in' to the steep Mantel-Bryan function without needing to be artificially added by the risk assessors. This is a very neat convergence of two entirely different methods for dealing with the dose-response characteristics of threshold-acting agents, and it could give extra "value-added" to the HEIDI II project as a whole.

c) C-R functions for Common Air Contaminants (CACs)

For the CACs, HEIDI estimates chronic health effects associated with exposure to $PM_{2.5}$ based on the extensively peer-reviewed American Cancer Society and Harvard Six City chronic epidemiology studies. The population health impacts associated with chronic exposure to $PM_{2.5}$ are estimated to be as large as or greater than those from acute exposure. It is recognized however, that HEIDI will likely underestimate the health effects associated with acute (daily) $PM_{2.5}$ exposure to some extent.

The epidemiological concentration-response (C-R) functions are linear non-threshold in form, so that the C-R function is based on a linear risk coefficent that incorporates a slope factor according to the following equation:

Incident Cases = Risk Coefficient * Conc PM2.5 * at-risk population

(both population prevalence and C-R slope factors are accounted for in the risk coefficient)

The reference documentation for PM health effects endpoints and concentration-response function information was obtained from the report by Abt Associates (2002):

Abt Associates. Nov. 2002. Particulate-Related Health Impacts of Emissions in 2001 From 41 Major US Power Plants. Prepared for the Environmental Integrity Project. Rockefeller Family Fund. http://www.abtassociates.com/reports/Abt_41_power_plant_report_Nov19.pdf

The HEIDI II model estimates annual case incidence rates for CAC-related mortality and morbidity endpoints based on Table 4 below of the Abt (2002) document. It is assumed that $PM_{2.5}$ is an overall indicator of ambient air quality and the concentration response functions for the various PM related health endpoints include effects associated with exposure to gaseous pollutants. This approach recognizes the high correlation between PM and gaseous co-pollutants and will avoid double counting of health effects. Therefore the primary emissions inventories of

gaseous CACs from oil refineries (SOx, NOx, VOCs) are exclusively used to obtain an estimate the concentration of secondary PM (using air modeling and chemical conversion factors). When added together with the primary PM inventory from oil refineries, the combined PM (primary and secondary) concentrations serve as the basis of the concentration-response functions used to provide risk estimates of PM-related air pollution for various chronic health endpoints attributable to the CACs. No gaseous co-pollutants were included in the risk estimates.

Table 4 PM-related health endpoints to be considered in HEIDI II

Health Effect	Population	PM Measure	Study
Mortality			
Associated with long-term exposure	Ages 30+	$PM_{2.5}$	Pope et al. (2002) Krewski et al. (2000)
Chronic Illness			
Chronic Bronchitis	Ages 30+	PM_{10}	Schwartz (1993)
Hospital Admissions			
COPD (ICD-9 codes 4490-492, 494-496)	Age 65+	PM_{10}	Samet et al. (2000)[a]
Pneumonia (ICD-9 codes 480-487)	Age 65+	PM_{10}	Samet et al. (2000)[a]
Cardiovascular (ICD-9 codes 390-429)	Age 65+	PM_{10}	Samet et al. (2000)[a]
Asthma (ICD code 493)	< 65	$PM_{2.5}$	Sheppard et al. (1999)
Asthma-related ER visits	< 65	PM_{10}	Schwartz et al. (1993)
Respiratory Symptoms/Illnesses Not Requiring Hospitalization			
Acute bronchitis	Ages 8-12	$PM_{2.5}$	Dockery et al. (1989)
Lower respiratory symptoms (LRS)	Ages 7-14	$PM_{2.5}$	Schwartz et al. (1994)
Upper respiratory symptoms (URS)	Asthmatics, ages 9-11	PM_{10}	Pope et al. (1991)

Note: this table has been adapted from Abt (2002) Exhibit 2-1.

Of the possible health effects endpoints identified in Abt (2002) Table 4, only four of the most important have been included for analysis in the HEIDI II model (see Table 5).

Table 5 Health Endpoints for $PM_{2.5}$ in HEIDI II

Health Endpoint	**Age**
Mortality Premature deaths associated with long term exposure to $PM_{2.5}$	30+
Morbidity (chronic) Chronic bronchitis associated with chronic PM_{10} exposure	30+
Morbidity (acute) Asthma attacks associated with chronic exposure to resulting in (1) hospital emergency room (ER) visits and (2) hospital admissions	<65

The remaining less severe and non-hospital end-points were not used as these are seen as less reliable health effects indicators compared to death, chronic illness, and hospitalization.

The HEIDI model computes the predicted number of incident cases of CAC-related health effects based on the size of the exposed population, the concentration-response (C-R) function for each endpoint, and the population prevalence (for acute endpoints, annual incidence) of the health condition. For C-R functions based on PM_{10} exposure , the C-R function was adjusted on the assumption that $PM_{2.5}$ is the actual toxic agent and that it comprised 50% of the PM_{10} fraction by weight. Age-dependence of $PM_{2.5}$ outcomes is accounted for in the risk coefficient applied to each age group. $PM_{2.5}$ outcomes are not dependent on gender. Age discounting was applied to adjust the actual population size of specific at-risk age groups with each health endpoint (e.g. 30+ for premature mortality).

3. Health Impacts Module

A considerable amount of basic theoretical work was required to characterize the means of achieving adequate summary measures of health effects and health impacts. This included an extensive review of methodologies related to the quantification of health impacts using a uniform metric for assessing mortality (deaths) and morbidity (illness). An extensive analysis of this issue is presented in a background paper on *Comparisons of Health Impacts for Different Classes of Air Emissions* in Appendix C.

The Health Impacts Module aggregates diverse health effects of varying severity using a common metric. A series of simplified Disability Adjusted Life Years (DALYs) are calculated based on the approach developed by the International Life Sciences Institute (ILSI) (Burke et al., 1996) which accounts for three basic levels of severity: 1) irreversible/life shortening 2) may be reversible, could be life shortening and 3) generally reversible, generally not life shortening.

The more complex form of DALYs, based on the World Health Organization 'global burden of disease' approach (WHO, 2000), uses 140 illness categories representing fatal and non-fatal outcomes according to age, sex and other demographic factors. The final output of the HEIDI package is a priority ranking of those NPRI substances deemed most suitable for emissions reduction, according to the predicted health effects case-incidence rates (which do not consider severity) or the predicted health impact DALY statistics (which attempt to take age of onset and severity of the health effect into account).

Definition of DALY

Disability-Adjusted Life Years, was selected by WHO (Murray, 1996) as the preferred measure of health impacts that combine mortality and morbidity. The measure is related to health only (intent is to rule out effects of type of risk, wealth, etc. that modify the individual utility) and is

intended to be a measure of society at large for purposes of public policy. The measure for society is constructed by the aggregation of individual measures of utilities of health status.

$$DALY = YLL + YLD \quad (1)$$

Where DALY is the Disability-Adjusted Life Years
YLL is the Years of life lost due to premature mortality
YLD is the Years Lived with Disability (morbidity)

$$YLD = DW \times L \quad (2)$$

Where YLD is the years lived with a disability
DW is the disability weight
L is the average duration of disability (years)

DALYs are a time measure of a health gap, the time lived with less than perfect health and the time lost due to death before a standard life expectancy (life expectancy at birth of 80 years for men and 82 years for women) (Pruss-Ustun, 2003). In many cases rather than the standard life expectancy the life expectancy for the country considered is used, e.g. in Canada the age specific life expectancy is use.

Equations (1) and (2) are for an individual case. Usually DALYs are estimated for a total population, for an exposed population, and so forth. The equations are modified by inserting the number of cases. For a total population estimate, such as for the WHO burden of disease study, the total population is used and the equations are modified from cases (incidence) to prevalence in order to estimate DALYs.

DW is a weight that reflects the severity of the disease on a scale from 0 (equivalent to perfect health) to 1 (equivalent to dead). Depending on how these weights are determined they are called disability weights, QUALY weights, health state valuations, health state preferences or health state utilities. (Pruss-Ustun, 2003). DW do not represent the lived experience of any disability or health state, or imply any societal value for the person in a disability or health state, but rather quantify societal preferences for health states in relation to the societal ideal of good health. (Pruss-Ustun, 2003) As a preference it means that society would be indifferent between any l person in the population living three years with a DW of .33 and any other person in the population dying one year prematurely, since the DALYs are the same.

DALY weightings - Simplified Approach for Air Toxics (SETAC)

The simplified approach of SETAC for toxic effects are generally defined for a variety of end points which may not necessarily correspond to health effects end points for CACs and carcinogens. SETAC (Owen, 2002) faced this same difficulty and have the same objective as our

study (i.e. characterizing chronic non-cancer toxicity with a view to establishing a screening indicator for organizing and aggregating information in order to provide meaningful direction for further policy analysis). Their solution as documented by Owens (2002) can be extended to provide a comparable estimate. It is noted that the SETAC approach case study covers many of the emissions considered in HEIDI II but there will be a need to provide estimates for missing data. The procedure proposed by SETAC is:

1. The procedure is a subjective scoring exercise, not a scientific or technical operation, but it does use the original toxicity data in an attempt to avoid hidden weighting and valuation schemes (e.g. ADI and RfD are not used) as a substitute for scientific characterization. When available, the toxicological ED05 or ED10 levels will be used to estimate toxicity. The method was developed for the case of an interrelated industrial system with environmental emissions as a focus and an objective of "identify and prioritize potentially important emissions and to facilitate risk assessment, including comparative risk assessment".

2. There were three classes of severity established: 1) irreversible/life shortening, 2) may be reversible, could be life shortening, and 3) generally reversible, generally not life shortening. The WHO uses general estimates of DALYs for these three broad classes of endpoints -- 6.7, .67, and .067 respectively (Pennington, 2002). For example, category 1 includes cancers and the DALY weighting of 6.7 is the average in Table 1 for all cancers considered by WHO.

IV. TECHNICAL FEATURES OF THE HEIDI II MODEL

1. Overview

HEIDI II is an Excel-based program that consists of a single Excel workbook comprising a series of worksheets that contain the data and algebraic formulas required to compute predicted health effects and health impacts for various NPRI air emissions, and their relative ranking for emissions reductions. The workbook allows the user to select several input parameters such as stack height and number of daylight hours in order to rank emissions from a particular refinery in Canada for health impacts. As the major purpose of each worksheet is discussed, it will be provided as a quoted heading highlighted in blue, i.e., "SCENARIO selection".

2. Structure of the HEIDI II model

Input

"SCENARIO selection"
To begin, the user clicks on the "SCENARIO selection" tab which is found at the far lower left of the screen. On this page, the user can select the refinery of interest from the dropdown menu.

The default stack height is set at 30 metres, but HEIDI can also perform rankings using stack heights of 15 metres or 5 metres. To change the default setting, enter the desired value (5 or 30) in the box provided. Entering values other than 5, 15, or 30 will result in a warning appearing under the box telling the user that s/he has entered an invalid number. Invalid numbers for stack heights will not produce any output.

The default setting for photodegradation time (important for predicting the amount of decay that a given chemical will undergo) is 12 hours. HEIDI II can also perform rankings using photodegradation times of 8 hours or 16 hours (for Class I and Class II air toxics only). Prediction of formation of secondary $PM_{2.5}$ is (currently) always based on 12 hours of sunlight. Altering the photodegradation time on the "SCENARIO selection" sheet will not alter predictions for $PM_{2.5}$.To change the default setting, enter the desired value (8 or 16) in the box provided. Entering values other than 8, 12, or 16 will result in a warning appearing under the box telling users that they have entered an invalid number. Invalid numbers for photodegradation times will not produce any output. Prediction of formation of secondary $PM_{2.5}$ is (currently) always based on 12 hours of sunlight. Altering the photodegradation time on the "SCENARIO selection" sheet will not alter predictions for $PM_{2.5}$.

HEIDI II normally ranks emissions based on what is reported by the NPRI for a given refinery in Canada. Because many of the emissions are reported in the NPRI as being zero, when in fact they may be under the NPRI reporting threshold for manufacturing, processing, or otherwise using a

chemical, an alternate ranking scenario is available within the HEIDI II model that permits imputed or hypothetical emissions scenarios to be specified by the user.

The user may choose a percentage value (1-99%) of the NPRI reporting threshold for substances that are recorded in the NPRI as zero emissions. The default setting for percent of the reporting threshold is 50%; the user may select any value between 1 and 99 by entering it in the box provided. Predictions based on this alternate ranking scenario are provided only on the "Health Impacts" sheet (see below).

Each time the user returns to the "SCENARIO" selection sheet, from another part of the excel workbook, the selected refinery will be cleared from the dropdown box so that a new selection can be made. As a result, it is important to ensure that the correct refinery is selected before leaving this sheet.

All toggles currently available to the HEIDI II user are located at the front end of the program on the "SCENARIO selection" worksheet. Toggles available in HEIDI are summarized in Table 6.

Table 6 Summary of Toggles currently available in HEIDI II

Toggle	**Description**	**Possible Values**
Refinery	A dropdown box which allows the user to select emissions from one of the 20 Canadian Refineries as the basis for the ranking	Twenty refineries are listed in the dropdown box, as well as a "hypothetical worst-case" refinery which represents the highest emissions recorded by any refinery for each chemical
Stack Height	The stack height can be altered to reflect emissions that occur primarily at ground level (represented by 5 metres) as opposed to those which occur further from the ground (15 or 30 metres). The height at which chemicals are emitted affects exposure at different locations around the refinery.	5, 15, 30 metres (default is set at 30 metres)
Photodegradation Time	Many substances degrade over time when exposed to sunlight. The photodegradation time represents the number of sunlight hours per day that will be assumed by the program. ****Note** that for $PM_{2.5}$, the photodegradation time is always assumed to be 12 hours.	8, 12, 16 hours (default is set at 12 hours)
% of Reporting Threshold	The NPRI reports zero emissions for many substances by many refineries. This means that emissions fall below the reporting threshold, but does not guarantee zero emissions. The user can create an additional health impacts ranking using a percent of the reporting threshold to stand in for values of zero reported by the selected refinery.	1-99 % (default is set at 50%)

Output Sheets

"Health Impact"
The output can be viewed on the worksheet called "Health Impacts". This is a comprehensive output and summary sheet which provides rankings for the emissions of the selected refinery based on disease incidence in the population, simple DALY values, and complex DALY values. Additionally, the "Health Impacts" sheet encapsulates most of the information used in producing the rankings. The rankings can be found at the far right of the worksheet. This sheet is the only place where predictions from the alternate ranking scenario (i.e., using a percentage of the NPRI reporting threshold for substances that are reported as zero emissions in the NPRI) are available.

This sheet should be viewed by all users to ensure that the sources of information used are adequately understood. Please see the section in this document entitled "Understanding the Output of HEIDI II" for more information about what the output means.

"Health Impact(print)"
A second sheet called "Health Impact(print)" provides a condensed version of the Health Impacts output sheet. This sheet should print in a readily readable form and provides information about each emission, the health endpoints used to rank its emissions, predicted incidence for these endpoints, DALY values, and rankings based on either incidence or on simple or complex DALYs. Rankings on this sheet are based on the primary emissions scenario: emissions as reported by the NPRI.

Data sheets

Several sheets in the HEIDI II workbook exist to provide source data for HEIDI II. These include:

"data_NPRI emissions": This sheet provides data as found in the NPRI database for emissions of each of the listed Class I and Class II chemicals in metric tonnes/year.

"data_Background": This sheet provides background concentrations for each of the chemicals being ranked at the location of each refinery. In this context, the term "background" refers to all ambient concentration of each substance other than that derived from refinery emissions. These data were collected mostly from the NAPS (National Ambient Pollution Surveillance) network in Canada. Data for some substances and locations were sparse, and for these cases, values were inferred using data from similar sites. For more information on which values were inferred, please see the notes on the "data_Background" worksheet.

"data_Population": This sheet provides the number of children (age 0-19), adults (age20-64) and seniors (age 65+) living in each of the 20 defined sectors around each refinery. The values are based on year 2001 Canadian census data that were mapped using ArcMap software[1]. The number of people residing in each sector was estimated using the average number of people living per square kilometer for each dissemination area (A DA, or dissemination area, is the smallest geographic area for which census data are reported. DAs vary in shape and size depending on the population density).

"data_PM conc" : This sheet provides the concentration of $PM_{2.5}$ predicted for each sector. This page provides calculations of secondary $PM_{2.5}$ formation from NO_x, and SO_x in addition to primary $PM_{2.5}$. Additionally, the proportion of total $PM_{2.5}$ attributable to each of these three "sources" is calculated. These totals include $PM_{2.5}$ from primary sources as well as NO_x and SO_x.

"data_Toxicity": This sheet provides toxicological parameters used in the equations which predict incidence of disease for each chemical in class I and class II. The preferred datum form for each substance was unit risk for carcinogenic endpoints, and ED05 values for noncarcinogenic endpoints. Ideally, all information would have been available from Health Canada. In practice, alternative values were also collected from the USEPA and from CalEPA where Health Canada values were not available. This sheet provides the toxicological value, the type of value, the endpoint on which the value is based, the source for the value, and EPA and IARC classifications for each substance. Carcinogenic parameters are highlighted in blue while noncarcinogenic parameters are highlighted in yellow.

"BTEX Tox": Because benzene, toluene, ethylbenzene and xylene have extremely similar endpoints, they are treated as a mixture by HEIDI II, and are ranked in terms of their cumulative (rather then individual) effects. Because each component of the BTEX class has a unique ED05 value and a unique concentration in each sector around the specified refinery, it was necessary to derive a weighted concentration and ED05 value for the class as a whole for each sector. A sample of the derivation method (simple "mixtures weighting") and the predicted weighted BTEX concentration for each sector around the given refinery are produced by this sheet. These serve as input to the predicted incidence equations for BTEX.

Users will also notice that weighted concentrations and ED05 values are also calculated for the "alternate ranking scenario" as well as for the specific background concentrations.

"PM Epidemiology2": This sheet provides the information and calculations underlying the risk coefficients used to related $PM_{2.5}$ exposure and outcome. The sheet provides sources for epidemiological data and calculates age-specific risk coefficients for each of the $PM_{2.5}$ "sources": NO_x, SO_x, and primary $PM_{2.5}$.

Refinery Sheets

There is a sheet named for each of Canada's 20 refineries. The structure of each page is identical, as is the purpose: to gather all the site-specific information for each refinery in one location. Site-specific information includes the specific refinery emissions, the specific background concentrations, the specific population in the area, and the refinery-produced $PM_{2.5}$.

"ActiveRefinery": When the user selects a specific refinery on the "SCENARIO selection" sheet, the site-specific information for that location is sent to the ActiveRefinery sheet using Visual Basic code. The data for the specified refinery are replicated exactly; the VBA code does not alter any parameters or perform calculations.

The "ActiveRefinery" page also contains a column listing the reporting thresholds for each substance. If desired by the user, 90% of the reporting threshold can be used for ranking substances that are reported in the NPRI as not being emitted. This is not currently the default option, but can be selected by the user (See "SCENARIO selection").

"Hypothetical Refinery": This page is used to create a hypothetical "worst-case" refinery. The emissions values for each substance are the maximum observed emissions across all refineries. For acetaldehyde and formaldehyde, which were not reported in the 2001 NPRI dataset provided, 95% of the reporting threshold was used. In order to create a ranking for this hypothetical refinery, background data and population data from Burnaby were used.

Because the CAC data required some rather complex manipulation, determining a parallel "worst-case" dataset for $PM_{2.5}$ was difficult. The benefit of including a hypothetical refinery is that it allows comparison of rankings for a case of all nonzero, emissions. In the case of the CACs, there are always emissions, and so performing a hypothetical ranking provides no added benefit for $PM_{2.5}$. As a result, no $PM_{2.5}$ data are included on this sheet.

Modeling Sheets

On all of the following modeling sheets, there are two sets of values, one set below the other. The methodology for each set is identical; the input is different: values are calculated for both the primary scenario (emissions as reported by the NPRI) and the alternate scenario (using a percentage of the reporting threshold where emissions are reported to be zero).

"Dispersed conc": This sheet calculates the concentration of each substance after dispersion to each of the 20 sectors around the refinery. The dispersion modeling adjustment factors were derived using the ISC3/AERMOD package and are specific to each sector and stack height. The concentration of dispersed chemical at each sector location is based on the amount of chemical emitted, the specific refinery, and the adjustment factors.

“Degraded Dispersed”: This sheet calculates the amount of substance remaining in each sector after a period of photodegradation. The default setting is 12 hours of photodegradation, but the user can select either 8 or 16 hours on the “SCENARIO selection” sheet. For more information on how degraded concentrations were calculated please see the notes on the sheet itself.

“Delivered Concentration”: This page adds background concentrations to the concentrations of each substance predicted to be in each sector as a result of refinery emissions after photodegradation. The values calculated on this page are used in predicting incidence for substances that act in a nonlinear threshold manner only. (Please see section 2. “Discounting in HEIDI II” in Chapter V. Discussion for further information).

“Case Incidence (undisc.)”: There are three of these sheets: one for each age group (child, adult, and senior). The values on these sheets predict case incidence based on concentrations from the “degraded Dispersed” sheet for class I substances, based on the “Delivered concentration” sheet for class II substances, and based on the “data_PM conc” sheet for class III substances. Therefore, for nonlinear, threshold-acting substances (the “noncarcinogens”), these sheets calculate the predicted incidence of disease for each substance resulting from both the background concentration and the concentration delivered from the refinery. For non-threshold-acting substances (the “carcinogens” and the “CACs”) the case incidence is predicted based on the concentration delivered by the refinery only. (Please see section 2. “Discounting in HEIDI II” in Chapter V. Discussion for further information).

“Case Incidence (disc)”: There are three of these sheets: one for each age group (child, adult, and senior). The purpose of this sheet is to account for case incidence resulting from exposure to background levels of each substance (i.e., ambient levels not originating from the refinery). Discounting is relevant only for those substances for which the dose-response curve is nonlinear and for which a threshold exists. As a result, for class I and class III substances, the case incidence that appears on this page and on the “Case Incidence (undisc)” page are identical, and based only on the concentrations delivered by the refineries to each sector. For class II substances, the values represent background-discounted case incidence (please see the section 2.“Discounting in HEIDI II” in Chapter V. Discussion for further information).

“Total case Incidence (disc)”: This page sums the case incidence across age groups (for children, adults and seniors) and across sectors (thus predicting total case incidence within a 25 km radius of the refinery).

There is also a column on this page which allows for age-sex discounting. This occurs when an endpoint identified for a give substance is appropriate for only one segment of the population (for example, ovarian cancer would only be predicted for females). (Please see section 2. “Discounting in HEIDI II” in Chapter V. Discussion for further information).

3. Understanding the Output of HEIDI II

The output sheets, "Health Impact" and "Health Impact(print)" contain a great deal of information. This section details the contents and meaning of the information in each column of the worksheet.

At the top left corner, the user can see which selections have been made: the specific refinery that the output relates to, as well and the stack height and photodegradation times selected. If these variables are not appropriate for the user's needs, s/he should return to the "SCENARIO selection" page to change them.

"Informational" Output columns:

The first eight major column headings in the "Health Impacts" output sheet (described below) are actually collections of information and data required by HEIDI to perform calculations and rankings. While some of this information is also available elsewhere in the workbook, it is convenient to present important health-related information with the impact and ranking predictions.

In the case of Class III substances CACs, health endpoints are related to death and disease using epidemiological coefficients. As a result, some of the "informational output" is not relevant to this class of emissions. Only the class and substance identifiers, relevant human endpoints and the DALY values are reproduced for these substances. For more information about the derivation of the risk coefficients, see the "PM Epidemiology2" sheet.

"effect class" The substances that are ranked in HEIDI II are categorized as belonging to one of three classes:

I – carcinogens
II – non-carcinogens
III – CACs (criteria air contaminants)

"NPRI substance" – this column contains the individual names of the chemicals being ranked as a part of HEIDI II. Note that there are two groups of chemical mixtures considered within HEIDI II: PAHs (polycyclic aromatic hydrocarbons) within class I (carcinogens), and BTEX (benzene, toluene, ethylbenzene, and xylene) within class II (non-carcinogens).

"CAS number (Ont MOE)" – provides an Ontario government reference number for each specific chemical.

"Reported refinery emissions" – recaps the information provided in the NPRI (in tonnes/year) for the specific refinery.

"Toxicity parameter" – essentially summarizes information available from the "data_Toxicity" worksheet. The first column, *source*, provides the agency that developed the toxicological parameter being used in HEIDI II, where:

> "HC" is Health Canada
> "EPA" is the U.S. Environmental Protection Agency,
> "HEAST" is the Health Effects Summary Table of the US EPA,
> "CalEPA" is the California Environmental Protection Agency,
> "NTP" is the U.S. National Toxicology Program,
> "WHO" is the World Health Organization,
> "Ontario MOE" is the Ontario Ministry of the Environment, and
> Wiaderna et al.[2] refers to a specific study and authors.

Numerical Value provides the actual value of the toxicological parameter used, and *type* describes what sort of value it is, where "IUR" is Inhalation Unit Risk, "LOAEL" is Lowest Observed Adverse Effect Level, "BMC" is Benchmark Concentration, "TC" is tolerable concentration, and "NOAEL" is "No-Observed Adverse Effects Level", and "REL" is the Reference Exposure Level.

"toxicological endpoints" – also provides summary information from the "data_Tox" sheet. These columns provide information about the research on which the toxicological parameters are based, where *reference species* is the type of animal on which the research was based, *most sensitive endpoint* is the health endpoint that was observable at the lowest doses and was also deemed to be relevant for human health. Some of these endpoints are very specific physiologic processes, but are indicative of disease states. These are summarized in the *most relevant endpoint* column.

"equivalent human endpoints" – summarizes the most important and relevant human endpoints associated with exposure to each substance being ranked, as well as providing an indication of the relative severity of each endpoint.

"target groups" – Some endpoints are not applicable to every member of the population. Ovarian cancer, for example, would only apply to females. The predictions for incidence of disease must account for the fact that not everyone is affected by each endpoint. These columns indicate which members of each populations subgroup (male/female) and (child/adult/senior) are considered to be affected by the given endpoint in HEIDI II. A "1" indicates that they are; a "0" indicates that they are not.

"health impact factors" - These values provide the DALY (Disability Adjusted Life Year) values that are used to weight the impact predicted incidence of various diseases relative to each other. HEIDI provides two alternate sets of DALY values. The first is based on work done by Pennington et al.[3] and consists of the value of 6.7 with an applied divisor of 10 or 100 if the disease is deemed by ILSI to be at a severity level of 2 or 3, respectively. The rationale for this is

provided in their paper and elsewhere in the report on HEIDI II. The complex DALY values are derived from a variety of sources, references to which are provided in the comprehensive "Health Impacts" sheet.

Impact and Ranking Output

The impact and ranking output on the comprehensive "Health Impacts" sheet provides information based on both the primary ranking scenario (using NPRI emissions as reported) and the alternate ranking scenario (using a percentage of the NPRI reporting threshold for cases where emissions are reported to be zero). The primary scenario predictions are presented in columns AB-AQ and the alternate scenario predictions are presented in columns AS-BH. These alternate scenario predictions are presented on this sheet only and not on the summary "Health Impacts(print)" sheet. The information presented in the following section is relevant to both scenarios. The user is encouraged to be mindful of the source data when comparing outputs from the two scenarios.

The values in the remaining columns of the "Health Impacts" sheet are all calculated by HEIDI. The following points should be noted:

- Values of "**N/A**" arise when the reported emissions for the given substance are zero.
- Values of zero that appear in these columns indicate that while an emission was reported in the NPRI for this substance, the predicted incidence of disease (and therefore impact) is so low that it cannot be displayed by excel (excel can display values down to 10^{-27}).
- In the case of BTEX, impact values are calculated for the mixture as a whole but not for the individual components of the mixture.
- In the case of the CACs, there are three refineries for which $PM_{2.5}$ emissions data were unavailable. These refineries are Husky Prince George, Nova Corunna, and Parkland, Bowden. If these refineries are selected, "no emiss. data" will appear in the Health impact cells of the worksheet. Rankings for $PM_{2.5}$ cannot be calculated for these refineries.
- All rankings for CACs are (currently) based on a stack height of 30 metres and a photodegradation time of 12 hours.
- There is no "alternate scenario" for CAC emissions because there are never cases where reported emissions for $PM_{2.5}$ or its precursors is zero (although there are missing data for three refineries, as noted above).

This section contains three major column headings, "population health impact", "impact fraction", and "priority ranking score". Within each of these, the results are presented based on predicted incident cases, simple DALYs and complex DALYs. Thus there are three possible bases on which to view each impact measure.

"Predicted population Health Impact" – these columns provide absolute values for predicted number of incident cases and apply the DALY factors directly to these predictions. As a result,

the DALY columns essentially predict absolute disability adjusted life years. Bold values represent totals for the relevant class.

"Predicted impact Fraction (within class)" – these columns determine the fraction of the total impact (whether measured in terms of incidence or DALYs) that is attributable to each substance. These fractions are calculated within each of the three classes. As in the case of the rankings, comparing health impacts across classes may be meaningless. For the CACs, the impact fraction is determined for each "source" (NO_x, SO_x, and primary $PM_{2.5}$), encompassing each of the endpoints related to that source.

"Priority Ranking Score (within class)" - These columns provide rankings for the emitted substances where 1 indicates the highest priority score for reduction. Rankings in these columns are separated by classes such that class I substances are only ranked against each other, as are class II and class III substances. (Thus, three separate rankings are produced in each column). Again, rankings are calculated based on predicted incidence as well as on DALYs, and rank is evaluated within each class only, and not across classes. If a value appears in the same column more than once, it will be ranked at the same level each time. An example of this can be observed with the zero values that are produced when predicted incidence is extremely low. If several compounds are associated with predicted incidence of 0, they will all be assigned the (same) lowest priority ranking score.

Again, for the CACs, a ranking is assigned to each "source" (NO_x, SO_x, and primary $PM_{2.5}$) only, although this ranking does take into account predicted incidence and DALY values associated with each endpoint considered.

"Priority Ranking Score (across classes)" - These columns also provide rankings for the emitted substances where 1 indicates the highest priority score for reduction. These columns rank across classes, however such that one ranking is produced which captures all substances evaluated by HEIDI II. There are some concerns about the validity of ranking predictions across different classes and due consideration should be given to these concerns before using the across-class rankings.

Again, rankings are calculated based on predicted incidence as well as on DALYs. If a value appears in the same column more than once, it will be ranked at the same level each time. An example of this can be observed with the zero values that are produced when predicted incidence is extremely low. If several compounds are associated with predicted incidence of 0, they will all be assigned the (same) lowest priority ranking score.

Again, for the CACs, a ranking is assigned to each "source" (NO_x, SO_x, and primary $PM_{2.5}$) only, although this ranking does take into account predicted incidence and DALY values associated with each endpoint considered.

V. DISCUSSION

1. Principles underlying the development of the HEIDI II model

At the outset of the project, NERAM and the NFPRER Health Prioritization Sub-group identified a set of underlying principles for the development of the HEIDI II model to ensure the model's acceptance as a useful decision-making tool by NFPRER members and other interested governmental and non-governmental stakeholders. These principles are articulated as follows:

a. statement of criteria objectives for the HEIDI II model

The criteria statement for the HEIDI II project in the Statement of Work was revised jointly by NERAM and the NFPRER Health Prioritization Sub-group as follows:

> *The criteria for decision-making on the approach and assumptions underlying the expansion of HEIDI 1 Model 4c (issues identified in Table 1: Summary of Proposed Parameters for Expansion of Model 4c) are to meet the project terms of reference which seek the development of a risk-based screening tool to assist in establishing rank order priorities for petroleum refinery emission reductions with the goal of protecting population health. Decisions on issues will be guided strictly by achieving a reasonably accurate estimate of health impacts with the available information. All choices and assumptions will be made transparent and explicit in the report.*

b. model transparency

The NERAM team and the Sub-group understand the importance of the criteria statement, and the need for model transparency to avoid possible misinterpretation of outputs of HEIDI II. NERAM agrees that critical modeling assumptions will often tend to drive the results obtained from the HEIDI II ranking system, therefore it is important that all underlying subjective judgments and modeling assumptions are accessible and transparent to the Sub-group and third-party reviewers. Transparency is critical for acceptance of the tool by the Sub-group. A summary listing of underlying assumptions and limitations of the HEIDI II model is provided in Table 7, as well as in the User Guide.

c. presentation of key results

It is recognized that the output format for presentation of key results in the HEIDI II package will likely affect the interpretation and emphasis placed on certain types of findings by the Sub-group and other users. It is recognized that providing a single unidimensional scale of numerical or ordinal priority rankings oversimplifies a complex set of decision-criteria while providing little insight into the underlying factors that determine the rank ordering process. HEIDI II carefully builds essential risk communication aspects into the spreadsheet outputs and

the narrative explanations in the User Guide. Principal results are structured in such a way as to be reasonably comprehensible to non-experts, while providing an output matrix of crucial information regarding the rank ordering of substances for priority reduction including the following columns: critical effect, severity index of critical effect, population-weighted estimate of exposure, risk incidence, and potential health impact (e.g. DALY).

In addition to the 'default' format provided as the standard output format for routine use, alternative output formats and toggles for the selection of alternate defaults allow for better exploratory and confirmatory analysis of the HEIDI results.

d. ranking of CACs and air toxics

The NFPRER Health Prioritization Sub-group decided that the output of HEIDI II should be configured to provide both a separate (within class) and unified (across classes) relative ranking of health impacts for pollutants in each of the three classes of health effects: i) air toxics - carcinogens ii) air toxics - noncarcinogens, and iii) common air contaminants (CACs).

However, it is important to recognize that calculating priority rankings within a given effect category is itself a difficult challenge given the dissimilar nature of diverse health effects endpoints that vary widely by severity, duration, age of onset, and reversibility. The DALY health impact weightings can capture some, but not all, of these diverse health impact criteria.

To a greater extent, producing priority rankings that carry across all three classes of health effects is subject to more questionable assumptions about the HEIDI model's capacity to compare different health endpoints, considering that the data sources and statistical modeling techniques used in the model are fundamentally different between the three classes.

Some of the most salient modeling differences between the three health effects classes is summarized in Table 8 (Newhook, 2004).

Table 7: Underlying Assumptions and Limitations of the HEIDI II Model with Respect to NPRI Emissions, Air Modeling, Health Effects Modeling, and Health Impacts Assessment

METHODOLOGY		assumptions	limitations		advantages	weaknesses
Emission Inventory						
Quantification of emissions		Reported quantities for NPRI air emissions are reasonably accurate	NPRI air emissions are quantified by various approximation methods that may underestimate some emissions and overestimate others		Reduced measurement burden; allows standard methods of emissions measurement; encourages consistent inventory reporting across all refineries	Approximation methods for quantifying emissions may lead to biased or inaccurate estimates; not all refineries may adhere to the same measurement and reporting guidelines
Individual chemicals and chemical classes		Reported quantities for NPRI air emissions are both comprehensive within classes and mutually exclusive between classes for relevant chemicals	Inclusionary and exclusionary criteria for some classes of chemicals is ambiguous or inconsistent		Complex chemical mixtures can often be better characterized as a chemical class than as individual agents	Composition of complex mixtures is often poorly understood or oversimplified; unintentional double-counting or discounting of chemicals in inventory may occur
Source characterization		NPRI air emissions are independent of various source types (e.g. stack vs. fugitive emissions) and source characteristics (e.g. number and location of stacks)	Lack of source types and source characteristics may lead to oversimplification of emissions release locations and time dynamics; distinction between petrochemical processes and thermal generation processes is lost		Simplicity, reduced reporting burden	Absence of systematic databases for source types and source characteristics means that each refinery must be treated as a single point source of emissions
Time averaging over one year		NPRI air emissions reported on an annual basis are released at a constant daily mass equal to 1/365 of the annual mass	Peak emissions that may occur over days or weeks cannot be quantified or modeled		Simplicity, reduced reporting burden	Peaks and valleys of air emissions rates are not quantified or modeled; seasonal effects cannot be directly assessed
Reporting thresholds		Assumes that NPRI emissions that are reported as 'zero' are not released in any quantity (default model) or as 50% of the NPRI reporting threshold (alternate model)	Unable to distinguish between 'true-zero' NPRI emissions and "below-threshold' NPRI emissions; underestimation of possible health effects for 'below-threshold' substances; reduced statistical reliability		Reduced measurement burden	Presupposition that small emissions are inherently harmless; discontinuity in health effects estimation

METHODOLOGY		assumptions	limitations		advantages	weaknesses
Air exposure						
Release rate		Modeling assumes that pollutant releases are continuous and at a uniform rate at each refinery.	In reality the releases may be periodic and be subject to process variations, and environmental weather effects.		Allows the concentrations of pollutants in the atmosphere to be considered as annual averages.	Periodic events of higher-than-average emissions may pose risks that are not evaluated in the annual average type of evaluation
Source Characteristics		The modeling assumes that the emission originate from a fixed point source such as a vent stack	Emissions may in fact originate from a variation of point, area, and volume sources such as process and storage leakages, and fugitive sources such as ground spills.		Allows the model to work from a simplified point of emission of defined location and properties.	Model may not adequately represent the emissions from fugitive and volume sources.
Point Source – Stack Height		The model has been run using three different stack heights (30 metres, 15 metres, and 5 metres)	The variation in stack height results in significant variation of the results in the near stack area.		The model uses a typical stack profile, which is generally representative of conditions in refineries.	The model may not adequately represent some fugitive emissions and spill-type releases
Meteorology		A single location meteorology has been employed to represent conditions for all refinery locations	The use of a single representative meteorology description will mean that special location-specific meteorological features will be missed.		A considerable cost and time saving occurs with the implementation of a single meteorological profile.	For specific locations where the impacts must be known with greater confidence, a location-specific meteorological data set will be needed.
Terrain		A single location terrain has been employed to represent conditions for all refinery locations	The use of a single representative flat terrain description will mean that special location-specific terrain features will be missed.		A considerable cost and time saving occurs with the implementation of a single terrain profile.	For specific locations where the impacts must be known with a greater degree of confidence, a location-specific terrain description will be needed.
Averaging Time		Pollutants are assessed over an annual period. The model reports annual averages.	Short-term exposures such as smog episodes and their impacts are not evaluated.		The model assumes that only chronic and long-term impacts will be evaluated.	Some short-term episodic events may impose an important risk.
Secondary Particulate Matter - ammonia availability		Secondary particulate matter is assumed to be a function of the SO2 and NOx inputs to the atmosphere.	Other controlling factors in secondary PM formation exist but not used in the estimation, specifically the alkalinization potential of ammonia is assumed to be available from non-refinery sources		The modeling assumes that ammonia is freely available to react with the SO2 and NOx precursor gases.	In many locations and at many times the ammonia may not be available to this extent and the reaction process will be limited.
Secondary Particulate Matter – reaction duration		Secondary particulate matter formation is assumed to occur under specific atmospheric conditions that occur for several hours per day.	Different locations and seasons will have varied atmospheric conditions.		A simple time toggle switch on the model allows the user to adjust to reaction period per day.	The time toggle switch does not guarantee the right conditions are being input.
Secondary Particulate Matter – reaction rate		Secondary PM formation is assumed occur at a rate of 5% per hour in smog-like conditions	A value of 4% per hour may be more applicable for “clean air” conditions.		The assumption simplifies the application.	Some potential for error exists in using one value, but the error is relatively small compared to others.

METHODOLOGY		assumptions	limitations		advantages	weaknesses
Health Effects - general						
Simple additivity of health effects		Simple additivity of health effects assumes that each substance (or mixture class) produces its health effects independent of any other; there are no biological interactions between substances that might amplify effects (synergism) or diminish effects (antagonism)	Fails to account for possible synergistic or antagonistic interactions when populations are co-exposed to many different air contaminants as a complex mixture (however, most existing studies suggest that such interactions are unlikely to occur at low exposure levels)		Permits simple priority ranking procedures; allows summation of overall health effects in various categories of air pollutants	Cannot account for possible (unspecified) interactions between different substances.
Interspecies extrapolation		Assumes that chronic toxicity studies in lab animals can provide a reasonable estimate of the toxic potency and type of health effects endpoint expected to occur in human populations	Lab animals may have different physiological responses to toxic substances than humans.		Permits more accurate concentration-response studies under controlled conditions, avoids the need for experimentation on human subjects; serves adequately for determining toxic potency most of the time	Questionable relevance to human health effects; cannot account for natural variability of susceptibility within human populations; cannot account for effects of age or ill health in human populations
Allometric adjustment (Human Equivalent Concentration)		Assumes that the differences in the body sizes of test animals and humans can be adjusted using standard conversion formulas to account for breathing rates, lung volumes ,etc.	Allometric scaling accounts only for body size differences; lab animals may have different physiological responses to toxic substances than humans		Human Equivalent Concentration (HEC) is a commonly used toxicity adjustment that allows animal data to be applied routinely to human populations	More sophisticated methods for interspecies extrapolation exist (e.g. PB-PK analysis), but are not commonly employed for air contaminants

METHODOLOGY		assumptions	limitations		advantages	weaknesses
Health Effects - air toxics carcinogens						
Dose-response linearity		Inhalation Unit Risk (IUR) assumes that a linear concentration-response (C-R) function applies for all air toxics carcinogens	IUR cannot model C-R functions that might be sublinear (lower than expected risk) or supralinear (higher than expected risk)		Has good theoretical and experimental support. General consensus among regulatory agencies for modeling carcinogens acting by genotoxic mechanisms.	Actual derivation of IUR numbers often vary between agencies (e.g. Health Canada, USEPA). Less applicable to carcinogens acting by non-genotoxic mechanisms
No threshold at low doses		IUR assumes that no threshold exists at low exposure levels; any level of exposure is expected to produce a cancer risk, although such risk may be small and could approximate zero	IUR would overestimate cancer risk if the exposed population were not susceptible to very low concentrations of air carcinogens		Helps to ensure protection of populations against carcinogens by assuming that exposure should be reduced towards zero whenever feasible	May be excessively biased towards a conservative (pessimistic) risk estimate at very low exposure levels
Ambient air background independence		Assumes background-independence for air concentrations, so that for each carcinogenic substance, the increased cancer risk from refinery emissions is independent of the risk from ambient background air concentrations of the same substance	Assumption cannot be easily confirmed by human studies.		Simplifies risk estimation methods for carcinogens, as ambient background air concentrations can be ignored.	Conventional but unconfirmed assumption
Case-incidence independence		Assumes a case-independence model for carcinogens, so that any increased cancer risk is related only to refinery emissions, regardless of the existing incidence of cancer in the exposed population	Although the case-independence assumption must be used when using data from animal models, estimation of human cancer risks can rely either on a case-independence model or a case-additivity model.		Simplifies risk estimation, as the baseline incidence of cancer cases in the exposed population can be ignored.	Conventional but unconfirmed assumption
Identification of carcinogens		Assumes air toxic substances with carcinogenic activity in humans have all been identified with high certainty	Only substances classified under CEPA as "carcinogenic to humans" or "expected to be carcinogenic to humans" are included		Includes only substances where evidence of carcinogenicity is reasonably strong	Substances with lower rankings of evidence of carcinogenicity are treated as non-carcinogens, even if evidence is lacking
Interspecies extrapolation for dose-response		Assumes that IUR values derived from animal tumour studies are adequate for predicting human cancer incidence	IUR values from animal studies are often "adjusted" for varying durations of exposure, and for Human Equivalent Concentration (HEC)		Animal studies usually have better experimental and statistical reliability than human studies	Interspecies extrapolation from test animals to humans includes several possible sources of uncertainty
Interspecies extrapolation on tumour endpoint		Assumes same tumour location and tumour type occurs in humans as in test animals	Animal tumour studies frequently produce tumours different that those found in human populations		Tumour location and type in humans are of secondary importance for health impact assessment	Some types of animal tumours may be irrelevant to the prediction of human cancer risk
Lifetime cancer risk		Assumes animal lifetime risk (12-24 months) is equivalent to human lifetime risk (70 years)	No obvious method of establishing lifetime equivalency between test animals and humans		Lower human cancer incidence rates balanced out by longer exposure duration in humans	Conventional but unconfirmed assumption

METHODOLOGY		assumptions	limitations		advantages	weaknesses
Health Effects - air toxics noncarcinogens						
Threshold dose		Assumes that the C-R function includes some form of toxicity 'threshold dose' below which adverse health effects are thought to be very small or negligible	Individual toxicity thresholds in human populations are broadly distributed (e.g. bell-curve) due to inherent differences in susceptibility to toxicants; thus it is difficult to establish a meaningful threshold dose for exposed populations		The predicted population response to very low exposures will be very small or negligible, reflecting the body's natural resistance to minor stressors	In any human population, a small fraction of persons may be susceptible to harmful health effects even at low levels of exposure below the nominal threshold
Continuous non-linear C-R function		Assumes that the C-R function should be continuous but non-linear, to reflect threshold-like C-R characteristics of non-carcinogenic air toxics	Some non-carcinogens may have no true toxicological threshold and may produce a linear dose-response (e.g. lead, mercury, other CNS neurotoxicants)		Allows threshold-like behaviour in C-R function without requiring a artificial cutoff between 'effect' and 'no-effect' dose levels	Not as intuitive to laypersons as a simple all-or-none threshold level; may not be suitable for CNS neurotoxicants
Distributional statistical model for concentration-response *log(dose):probit function*		Assumes a conventional 'distributional' model for the concentration-response (C-R) function based on a log(dose):probit statistical function; can model a complex non-linear threshold-like C-R behaviour, and transforms to a simpler linear function for extrapolation to low doses	The model may not adequately reflect more complex C-R patterns; some substances may follow linear(dose):probit function		Standard toxicological model for characterizing C-R relationships; enables prediction of population case-incidence at exposures below the nominal threshold	Cannot account for bimodal C-R distributions when a large hypersusceptible group exists in the exposed population; log(dose):probit function may not always hold in low-dose situations
Default C-R slope is set to a constant of 1.5 *modified Mantel-Bryan model*		Assumes that the slope of the log(dose):probit C-R function is 1.5 (modified Mantel-Bryan model) for all noncarcinogenic substances	Observed log (dose):probit slopes in animal studies are typically found in the range 2-3. A slope of 1.5 is intended to be somewhat more 'conservative', i.e. tends to overpredict possible human case-incidence at very low exposures		Avoids the need to obtain an observed C-R slope for each substance; slope value is protective of public health because it is conservative (pessimistic)	Conservative slope may overestimate true case-incidence in exposed populations, especially at very low doses; lacks empirical validation by data
ED05 as a surrogate measure of threshold dose		Assumes that the experimental ED05 value (i.e. the dose producing a 5% response in exposed test animals) can adequately represent a reasonable surrogate measure of threshold dose in animal studies, and that this value can be applied to humans after adjustment	Reliable ED05 values are not always available for some substances; other toxicity values with poorer statistical properties must sometimes be used (e.g. NOAEL); the ED05 does not provide information about other C-R data points or the C-R slope		Preferred toxicity parameter for C-R studies of noncarcinogens (Health Canada); more statistically reliable than alternate measures; reflects toxicity data without regulatory bias	Does not account for any of the major sources of uncertainty in estimating the toxicity parameter; no uncertainty factors are included to account for scientific uncertainty
Critical endpoint		For ED05 determination, assumes that the "critical endpoint' of toxic effects in animal studies corresponds to the most sensitive and most relevant health effect in human populations	Critical endpoint for ED05 determination in animals may not always correspond to relevant ED05 in humans		Standard assumption in risk assessment; allows use of animal toxicity data for assessing health risk in human populations	Major source of uncertainty in ED05 determination
Background discounting		Assumes that the health effects of noncarcinogen emissions from refineries can be obtained by calculating the predicted incidence due to the combined background and refinery-specific contaminants, then subtracting the predicted incidence due to background air contaminants	Refinery emissions and background air contaminant levels may not always superimpose in the same time frame, due or dissimilar day-night cycles or seasonal effects not reflected in annual averages		Accounts for the combined non-linear health effects of ambient background air emissions and refinery emissions, but provides the predicted net health effects due to refinery emissions separately	Refinery emissions are treated as equivalent toxicity to ambient background emissions; toxicity of refinery emissions might be greater (or lesser) due to secondary chemical reactions in air
Age-sex discounting		Assumes that health effects will occur only in the age and sex groups specified in the existing toxicity data; all other age and sex groups are discounted so as to produce no additional predicted cases	Toxic effects observed in one sex group of test animals may go unobserved in corresponding organs the other sex; characterization of age-specific health effects are not very reliable		Avoids counting age and sex groups "at-risk", when no actual risk would occurs in that population (e.g. risk to fetus in males over 65)	Age-sex discounting is a partly subjective process and available data are not complete for many substances.

METHODOLOGY		assumptions	limitations		advantages	weaknesses
Health Effects: Common Air Contaminants (CACs)						
PM predominant		Assumes particulate matter (PM) is the predominant CAC contributing to chronic health effects from smog constituents	Does not address the possible health effects of gaseous copollutants in CACs such as ozone, NOx, SOx, and CO.		Major cohort studies (Six-Cities, ACS), indicate PM is the predominant contributor to chronic health effects	Possible chronic health effects of gaseous copollutant CACs cannot be assessed.
PM2.5 predominant		Assumes that PM2.5 (fine particulate) is the predominant fraction of PM that contributes to chronic health effects	Does not address the possible health effects of PM10 (coarse particulate)		Many of the more serious chronic health effects of PM air pollution has been attributed to PM2.5	Several less serious health effects are apparently associated with PM10 (coarse + fine fractions)
PM2.5 is 50% of PM10		Where the C-R risk coefficient is reported only for PM10, assumes that PM2.5 constitutes 50% of PM10 mass	Fraction of PM2.5 within PM10 varies considerably by location; 50% is a relatively high fraction		Simplifies conversion of C-R risk coefficients based on PM10 to C-R for PM2.5 (2-fold adjustment factor)	Rough approximation of actual PM2.5 content in PM10.
NOx and SOx in PM2.5		Assumes that inorganic salts derived from NOx and SOx are the only relevant contributors to PM2.5 health effects	Elemental carbon, ammonium ion, metals, PAHs, and other salts may contribute to the relevant health effects of PM2.5		Simplifies attribution of predicted PM2.5 health effects to the major refinery emissions of NOx and SOx	Oversimplification of complex particulate health effects
Additive risks model		Assumes that C-R function for PM2.5 follows a conventional epidemiological 'additive risks' model for predicting case-incidence in exposed population	Requires accurate knowledge of the underlying incidence rate of relevant health conditions in the exposed population, by age groups; other models are possible (e.g. 'independent risks')		Conforms to prior assumptions used by epidemiologists to estimate the C-R risk coefficients in published studies	Accurate underlying incidence rates for relevant health conditions may not be available for some age groups (e.g. children)
Linear, nonthreshold assumption		Assumes that a linear, non-threshold C-R function is the appropriate model for predicting case incidence at low exposures	Possible threshold dose may exist for health effects at sufficiently low exposures, or the C-R function may be sublinear at low exposures		Simple model is supported by epidemiological evidence; conservative linearity assumption helps protect public health	May tend to overestimate predicted case incidence at low exposure levels
Prevalence to incidence conversions		Assumes that C-R risk coefficients based on underlying population prevalence data can be reliably converted to underlying annual incidence for each health endpoint	Conversion of prevalence data to incidence data requires additional information on onset and duration of each chronic condition; otherwise simplifying assumptions may produce inaccuracies		Allows use of published C-R risk coefficients based on prevalence data; data conversions are provided in Abt 2002 report	Abt 2002 conversion factors may not be applicable in all cases
Annual case incidence		Assumes that for each exposed individual, only one predicted incident case will occur per year for a specified endpoint	For less serious health effects, several incident cases might occur in each person in a given year (e.g. ER visits) but only one would be counted		Simplifies analysis	Tends to discount chronic health effects where repeated episodes might occur in a single year
Cardiopulmonary mortality		For the purposes of deriving C-R values, assumes that cardiopulmonary mortality related to PM exposure is equivalent to all-cause mortality	Other causes of death might possibly be related to PM2.5 exposure (e.g. stroke)		Standard simplifying assumption, reasonably supported by epidemiology studies	Cardiopulmonary mortality may not capture other possible causes of PM mortality

METHODOLOGY		assumptions	limitations		advantages	weaknesses
Health Impacts						
Utility function approach		Assumes health impacts are best quantified using a unified metric based on a 'utility function' approach that assigns weighting factors to different health effects endpoints using standard weighting criteria	Utility functions on a quantitative approach, which may fail to capture more subjective impact criteria such as equity, high-risk groups, and risk tolerance		Provides unified quantitative measure of health impacts for diverse types of health effects; permits objective ranking of emissions reduction priorities within and across health endpoints	Rankings within a particular category must quantify widely dissimilar health effects endpoints; rankings across various categories (e.g. carcinogens, CACs) are even more problematic
Disability Adjusted Life Years" (DALY)		Assumes that the Disability Adjusted Life Years" (DALY) approach is the best method for producing a health impacts utility function for the prioritization' of air contaminants	Other utility functions such as Willingness to Pay (WTP) may provide a better measure of societal preferences as they are quantified as personal choices expressed in monetary terms (dollars)		DALY is most commonly used utility function for environmental health impact assessment; based on strong methodological foundations	DALY approach focuses on physical health and functional disability, ignores some individual preferences and risk perception issues
'Global Burden of Disease' (GBD) approach		Assumes that the 'Global Burden of Disease' (GBD) approach for establishing DALY values is the best method for weighting health effects	GBD approach is one of several different methods for establishing DALY values; national DALY values may diverge from global values; not yet a standard system		Supported by WHO GBD research program; widely employed in EC countries, Australia, and Canada	Main focus to date is on infectious diseases and malnutrition; chronic diseases less well studied
DALY weights		Assumes GBD approach and related DALY methods (EBD) are sufficient developed to provide consistent DALY weights for air pollutant health effects	Inconsistencies across various GBD/EBD weighting systems regarding disease classifications, DALY weights, and underlying assumptions		GBD and EBD systems are gradually evolving towards a unified global consensus on DALY weights	More work needs to be done to establish unified DALY weight tables worldwide
Health endpoints		Assumes that health endpoints of widely varying severity and duration can be consistently assigned DALY weights in human populations	Toxicological endpoints in animal studies often do not correspond clearly with conventional human health effects endpoints (e.g. ICD disease classification)		Translates health effects incidence data to DALY health impact data for a given toxic effect in humans	Relies on subjective judgment by scientific experts.
Time-discounting and age discounting		Assumes that time-discounting and age discounting factors in DALYs are unimportant for environmental health	Time-discounting and age discounting factors are often important in developing countries; may be important in Canada		Simplifies DALY method in time-discounting and age-discounting are ignored	Possible oversimplification; fails to address importance placed on middle age population.
Prevalence and incidence data		Assumes that suitable adjustment factors can be applied for DALYs based on disease prevalence to obtain DALYs for disease incidence	Prevalence data is not readily convertible to incidence data; underlying assumptions about disease duration is unreliable		Allows available DALY weights based on prevalence to be adapted for health impacts based on incidence	Rough approximation using uncertain underlying conversion factors such as disease duration
Europe-Canada equivalency		Assumes the DALY weights derived in Europe (Netherlands) are equivalent to Canada DALY weights; similar assumption for N. America DALY data	Some DALY differences between Europe and Canada are possible, although they are likely to be small		Published DALY weights from Europe (Netherlands) should be a good equivalent of DALY weights in Canada	Ideally, DALY weights derived in Canada would be preferable, but are not available for air pollutants
Prioritization across pollutant classes* (*optional alternative to default analysis)		Assumes that DALY weights can be used to prioritize emissions reductions across the 3 major classes of pollutants (carcinogens, noncarcinogens, CACS	Health effects models used to derive DALY health impacts are dissimilar in data, structure, and assumptions between the 3 major classes of pollutants		DALY approach can provide prioritization rankings of health impacts across all 3 classes	The underlying statistical validity of DALY priority rankings across all 3 classes is open to question.

Table 8 Comparison of Attributes of Chemical Classes in Overall Estimation of Health Impacts

ATTRIBUTE	CHEMICAL CLASS			COMMENTS
	Carcinogenic air toxics	**Noncarcinogenic air toxics**	**CACs**	
Population in which critical effect observed	Relatively homogeneous groups of adult male workers or experimental animal models	Relatively homogeneous groups of adult male workers or experimental animal models	Relatively heterogeneous general population, including susceptible populations (the young, the old and/or compromised individuals)	Animal studies with PM indicate that there are vast differences in susceptibility between normal and compromised individuals, though it is hard based on available data to determine exactly how large this difference is. It is unclear how the measures of concentration response would be affected if studies of the general population, including susceptible subgroups, were available for carcinogenic and noncarcinogenic air toxics, instead of only for CACs. It is easy to hypothesize susceptible sub-groups for the air toxics that would not be revealed in currently available measures of exposure-response (e.g., renal toxicants affecting people with end-stage kidney disease)
Modeling of concentration-response	Assumed linear	Assumed threshold-like	Assumed linear	These assumptions are somewhat simplistic, though conventional
Nature of effect	Initiates new disease (case- independence)	Initiates new disease (case-independence)	Exacerbation of pre-existing health conditions (case-additivity)	While it is difficult to conceive how some of the outcomes associated with chronic exposure to PM would reflect exacerbation of a pre-existing condition (e.g., lung cancer), others may well be such a result (e.g., cardiovascular hospital admissions). The acute effects of PM are widely regarded as the result of exacerbation of a pre-existing condition
Reliance on animal data for critical effect and C-R measure	Mostly human data, animal data used for some substances	Generally heavier reliance on animal data than for cancer	Human data by definition, due to reliance on epidemiological studies	Owens (2002) concludes that – "For most chemicals, there is not apparent means to convert the critical effects in animal studies into time of human deaths or length and severity of disability for a DALYs approach. The current judgment then is that DALYs are not technically feasible to give an overall human health score incorporating chemical toxicity until the quantitation problem is addressed."
'Goodness of fit' of DALYs to critical effect	Quite good	Less good	Very good	The DALYs rely heavily on epidemiological evidence and have a strong clinical focus. The conditions for which DALYs have been developed correspond fairly well to the outcomes of interest for cancer and for PM health effects, but less well for non-cancer effects, for which the toxicological endpoints often do not correspond clearly to a clinical condition. For non-carcinogenic air toxics, the simple DALY scheme gets around this by having few classes, for which it is fairly straightforward to classify an effect, but the DALY value is based on little or no data. The complex DALY scheme has more categories, but there is difficulty in relating the toxicological endpoint to a corresponding clinical condition
Geographical scale of concern for critical effect	Principally local	Principally local	Principally regional	May tend to underestimate health effects of PM, since presumably secondary PM formation would continue well outside of 25 km limit
Atmospheric chemistry	Estimates photodegradation	Estimates photodegradation	Estimates secondary formation based on conversion factors	Unclear whether photodegradation effects modeled within a 25 km radius can adequately capture wider regional air pollution impacts

Source: Newhook, R. (2004).

e. interpretation of numerical data

The HEIDI II package and accompanying documentation includes a cautionary warning stating that numbers used to determine the output results are rough approximations and are only useful in a relative sense to derive ranking orderings for priority-setting purposes. These rankings rely on rough statistical estimates of predicted incidence rates for a variety of health endpoints of widely differing severity. Absolute estimates may include significant sources of bias that could introduce systematic over-estimates or under-estimates of absolute disease risk. Accordingly, the HEIDI II results must NOT be seen as a unidimensional ranking score that directly reflects the absolute estimates of risk in the exposed population. The comparative aspects of the tool are emphasized and incidences are expressed as a relative ratio.

2. Discounting in HEIDI II

"Discounting" refers to subtracting a certain number of predicted incident cases from the initial predictions, and is introduced to account for situations where a specific portion of the total incident cases would be predicted inappropriately for certain reasons. There are two types of discounting in HEIDI II: background discounting and age-sex discounting.

a. Background Discounting

HEIDI II predicts incidence due to emissions of a number of specific substances which are emitted from refineries. These substances also exist in varying amounts in ambient air as a result of other anthropogenic activities or from natural sources. In HEIDI II, the concentration of each substance in ambient air that is present and not due to refinery emissions is called the "background concentration". If the dose-response relationship for the substance is linear, discounting is not an issue: the amount emitted from the refinery can be used directly to predict the change in response, which in HEIDI II is equivalent to the predicted case incidence attributable to refinery emissions.

In the case of class II chemicals, the dose-response relationship used to predict case incidence in HEIDI II is not linear – it is a threshold-type relationship based on a log dose:probit function (for reasons that are fully explained elsewhere). Because of this, an incremental increase in "dose" from background levels could result in no increase in incidence if the total concentration of the given substance remains below the threshold, a dramatic increase if the threshold is crossed, or a moderate increase if the background concentration was already above threshold.

As a result, HEIDI II makes two predictions for incidence from each emitted chemical: one prediction of incidence due to the total amount in air (i.e., background + emitted, called "undiscounted") and one prediction of incidence from background only. The latter is subtracted (discounted) from the former to provide a best representation of the change in incidence arising from the refinery emissions only.

b. Age-Gender Discounting

HEIDI II initially predicts incidence under the assumption that each chemical has the same probability of affecting each member of the population. However, a careful examination of the toxicological endpoints that are most relevant for each chemical in HEIDI II's ranking list reveals that not all endpoints are appropriate for all members of the population. For example, the endpoint specified for cyclohexane is "reproductive/developmental." This endpoint could only apply to females of an age capable of reproduction (adults).

Because of this issue, HEIDI II identifies whether the relevant endpoint is applicable to males and/or females, as well as the appropriate age groups. To see which endpoints are age and/or gender-specific, go to the "target groups" columns on the "Health Impacts" Sheet. A "1" signifies that the endpoint applies to the group denoted in the column heading. A "0" signifies that it does not. A review of these columns reveals that the majority of endpoints are applicable to all members of the population.

3. Special Issues to note when using HEIDI II

HEIDI II was developed using the best available data in spring 2004. Some of the data used are subject to change as research is completed and knowledge expands. Users should note that some of the toxicological data in particular will be subject to change. Vanadium is one substance for which toxicological parameters in use by major agencies such as Health Canada and the USEPA may be updated in the near future.

The NPRI emissions data are from the 2001 database. The NPRI requires emissions to be reported each year, and ideally, the most recent available dataset should be used to produce rankings.

VI. CONCLUSIONS

The main conclusions listed below are based on the findings and analyses of the NERAM research group. They were obtained during the literature review, model development, data accumulation, and program coding phases of the HEIDI II project. These conclusions are entirely those of the NERAM project group, and do not necessarily reflect the views of CCME, NAICC-A or the Health Prioritization Subgroup of NFPRER.

Conclusion1.
The HEIDI II priority ranking tool has successfully demonstrated that it is possible to develop a consistent and objective methodological approach for ranking priority reductions of air emissions within the oil refinery sector in Canada.

Conclusion 2.
As HEIDI II is a fully functional prototype computer program, it can be used by decision-makers and other concerned parties to help inform the process whereby emissions reductions decisions are achieved. It can support decision-making with several user-configurable features that enable informed judgment about the interpretation of the ranking results -- these included program transparency, detailed descriptive information regarding health effects, alternate modes of output rankings within and across classes of substances, and sensitivity analysis of critical input parameters (stack heights, photodegradation time, imputed values for 'zero' reported emissions).

Conclusion 3.
The process by which HEIDI II was conceived and developed involved the active participation and input by the Health Prioritization Subgroup of NFPRER. Accordingly, HEIDI II represents a decision tool that has been developed by ongoing consultation with concerned stakeholders speaking as members of various organizations, including governmental, industrial, and nongovernmental organizations. Notwithstanding the inherent limitations of multistakeholder consensus-making in a technically-demanding field, the development of the HEIDI II ranking tool has been facilitated and informed by the active consultation process.

Conclusion 4.
The HEIDI II ranking tool requires numerous types and sources of input data in order to produce valid results. Much of the required data is readily available from published sources or governmental and industry records. However, the variable quality, immediacy, completeness, and relevance of the input data limits the ability of HEIDI II to produce unequivocal results in some cases. Careful attention to the underlying model assumptions and their limitations in HEIDI II must be maintained by the technical analyst and the decision-maker in order to avoid inappropriate conclusions based on excessive reliance on a single decision tool.

Conclusion 5.

As a working prototype model, the HEIDI II program will require further revision and updating over time to ensure the continuing validity and currency of the ranking results. The NPRI emissions data included in the HEIDI II prototype is based on the NPRI 2001 emissions inventory, and this will need to be updated annually. In addition, the input data used for other important model variables such as population density distributions, toxicological and epidemiological risk coefficients, and DALY health effects weights will require periodic revision.

VII. RECOMMENDATIONS

The recommendations listed below are those of the NERAM research group.

Recommendation 1:
The HEIDI II should be considered for adoption by NFPRER and CCME as one of the recommended decision tools to help inform the priority ranking of air emissions from oil refineries in Canada.

Recommendation 2:
Further research by NERAM on the validation and refinement of the HEIDI II program should continue, with the objective of assuring that the program produces ranking results that are computationally accurate and that adequately reflect the underlying assumptions and uncertainties inherent in the model formulation, input data, calculations, and output formats.

Recommendation 3.
Extension of the HEIDI II program to include new features deemed useful to decision-makers should be undertaken. A particularly useful extension of the program is development is a fully probabilistic version of the HEIDI II program using Monte-Carlo simulation or other statistical techniques that better characterize uncertainty.

Recommendation 4.
An administrative and technical arrangement for updating HEIDI II annually to reflect the most recent NPRI emissions inventory in the oil refinery sector should be put in place.

Recommendation 5.
To ensure that both technical analysts and non-expert users are able to become familiar with the operation and interpretation of the HEIDI II program, one or more training workshops should be instituted, where the NERAM project group has the opportunity to provide training and support to potential users of the program.

Recommendation 6.
As the HEIDI II priority-ranking program is generally adaptable to ranking of NPRI air emissions from other types of Canadian industrial and energy sectors, the program should be considered for adaptation and use in other sectors.

Recommendation 7.
The further application of more sophisticated health impacts measurement tools based on the DALY approach should be considered, as this field is in a rapid state of evolution towards better and more consistent methods in Canada and worldwide.

Recommendation 8.

The HEIDI II ranking tool should be compared and contrasted with similar priority ranking tools in other North American and European jurisdictions, to identify commonalities and critical differences regarding desirable features and model sophistication. A coordinated international movement towards 'best-practices' in priority ranking of air pollutants is deemed desirable, using approaches that may include the HEIDI II model.

References

Bennett, D.H., Margni, M.D., McKone, T.E., and Jolliet, O. 2002. Intake Fraction for Multimedia Pollutants: A Tool for Life Cycle Analysis and Comparative Risk Assessment. *Risk Analysis* 22, 5, 905-918.

Bouwes, N.W. Sr., and Hassur, S.M. 1999. User's manual for OPPT's risk-screening environmental indicators model: Version 1.0. Washington, DC: U.S. Environmental Protection Agency Office of Pollution Prevention and Toxics Economics, Exposure, and Technology Division.

Burke, T.A., Doull, J., McKone, T.E., Paustenbach, D.J., Scheuplein, R., Udo de Haes, H.A., Young, J.L. 1996. Human Health Assessment and Life Cycle Assessment: Analysis by an Expert Panel. International Life Sciences Institute, Washington, DC.

Brown, D. 2001. Air Quality Measures and Public Health Policy. *Public Health Reports*. 116, p. 61-63.

The Cadmus Group, Inc. 2000. A review of risk-based approaches for developing and implementing environmental standards. Winnipeg, Manitoba: Canadian Council of Minister of the Environment.

Cantox Environmental Inc. 1999, December. Public Health Risk Assessment of the Irving Oil Refinery Upgrade, Final Technical Document.

Carter, W.P.L. 1994. Development of ozone reactivity scales for volatile organic compounds. Journal of the Air & Waste Management Association, 44:881-899.

Davis, G., Swanson, M., Jones, S. 1994. Comparative Evaluation of Chemical Ranking and Scoring Methodologies. EPA Order No. 3N-3545-NAEX.

Kennedy, S.M., Copes, R., Henderson, S., Na, S. and MacKay, C. 2002. Air Emissions from the Chevron North Burnaby Refinery. Human Health Impact Assessment. Final Report. UBC School of Occupational and Environmental Hygiene.

Kyle, A.D., Wright, C.C., Caldwell, J.C., Buffler, P.A., and Woodruff, T.J. 2001. Evaluating the Health Significance of Hazardous Air Pollutants Using Monitoring Data. *Public Health Reports*. 116, p. 32-44.

Levelton Engineering Ltd and Purvin & Gertz Inc. Feb. 15, 2003. Benchmarking of Refinery Emissions Performance - Draft Report.

Levy, J., Wolff, S.K., and Evans, J.S. 2002. A Regression Based Approach for Estimating Primary and Secondary Particulate Matter Intake Fractions. *Risk Analysis*. 22, 5, p. 895-904.

Mantel, N, Bohidar, N.R, Brown, C.C, Ciminera, J.L, Tukey, J.W.1975. An improved Mantel-Bryan procedure for "safety" testing of carcinogens. *Cancer Res.* 1975 Apr;35 4:865-72.

Matthews, H.S., Lave, L. and MacLean, H. 2002. Life Cycle Impact Assessment: A Challenge for Risk Analysts. *Risk Analysis*,. 22, 5, p. 853-860.

McColl, S., Hicks, J., Craig, L., and Shortreed, J. 2003. Assessment of Comparative Human Health Risk-based Prioritization Schemes for Petroleum Refinery Emission Reductions. NERAM Final Report Submitted to CCME National Framework for Petroleum Refinery Emissions (NFPRER) Health Prioritization Sub-Group. Available at http://www.irr-neram.ca/.

Morgan, M.G. 2000. Categorizing Risks for Risk Ranking. *Risk Analysis*, 22, 5, p. 853-860.

Newhook, R. 2004. personal communication, on behalf of the NAICC-A National Framework for Petroleum Refinery Emissions (NFPRER) Health Prioritization Sub-Group.

Pandey, M. D., and Nathwani, J. S. 2003. Canada wide Standard for Particulate Matter and Ozone: Cost-Benefit Analysis Using a Life Quality Index. *Risk Analysis*, 23, 1, 2003.

Pennington, D.W., and Bare, J.C. 2001. Comparison of chemical screening and ranking approaches: The waste minimization prioritization tool versus toxic equivalency potentials. *Risk Analysis*, 21, 5, 897-912.

Pennington, D.W. and Yue, P.L. 2000. Options for the comparison of process design alternatives in the context of regional toxicological impacts. *Journal of Cleaner Production*,.8, p. 1-9.

Petro-Canada. 2002, December. Application for the Approval of the Edmonton Refinery Feed Conversion Program. Submitted to Alberta Energy and Utilities Board and Alberta Environment. Volumes 1 to 5.

Royal Society of Canada. 2001. Report of an Expert Panel to Review the Socio-Economic Models and Related Components Supporting the Development of Canada-Wide Standards for Particulate Matter and Ozone. http://www.rsc.ca/ozone_&_pm/indexEN.html

Shell Canada Scotford Upgrader June 1998. Environmental Impact Assessments (human health risk component).

Slubik, D.A. and Zemanek, M.G. 2000. Rating National Pollutant Release Inventory Releases Using a Comparative Risk Scheme. Presented at the Air and Waste Management Association International Specialty Conference, Exhibition and Workshop: Recent Advances In The Science And Management Of Air Toxics Banff, Alberta, Canada April 9-12, 2000.

Smuts, M. 2001. Hazardous Air Pollutants: Inside and Out. Public Health Reports. 116, p. 58-60.

Thornton, J.W., McCally, J. Houlihan, J. 2002. Biomonitoring of Industrial Pollutants: Health and Policy Implications of the Chemical Body Burden. *Public Health Reports*. 117, p. 315-323.

WHO, 2000. Global Burden of Disease Estimates. Available at http://www3.who.int/whosis/menu.cfm?path=evidence,burden,burden_estimates&language=english

Wu, C.Y., and Berglund, O. 2001. Analysis of air toxics emission inventory: toxicity-based ranking. International Emission Inventory Conference, One Atmosphere, One Inventory, Many Challenges. U.S. Environmental Protection Agency.

Wu, C.Y., and Pratt, G. 2001. Analysis of air toxics emission inventory: toxicity-based ranking. AWMA Technical Paper.

Appendices

APPENDIX A Summary of Parameter and Modeling Decisions

Table 9: Summary of Parameter and Modeling Decisions for HEIDI II

Parameter	Description of Approach	Documentation of Decisions
Criteria for Decisions on HEIDI II Approach and Assumptions	The criteria for decision-making on the approach and assumptions underlying the expansion of HEIDI Model 4c are to meet the project terms of reference which seek the development of a risk-based screening tool to assist in establishing rank order priorities for petroleum refinery emission reductions with the goal of protecting population health. Decisions on issues will be guided strictly by achieving a reasonably accurate estimate of health impacts with the available information. All choices and assumptions will be made transparent and explicit in the report.	Minutes of Oct. 2/03 meeting of NERAM Project Team and NFPRER Health Prioritization Sub-group October 31/03 NERAM Progress Report
Modeling Framework for CACs and air toxics	A single modeling framework is used for both criteria air contaminants and air toxics and the output is in the form of a matrix including relative rankings and several columns that inform on the rankings (critical effect, severity of critical effect, population weighted estimate of exposure, risk incidence, potential health impact e.g. DALY).	Minutes of Oct. 2/03 meeting of NERAM Project Team and NFPRER Health Prioritization Sub-group October 31/03 NERAM Progress Report
Output Format	The output sheets of HEIDI II provide both separate and combined aggregation of relative ranking of the three chemical classes (carcinogenic air toxics, non-carcinogenic air toxics and criteria air contaminants) by incident cases, simple DALYs, and complex DALYs. Whether the separate or aggregated priority rankings are considered, the default output for prioritization of NPRI reductions is a policy decision that lies outside of the HEIDI II package itself.	Prototype health impact output spreadsheet v.3 circulated to Andrew Snider and Ron Newhook on Feb. 16 as per Feb. 16 conference call minutes. Output format discussed between Ron Newhook and Steve McColl on March 31, 2004 and documented in email to NFPRER from Ron Newhook.
Pollutant Emissions	HEIDI II includes 29 air toxics including all polyaromatic hydrocarbons (PAHs) as a single class and benzene, toluene, ethylene, and xylene (BTEX) substances as a single class. Benzene is also treated as a carcinogen on its own but also has potential for chronic noncarcinogenic effects as a member of the BTEX class compounds. Refinery specific annual pollutant emission data is taken from the National Pollution Release Inventory 2001. HEIDI models fine primary particulate matter ($PM_{2.5}$) based on emissions reported in NPRI 2002 and secondary PM based on emissions of precursor gases (NO2 and SO2) reported in NPRI 2002. The formation of ozone from volatile organic compounds (VOC) is also modeled. In cases where zero emissions of a substance are reported in NPRI (i.e. true-zero or below the minimum NPRI reporting threshold), HEIDI II uses zero in ranking health impacts. HEIDI II offers an alternate ranking scenario for these substances whereby 50% of the minimum NPRI reporting level is assumed as the default value for estimating health impacts. Alternatively, the user can modify the default value from 1- 99% of the reporting threshold. The health impacts and priority ranking scores associated with the reported and default (50% or custom selected %) are displayed side by side for ease of comparison.	The air toxics were selected though ongoing consultation with the NAICC-A NFPRER Sub-group initially based on a review and discussion on Oct. 2/03 of Ron Newhook's proposed list of substances. The following selection criteria were used – quantity of emissions reported in NPRI 2001, CEPA-toxic substances, substances included on Health Canada Priority Substance List 2, and PSL scores for toxicity, persistence and bioaccumulation. Environment Canada draft record of decisions/action items from the March 26, 2004 NAICC-A NFPRER Health Prioritization Sub-group meeting Item 3.3.

Parameter	Description of Approach	Documentation of Decisions
	(50% or custom-selected %) are displayed side by side for ease of comparison.	
Air Dispersion Modeling	HEIDI II uses the US EPA air dispersion model AERMOD. Physical air distribution patterns are not site- specific. A generic meteorological profile representing the southern Ontario region is used as the default scenario. The report will include a non-quantitative discussion on how meteorology might affect the results in other Canadian locations. The modeling assumes that the refinery is located in flat terrain. This assumption is generally applicable to all Canadian refineries with the exception of the two refineries in British Columbia. A default of 30 metre stack height in the centre of the refinery is assumed. The stack height can be altered to 5 or 15 metre to reflect emissions that occur primarily at ground level. Steady state emissions are assumed over 365 days. This assumption fits with the NERAM proposal to focus on chronic annual health effects as the health endpoint -- all of the toxicological concentration-response (C-R) functions are based conceptually on 365 days continual exposure, while the epidemiological C-R risk coefficients have similar time dimensions.	Dec. 2/03 NERAM Progress Report Jan. 8/04 conference call minutes note Sub-group recommendation that sensitivity analysis of model assumptions for stack height and stack location be performed and discussed in the final report. March 03/04 NFPRER conference call minutes note Sub-group concern that default wind rose may not be representative of all refineries. Action item was noted for NERAM to include discussion of this limitation in the final report
Photooxidants and Secondary PM	HEIDI II does not perform nonlinear simulation modeling of photooxidant formation and secondary PM formation. Photooxidants and secondary PM are estimated using algebraic transfer functions and extrapolations based on conversion factors available in the research and technical literature -- primarily from controlled chamber studies. The intent is to keep the analysis method simple, so as not to obscure the transparency and interpretability of the HEIDI II results. NOx is considered both as a primary pollutant and as a precursor for PM2.5, but health effects modeling of these CACs in HEIDI II is restricted to PM only. The study team is aware of the double counting issue, and has ensured that the total formation of NO2 and nitrate ion is not more than 100% -- i.e. a mass-balance will be maintained for NOx and nitrate (also SO2 and sulfate). It is assumed that all secondary PM formation from photoxidants is <PM2.5	Dec. 2/03 NERAM Progress Report. NERAM summary of Feb. 16/04 NFPRER Sub-group conference call notes Sub-group comment that it would be useful if HEIDI II could estimate contribution of refinery emissions to local ozone concentrations.

Parameter	Description of Approach	Documentation of Decisions
Background Concentrations	The estimates of air contaminant concentrations are a function of the emissions from the refinery source and the "background" concentrations that exist in that refinery location from other anthropogenic and natural sources. Background concentrations were extracted primarily from Environment Canada's National Air Pollution Surveillance network based on monitoring information for sites represented in HEIDI II. Background concentrations for VOCs were provided by Tom Dann of Environment Canada. In cases where data were not available assumptions were made which are documented in the HEIDI II data_Background output sheet.	Minutes of Oct. 2/03, Jan. 28/04 conference calls with NFPRER Health Prioritization Sub-group. Dec. 2/03 NERAM progress report
Exposed Population	HEIDI II uses site-specific population distribution data based on Canada Census data. Generic age/sex distribution profiles are employed. HEIDI II assumes a 25 km outer boundary of effects with radial zones at 1, 2.5, 5, and 10 and 25 kms. For each refinery, HEIDI II estimates population average exposure for 20 locations at each refinery site based on 4 quadrants (NE, SE, SW, NW). Site specific population density information and generic Canadian age/sex distribution profiles were derived from 2001 Canadian Census data using ArcMap software. The following aggregated age classes were used: children (age 0-19), adults (age 20-64) and seniors (age 65+) to correspond with the concentration response functions for PM and to simplify the output.	Dec. 2/03 NERAM progress report.
Toxicological Parameters	HEIDI II uses Federal Canadian toxicity benchmarks (e.g. Health Canada CEPA PSL1/PSL2) to derive default dose-response parameters for air toxics. Alternative values were also collected from the USEPA and from CalEPA where Health Canada values were not available. The preferred datum form for each substance was unit risk for carcinogenic endpoints, and ED05 values for noncarcinogenic endpoints For criteria (ambient) air pollutants, dose-response parameters for mortality and morbidity endpoints are derived from chronic epidemiological studies as documented in Abt Associates (2002).	NFPRER Subgroup members (Ron Newhook and Geoff Granville) provided NERAM with ongoing feedback on the various iterations (11 versions) of the toxicity look up table (as per NERAM's Feb. 16/04 conference call minutes)
Physicochemical degradation rates	HEIDI II assumes a linear first-order rate constant of photochemical degradation for air toxics, based on the best available values obtained from the research literature.	Dec. 2/03 NERAM progress report

Parameter	Description of Approach	Documentation of Decisions
Concentration Response Parameters for Estimation of Population Health Effects	For CACs, HEIDI II estimates annual mortality incidence and three annual morbidity endpoints (e.g. annual chronic bronchitis, annual asthma hospitalizations, and asthma emergency room visits) for PM2.5 using dose-response parameters from cohort epidemiological studies reporte in the following reference: Abt Associates. Nov. 2002. Particulate-Related Health Impacts of Emissions in 2001 From 41 Major US Power Plants. Prepared for the Environmental Integrity Project. Rockefeller Family Fund. http://www.abtassociates.com/reports/Abt_41_power_plant_report_Nov19.pdf Among the PM-related health endpoints described in Exhibit 2-1 of Abt (2002) are several health endpoints with relatively low severity and duration -- specifically these are Minor restricted activity day (MRAD) (adjusted for asthma attacks), ages 18-65; Work loss days (WLDs), ages 18-65; and Asthma Attacks, all ages. To avoid double-counting of these cases with more severe disease conditions that are already captured under 'Hospital Admissions', these health endpoints are excluded from the HEIDI analysis. It is assumed that PM2.5 is an overall indicator of ambient air quality and the concentration response functions for the various PM related health endpoints include effects associated with exposure to gaseous pollutants. This approach recognizes the high correlation between PM and gaseous co-pollutants and will avoid double counting of health effects. The primary emissions inventories of gaseous CACs from oil refineries (SOx, NOx, VOCs) will be exclusively used to obtain an estimate the concentration of secondary PM (using air modeling and chemical conversion factors). When added together with the primary PM inventory from oil refineries, the combined PM (primary and secondary) concentrations serves as the basis of the concentration-response functions used to provide risk estimates of PM-related air pollution for various chronic health endpoints attributable to the CACs. No gaseous co-pollutants are included in the concentration-response analysis or risk estimates For carcinogenic air toxics, annual cancer incidence is estimated using Health Canada ED05 values or their equivalent inhalation unit risk (UR) values. The UR approach conservatively assumes a linear non-threshold C-R function. All substances rated as CEPA category I and II carcinogens will be modeled by this method. The Sub-group committee recommended that CEPA category III carcinogens should	Dec. 2/03 NERAM progress report

Concentration Response Parameters for Estimation of Population Health Effects	generally be modeled in the same manner as Category I and II (i.e. using inhalation unit risk parameters) if the PSL assessment includes an estimate of cancer potency (i.e., a TC05) value which would indicate that the weight of evidence for cancer was relatively strong. For estimating the population health effects of threshold-acting (mainly non-carcinogenic) air toxics, a toxicological log(dose):probit function based on the Mantel-Bryan extrapolation method is used to provide estimates of the expected case incidence rates in the exposed population from exposure to each NPRI substance (or class of substances). The Mantel-Bryan extrapolation is a reasonable statistical approach to estimating chronic health effects at low air concentrations. The Mantel-Bryan extrapolation is calculated with a default slope of 1.5 (modified from NERAM's original proposal to use a default slope of 1 following NERAM simulation studies) regardless of whether empirical concentration-response data is available to estimate a custom slope. The HEIDI II results and final report will make clear that the project is based solely on chronic studies based on 365-day time frame and that long-term chronic effects should be able to capture a considerable portion of clinical acute effects, however the overall public health impacts may be underestimated.	NERAM minutes of Jan. 8/04 NFPRER Health Sub-group conference call document decision to use default slope of 1.5 for the Mantel Bryan model (Item 3 of minutes).
Health Impact Measures	The issue of aggregation of diverse health effects (mortality, morbidity), of varying severity, with fundamentally different pathogenesis mechanisms will need to be addressed by the use of a common metric (or metrics) of health impact. HEIDI II provides output priority rankings in the Health Impacts section of HEIDI II by presenting 3 types of health impact measures: (1) case incidence (with descriptive note of type and severity of health effect for each substance); (2) simplified DALYs (based on the Pennington Table 2 categories and weights); and (3) complex DALYs (breaking out the Pennington categories into more detailed disease endpoints using WHO (or SETAC) DALY weights). Incidence for annual mortality and morbidity rates are converted to DALYs as the preferred 'burden of disease' metric -- the current best available DALY approach is used by WHO and SETAC. The SETAC consensus DALY methodology is used in HEIDI as it is the most closely related to the domain of environmental risk assessment of chemical hazards.	Dec. 2/03 NERAM progress report NERAM notes of Jan. 8/03, Jan. 28, Feb. 16 NFPRER Sub-group conference calls document NERAM rationale for use of DALYs supported by John Shortreed's discussion document on *Comparisons of Health Impacts for Different Classes of Air Emissions* (see Appendix C).

APPENDIX B Air Exposure Module Supplementary Outputs

Table 3.a Dispersion Modeling Factors For A 30 metre Stack Emission Annual Averaging Factor (To convert from source emission in g/s to receptor

Direction Degrees	500 metres	1,000 metres	2,500 metres	5,000 metres	10,000 metres	25,000 metres
45	0.688	0.334	0.104	0.048	0.029	0.013
135	0.692	0.270	0.083	0.045	0.033	0.016
225	0.219	0.102	0.037	0.020	0.015	0.007
315	0.349	0.149	0.058	0.031	0.023	0.013

Table 3.b Dispersion Modeling Factors For A 15 metre Stack Emission Annual Averaging Factor (To convert from source emission in g/s to receptor

Direction Degrees	500 metres	1,000 metres	2,500 metres	5,000 metres	10,000 metres	25,000 metres
45	1.633	0.787	0.442	0.223	0.086	0.019
135	1.390	0.765	0.547	0.295	0.117	0.027
225	0.538	0.334	0.263	0.150	0.060	0.014
315	0.787	0.486	0.486	0.228	0.095	0.023

Table 3.c Dispersion Modeling Factors For a 5 metre Stack Emission Annual Averaging Factor (To convert from source emission in g/s to receptor

Direction Degrees	500 metres	1,000 metres	2,500 metres	5,000 metres	10,000 metres	25,000 metres
45	10.1	3.51	0.783	0.240	0.071	0.013
135	12.9	4.58	1.038	0.321	0.095	0.017
225	6.70	2.37	0.535	0.165	0.048	0.009
315	9.63	3.44	0.791	0.249	0.075	0.014

Table 10 Estimated Particulate Matter from Primary and Secondary Pollutants

$PM_{2.5}$ in ug/m^3 assuming half day unreactive, half day reactive SO_2 and NOx*

Refinery	Distance from facility (metres)	Direction (Degrees)			
		45	135	225	315
Shell Montreal East	500	9.3	9.3	3.0	4.7
	1000	5.9	4.8	1.8	2.7
	2500	2.5	2.0	0.9	1.4
	5000	1.6	1.5	0.7	1.1
	10000	1.6	1.8	0.8	1.2
	25000	1.2	1.6	0.7	1.2
Shell Sarnia	500	18.5	18.7	5.9	9.4
	1000	11.4	9.2	3.5	5.1
	2500	4.7	3.7	1.7	2.5
	5000	3.0	2.8	1.2	1.9
	10000	2.8	3.1	1.4	2.2
	25000	2.1	2.8	1.2	2.1
Shell Scotford	500	1.1	1.1	0.4	0.6
	1000	0.7	0.6	0.2	0.3
	2500	0.3	0.2	0.1	0.2
	5000	0.2	0.2	0.1	0.1
	10000	0.2	0.2	0.1	0.1
	25000	0.1	0.2	0.1	0.1
Chevron Burnaby	500	3.6	3.6	1.1	1.8
	1000	2.0	1.6	0.6	0.9
	2500	0.7	0.6	0.3	0.4
	5000	0.4	0.4	0.2	0.3
	10000	0.4	0.4	0.2	0.3
	25000	0.3	0.3	0.2	0.3
Imperial Oil Dartmouth	500	9.0	9.0	2.9	4.6
	1000	5.9	4.7	1.8	2.6
	2500	2.5	2.0	0.9	1.4
	5000	1.7	1.6	0.7	1.1
	10000	1.6	1.8	0.8	1.3
	25000	1.3	1.7	0.7	1.3
Imperial Oil Nanticoke	500	5.5	5.6	1.8	2.8
	1000	4.2	3.4	1.3	1.9
	2500	2.0	1.6	0.7	1.1
	5000	1.5	1.4	0.6	1.0
	10000	1.5	1.7	0.8	1.2
	25000	1.3	1.7	0.7	1.3
Imperial Oil Sarnia	500	17.7	17.8	5.6	9.0
	1000	13.3	10.8	4.1	6.0
	2500	6.3	5.0	2.2	3.4
	5000	4.5	4.2	1.9	2.9
	10000	4.6	5.2	2.4	3.6
	25000	3.8	5.1	2.2	3.8

Refinery	Distance from facility (metres)	Direction (Degrees)			
		45	135	225	315
Imperial Oil Strathcona	500	13.5	13.5	4.3	6.8
	1000	7.6	6.2	2.3	3.4
	2500	2.9	2.3	1.0	1.6
	5000	1.7	1.6	0.7	1.1
	10000	1.5	1.7	0.8	1.2
	25000	1.1	1.4	0.6	1.1
Ultramar Jean Gaulin	500	3.0	3.0	0.9	1.5
	1000	2.3	1.9	0.7	1.0
	2500	1.1	0.9	0.4	0.6
	5000	0.8	0.8	0.3	0.5
	10000	0.8	0.9	0.4	0.7
	25000	0.7	0.9	0.4	0.7
Husky Lloydminster	500				
	1000				
	2500	Inadequate Input Data			
	5000				
	10000				
	25000				
Husky Prince George	500				
	1000				
	2500	Inadequate Input Data			
	5000				
	10000				
	25000				
Nova Corunna	500				
	1000				
	2500	Inadequate Input Data			
	5000				
	10000				
	25000				
Petro-Canada Edmonton	500	4.4	4.5	1.4	2.2
	1000	2.9	2.3	0.9	1.3
	2500	1.2	1.0	0.4	0.7
	5000	0.8	0.8	0.3	0.5
	10000	0.8	0.9	0.4	0.6
	25000	0.6	0.8	0.4	0.6
Petro-Canada Mississauga	500	2.6	2.6	0.8	1.3
	1000	1.7	1.3	0.5	0.7
	2500	0.7	0.6	0.2	0.4
	5000	0.5	0.4	0.2	0.3
	10000	0.4	0.5	0.2	0.3
	25000	0.3	0.5	0.2	0.3
Petro-Canada Montreal	500	5.3	5.3	1.7	2.7
	1000	3.5	2.8	1.1	1.6
	2500	1.5	1.2	0.5	0.8
	5000	1.0	1.0	0.4	0.7
	10000	1.0	1.1	0.5	0.8
	25000	0.8	1.1	0.5	0.8

Refinery	Distance from facility (metres)	Direction (Degrees)			
		45	135	225	315
Petro-Canada Oakville	500	10.2	10.2	3.2	5.2
	1000	6.1	4.9	1.9	2.7
	2500	2.4	1.9	0.9	1.3
	5000	1.5	1.4	0.6	1.0
	10000	1.4	1.6	0.7	1.1
	25000	1.0	1.4	0.6	1.0
Sunoco Sarnia	500	6.1	6.1	1.9	3.1
	1000	3.5	2.8	1.1	1.6
	2500	1.3	1.1	0.5	0.7
	5000	0.8	0.7	0.3	0.5
	10000	0.7	0.8	0.4	0.5
	25000	0.5	0.7	0.3	0.5
Irving NB	500	9.7	9.7	3.1	4.9
	1000	6.3	5.1	1.9	2.8
	2500	2.7	2.2	1.0	1.5
	5000	1.8	1.7	0.8	1.2
	10000	1.8	2.0	0.9	1.4
	25000	1.4	1.9	0.8	1.4
Consumers' Cooperative	500	3.3	3.3	1.0	1.7
	1000	2.6	2.1	0.8	1.1
	2500	1.2	1.0	0.4	0.7
	5000	0.9	0.8	0.4	0.6
	10000	0.9	1.1	0.5	0.7
	25000	0.8	1.0	0.5	0.8
North Atlantic	500	18.1	18.2	5.8	9.2
	1000	12.2	9.8	3.7	5.5
	2500	5.3	4.3	1.9	2.9
	5000	3.6	3.4	1.5	2.3
	10000	3.5	4.0	1.8	2.8
	25000	2.8	3.8	1.6	2.8
Parkland Bowden - 2001	500				
	1000				
	2500	Inadequate Input Data			
	5000				
	10000				
	25000				

** Note: primary particulate emitted all day, but secondary PM forms only in 12 hrs of daylight*

.

Table 11 Estimation of Maximum Amounts of Ozone Produced from Refinery Emissions (Tonnes per year)

This method uses a factor to estimate the maximum incremental reactivity of the available emission that are capable of producing ozone secondary products. The numerical listings in this table represent the estimated total annual tonnes of ozone that may be produced from refinery emissions, if optimum ozone production conditions existed every hour of the year. In reality these conditions are impossible to achieve and a much smaller amount of ozone is produced in periodic episodes. However, the table is useful as a relative measure of the contribution that different emissions are capable of contributing to photochemical smog.

	Facility Name										
Chemical Name	**Chevron Burnaby**	**Consumers' Co-operative Regina**	**Husky Prince George**	**Imperial Dartmouth**	**Imperial Nanticoke**	**Imperial Sarnia**	**Imperial Strathcona**	**Irving Saint John**	**North Atlantic Refining Come by Chance**	**NOVA Corunna**	**Parkland Bowden (Active?)**
1,2,4-Trimethylbenzene	5.897	136.333	0.000	33.958	24.339	61.197	41.304	9.479	127.133	21.063	26.305
1,3-Butadiene	0.000	0.000	0.000	0.000	0.000	1.690	0.000	1.243	0.000	281.551	0.000
7H-Dibenzo(c,g)carbazole	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Aluminum oxide (fibrous forms)	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Ammonia (Total)	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Anthracene	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.013	0.000
Asbestos (friable form)	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Benzene	5.830	65.218	0.000	24.681	16.697	47.146	16.188	5.519	71.744	205.137	18.159
Benzo(a)anthracene	0.001	0.000	0.000	0.057	0.075	1.347	0.029	0.000	0.008	0.004	0.000
Benzo(a)phenanthrene	0.000	0.000	0.000	0.409	0.310	1.557	0.102	0.000	0.000	0.042	0.000
Benzo(a)pyrene	0.002	0.000	0.000	0.026	0.045	1.946	0.019	0.000	0.000	0.008	0.000
Benzo(b)fluoranthene	0.000	0.000	0.000	0.026	0.025	0.672	0.012	0.000	0.004	0.017	0.000
Benzo(e)pyrene	0.000	0.000	0.000	0.051	0.044	0.080	0.019	0.000	0.002	0.000	0.000
Benzo(g,h,i)perylene	0.000	0.000	0.000	0.031	0.027	0.048	0.016	0.000	0.001	0.013	0.000
Benzo(j)fluoranthene	0.000	0.000	0.000	0.024	0.023	0.669	0.010	0.000	0.002	0.000	0.000
Benzo(k)fluoranthene	0.001	0.000	0.000	0.019	0.015	0.659	0.008	0.000	0.002	0.003	0.000
Biphenyl	0.270	0.000	0.000	0.000	0.000	0.000	3.942	0.000	0.000	0.000	0.000
Cadmium (and its compounds)	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Calcium fluoride	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Chlorine	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Chromium (and its compounds)	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Cobalt (and its compounds)	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Copper (and its compounds)	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Cresol (all isomers)	0.823	0.000	0.000	4.892	3.370	8.045	3.354	0.000	0.000	0.000	0.000
Cumene	0.438	1.169	0.000	0.000	2.172	5.737	3.701	0.584	0.000	0.000	0.000
Cyclohexane	3.337	201.924	8.153	31.853	17.999	76.763	9.606	15.231	97.328	20.597	15.971
Dibenz(a,j)acridine	0.000	0.000	0.000	0.001	0.002	0.002	0.002	0.000	0.014	0.000	0.000
Dibenzo(a,h)anthracene	0.000	0.000	0.000	0.001	0.004	0.004	0.002	0.000	0.000	0.001	0.000
Dibenzo(a,i)pyrene	0.000	0.000	0.000	0.001	0.004	0.004	0.001	0.000	0.001	0.000	0.000
Dicyclopentadiene (MIR est'd-JH)	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	122.585	0.000
Diethanolamine (and its salts)	0.000	0.000	0.000	7.464	6.492	0.000	0.000	0.000	25.816	0.000	0.000
Ethylbenzene	5.488	94.777	0.000	24.155	16.603	50.248	22.184	9.685	47.504	58.203	9.282

Ethylene	2.263	758.354	37.069	44.619	32.855	66.744	119.811	330.499	0.000	1560.937	0.000
Ethylene glycol	0.000	0.000	0.000	0.424	0.000	0.000	5.757	0.000	22.725	0.000	0.000
Fluoranthene	0.000	0.000	0.000	0.058	0.048	0.873	0.033	0.000	0.024	0.017	0.000
Hydrochloric acid	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Hydrogen fluoride	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Hydrogen sulphide	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Indeno(1,2,3-c,d)pyrene	0.002	0.000	0.000	0.033	0.018	0.039	0.013	0.000	0.000	0.000	0.000
Isoprene	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	46.114	0.000
Isopropyl alcohol	0.000	0.000	0.000	0.000	0.014	0.007	0.000	0.000	0.000	0.000	0.000
Lead (and its compounds)	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Manganese (and its compounds)	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Mercury (and its compounds)	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Methanol	0.020	0.000	0.000	1.931	1.360	50.505	0.013	0.000	0.000	0.000	0.000
Methyl ethyl ketone	0.000	0.000	0.000	0.000	0.000	605.682	264.154	0.000	0.000	0.000	0.000
Methyl isobutyl ketone	0.000	0.000	0.000	0.000	0.000	1317.868	309.805	0.000	0.000	0.000	0.000
Methyl tert-butyl ether	30.438	0.000	0.000	0.000	0.000	0.000	0.000	28.183	285.467	0.000	0.000
Molybdenum trioxide	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Naphthalene	0.899	11.813	0.000	7.665	4.687	10.914	15.613	1.027	0.000	18.464	0.000
n-Hexane	14.267	0.000	20.004	84.825	38.781	197.955	46.828	16.492	59.553	39.290	61.492
Nickel (and its compounds)	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Nitrate ion in solution at pH >= 6.0	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
N-Methyl-2-pyrrolidone	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
PAHs, total Schedule 1, Part 3	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Perylene	0.000	0.000	0.000	0.014	0.013	0.026	0.007	0.000	0.000	0.000	0.000
Phenanthrene	0.001	0.000	0.000	1.276	0.991	2.042	0.909	0.000	0.220	0.430	0.000
Phenol (and its salts)	0.230	0.531	0.000	6.322	2.692	30.337	5.065	0.000	2.302	0.000	0.000
Phosphorus (yellow or white)	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
p-Phenylenediamine	0.000	0.000	0.000	9.155	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Propylene	46.165	245.145	47.740	82.757	378.292	755.295	188.979	833.063	0.000	1109.573	0.000
Pyrene (est'd JH)	0.000	0.000	0.000	0.348	0.262	1.258	0.180	0.000	0.008	0.093	0.000
Styrene	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.323	0.000	39.696	0.000
Sulphuric acid	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Tetrachloroethylene	0.000	0.012	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Tetraethyl lead	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Toluene	29.052	379.722	13.518	133.096	75.606	230.175	126.011	43.090	377.308	275.618	93.084
Vanadium (except when in an alloy) and its compounds	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Xylene (all isomers)	14.760	407.100	0.000	167.340	98.085	287.925	198.285	56.100	315.600	140.610	66.090
Zinc (and its compounds)	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Grand Total	160.184	2302.097	126.484	667.513	721.950	3815.455	1381.964	1350.518	1432.766	3940.078	290.383

Chemical Name	Petro-Canada Edmonton	Petro-Canada Mississauga	Petro-Canada Montreal	Petro-Canada Oakville	Shell Montreal East	Shell Sarnia	Shell Scotford	Sunoco Sarnia	Ultramar St. Romuald	Grand Total	Percent of Total
1,2,4-Trimethylbenzene	14.874	43.995	30.584	4.726	37.359	50.993	37.317	50.923	4.154	761.933	2.8
1,3-Butadiene	0.000	0.000	0.124	2.510	0.000	0.000	0.000	0.000	0.000	287.117	1.0
7H-Dibenzo(c,g)carbazole	0.000	0.000	0.000	0.003	0.004	0.001	0.004	0.000	0.000	0.011	0.0
Aluminum oxide (fibrous forms)	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.0
Ammonia (Total)	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.0
Anthracene	0.000	0.000	0.000	0.000	0.244	0.000	0.077	0.000	0.000	0.334	0.0
Asbestos (friable form)	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.0
Benzene	10.876	14.988	78.954	3.504	34.558	143.103	48.916	97.600	18.181	927.001	3.3
Benzo(a)anthracene	0.005	0.000	0.095	0.078	0.101	0.123	0.006	0.093	0.025	2.046	0.0
Benzo(a)phenanthrene	0.005	0.000	0.252	0.190	0.185	0.368	0.010	0.202	0.075	3.707	0.0
Benzo(a)pyrene	0.002	0.000	0.056	0.053	0.072	0.060	0.005	0.147	0.011	2.451	0.0
Benzo(b)fluoranthene	0.002	0.000	0.038	0.002	0.028	0.084	0.003	0.037	0.030	0.978	0.0
Benzo(e)pyrene	0.003	0.000	0.076	0.049	0.046	0.023	0.032	0.099	0.016	0.539	0.0
Benzo(g,h,i)perylene	0.004	0.000	0.055	0.031	0.032	0.055	0.306	0.153	0.003	0.774	0.0
Benzo(j)fluoranthene	0.002	0.000	0.036	0.040	0.033	0.002	0.002	0.035	0.030	0.908	0.0
Benzo(k)fluoranthene	0.003	0.000	0.034	0.015	0.010	0.031	0.005	0.012	0.001	0.817	0.0
Biphenyl	0.000	0.000	0.655	0.000	0.000	0.000	17.848	0.000	0.000	22.714	0.1
Cadmium (and its compounds)	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.0
Calcium fluoride	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.0
Chlorine	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.0
Chromium (and its compounds)	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.0
Cobalt (and its compounds)	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.0
Copper (and its compounds)	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.0
Cresol (all isomers)	0.000	0.000	0.000	0.000	0.000	3.789	0.000	0.000	0.000	24.273	0.1
Cumene	0.000	0.000	2.688	0.000	2.143	0.000	0.010	13.733	0.000	32.376	0.1
Cyclohexane	20.306	2.395	12.710	5.795	14.157	46.572	1.106	24.471	23.997	650.271	2.3
Dibenz(a,j)acridine	0.000	0.000	0.002	0.005	0.005	0.000	0.008	0.014	0.000	0.056	0.0
Dibenzo(a,h)anthracene	0.000	0.000	0.000	0.006	0.005	0.002	0.004	0.000	0.001	0.029	0.0
Dibenzo(a,i)pyrene	0.000	0.000	0.002	0.008	0.008	0.004	0.002	0.015	0.001	0.050	0.0
Dicyclopentadiene (MIR est'd-JH)	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	122.585	0.4
Diethanolamine (and its salts)	6.479	0.000	0.000	6.812	20.320	59.005	0.000	0.000	0.000	132.388	0.5
Ethylbenzene	13.375	10.608	36.193	3.079	45.198	113.824	35.559	256.266	5.973	858.201	3.1
Ethylene	49.751	0.000	24.368	138.111	93.843	178.926	0.000	0.000	240.909	3679.059	13.3
Ethylene glycol	2.788	0.000	0.000	11.565	81.911	0.000	37.259	0.000	0.000	162.428	0.6

Fluoranthene	0.004	0.000	0.069	0.002	0.040	0.118	0.008	0.049	0.013	1.358	0.0
Hydrochloric acid	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.0
Hydrogen fluoride	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.0
Hydrogen sulphide	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.0
Indeno(1,2,3-c,d)pyrene	0.003	0.000	0.035	0.015	0.005	0.031	0.031	0.002	0.000	0.228	0.0
Isoprene	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	46.114	0.2
Isopropyl alcohol	0.000	0.000	0.000	0.000	79.611	8.140	0.000	0.000	0.000	87.772	0.3
Lead (and its compounds)	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.0
Manganese (and its compounds)	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.0
Mercury (and its compounds)	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.0
Methanol	0.000	0.000	3.887	0.000	0.000	3.463	0.000	0.337	0.000	61.515	0.2
Methyl ethyl ketone	0.000	203.011	0.000	0.000	200.258	0.000	0.000	0.000	0.000	1273.105	4.6
Methyl isobutyl ketone	0.000	0.000	0.000	0.000	123.806	0.000	0.000	0.000	0.000	1751.479	6.3
Methyl tert-butyl ether	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	344.088	1.2
Molybdenum trioxide	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.0
Naphthalene	0.000	29.969	5.252	4.558	11.684	8.898	5.598	6.805	0.578	144.424	0.5
n-Hexane	46.263	27.557	46.925	24.366	77.114	46.085	43.234	57.095	88.927	1037.052	3.7
Nickel (and its compounds)	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.0
Nitrate ion in solution at pH >= 6.0	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.0
N-Methyl-2-pyrrolidone	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.0
PAHs, total Schedule 1, Part 3	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.0
Perylene	0.002	0.000	0.026	0.017	0.000	0.010	0.007	0.038	0.004	0.163	0.0
Phenanthrene	0.080	0.007	1.296	1.028	1.647	2.036	0.743	0.560	0.292	13.558	0.0
Phenol (and its salts)	0.000	0.000	0.000	0.000	0.354	0.000	0.000	0.000	0.000	47.835	0.2
Phosphorus (yellow or white)	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.0
p-Phenylenediamine	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	9.155	0.0
Propylene	60.391	0.000	345.924	373.900	399.106	580.351	0.000	0.000	494.467	5941.148	21.5
Pyrene (est'd JH)	0.016	2.127	0.416	0.214	0.001	0.631	0.114	0.510	0.163	6.340	0.0
Styrene	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	40.019	0.1
Sulphuric acid	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.0
Tetrachloroethylene	0.000	0.000	0.000	0.000	0.000	0.288	0.000	0.000	0.008	0.308	0.0
Tetraethyl lead	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.0
Toluene	68.473	222.692	264.043	22.366	608.569	663.295	266.409	587.990	50.911	4531.030	16.4
Vanadium (except when in an alloy) and its compounds	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.0
Xylene (all isomers)	61.320	78.090	235.425	17.085	390.300	438.975	267.000	1390.770	41.535	4672.395	16.9
Zinc (and its compounds)	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.0
Grand Total	355.023	635.438	1090.219	620.131	2222.760	2349.287	761.622	2487.956	970.305	27682.130	100.0

Reaction of VOCs and the production of photochemical smog

This Appendix provides estimates of the contribution of primary emissions of VOCs from refineries to the production of secondary photochemical smog. In particular, estimates of the production of ozone under idealized conditions are provided. A description of this complex atmospheric reaction sequence and the conditions that contribute to photochemical reactions are summarized below. The estimation methodology is provided and the results for the Canadian refinery analyses are discussed.

Volatile organic compounds (VOCs) in the atmosphere are precursors to the formation of photochemical ozone in the troposphere. Other compounds such as NOx and carbon monoxide are also precursors and the rate of ozone production is dependent on the concentration of these materials in the tropospheric atmosphere. In addition, sunlight drives the photochemical reaction sequence, and the solar energy profile at different latitudes, seasons, and times of the day will change the ozone production rate dramatically.

It is necessary to note that a different set of reaction sequences and atmospheric mechanisms generate ozone in the stratosphere. The estimation of stratospheric ozone production is not considered as a component in the scope of this project and we have limited our assessment to the formation of tropospheric ozone only.

It is possible for natural sources to contribute organic emissions to the atmosphere that participate in the development of ozone. In addition, there are several anthropogenic sources of VOC's including vehicle emissions, coating industry releases, commercial and residential heating, as well as refinery emissions.

The reactions that contribute to the formation of photochemical ozone are complex and diverse. Different VOCs will react in different ways and utilize different chemical mechanisms to form smog. There are considerable temporal and spatial variations in the precursor conditions that are necessary for ozone formation. (Note: VOCs refer to organic materials in the troposphere with the exclusion of methane. Methane is considered to be relatively unreactive under most lower atmospheric conditions and is not an important contributor to photochemical smog, however methane may be found in such large quantities in some conditions, that the large quantity will make up for its slow reaction rate, and the contribution to total ozone creation may be significant.). In a general methodology, the formation process can be reduced to four major steps:

1) Reaction between VOCs or CO and OH radicals to form peroxy radicals $ROO^{\cdot}$.
2) The peroxy radicals oxidize NO to NO_2.
3) NO_2 is split by sunlight with the formation of NO and the release of atomic oxygen.
4) Oxygen atoms react with molecular oxygen to for ozone (O_3).

The production of ozone may be simplified to the following factors:

- availability of VOCs and CO,
- availability of NOx, and
- availability and intensity of sunlight to produce OH radicals and photolyze NOx.

It is assumed that the conditions for maximal production of ozone exist, and that there is ample NOx and sunlight available to drive the reaction forward. The availability of VOCs is considered to be the rate determining limitation on the model production of ozone. This is an assumption and is used only for the purposes of this exercise. The production of ozone is assumed to occur in maximal production conditions, and this will not occur on a regular basis but only on the rare episode where ideal conditions exist.

A relatively new methodology is used to estimate the contribution to ozone formation from the refinery emissions. The "Incremental Reactivity" method is based on laboratory measurements of VOCs and their contribution to ozone formation. Using controlled smog chamber studies, the atmospheric concentration of the VOC is multiplied by a "Maximal incremental reactivity" factor to estimate the potential ozone concentration under ideal formation conditions. The estimation method has been developed by Carter (1994) and improved over the past ten years to apply to a wide range of air contaminants.

It is important to note that the estimation methodology assumes that ideal ozone forming conditions exist. In fact, these conditions only exist periodically during a portion of the spring and summer days in Canada. In addition, the chemical mixture of reactive ozone forming species may occur in atmospheres that may exist in locations such as the lower mainland region of British Columbia, Southern Ontario and Quebec, and portions of the eastern coastline. Other locations, such as some prairie locations do not receive the high smog indices that occur elsewhere in the country. Consequently the estimates of ozone production are not precise and represent a worst case scenario.

Aside from the location differences noted above, there are several differences in the temporal formation of ozone. In the southern Ontario region, there may be 20 ozone events per year, and each one may last less than one-half of a 24 hour day. The HEIDI II model is designed to estimate health impacts from annual averages of air contaminant exposures. The formation of ozone in the atmosphere is based on periodic events where conditions exist to increase the airborne concentration to significant levels. The episodic nature of ozone formation, and the acute nature of ozone impact on human health are such that the photochemical oxidants are outside of the range of HEIDI's capabilities. Therefore HEIDI II does not provide an assessment of acute health risks associated with short term exposures to photochemical smog. This Appendix provides a basic estimation of the potential ozone producing capabilities of each VOC contaminant emitted from the refineries. The results are useful as a relative measure of ozone production capability of the refinery emissions. For example, the estimates indicate that the emissions of ethylene from the Regina Consumer Cooperative refinery contribute approximately 33% of the total refinery contribution to form ozone downwind.

Of the national refinery emissions, the specific contaminants that contribute most to the formation of photochemical smog are:

- Propylene with 21% of the contribution,
- Xylene with 17% of the contribution,
- Toluene with 16% of the contribution,
- Ethylene with 13% of the contribution,
- Methyl isobutyl ketone with 6%, and
- Methyl ethyl ketone with 4% of the total contribution.

It is notable that some refineries have a significant potential contribution to smog formation associated with certain contaminant emissions. For example the North Atlantic Refinery in Come-by-Chance, Nfld and Saskatchewan's Regina Consumer's Cooperative refinery both emit cyclohexane emissions that make up 6-7% of their total photochemical smog potential. At other refineries, cyclohexane is only 2% of the oxidant contribution

APPENDIX C Comparisons of Health Impacts for Different Classes of Air Emissions

BACKGROUND PAPER ON

COMPARISONS OF HEALTH IMPACTS FOR

DIFFERENT CLASSES OF AIR EMISSIONS

Prepared by

John Shortreed

NERAM

March 31, 2004.

1.0 Introduction

This document provides background discussion on the issues and methods for health impact comparisons between three classes of air emissions from refineries. The three classes are:

1. Carcinogens have significant mortality impacts. The magnitude of their impact is seldom directly observable and is usually estimated from animal studies.
2. Non-carcinogen emissions that are "toxic" and have mainly morbidity health impacts. The magnitude of their impacts is often estimated from animal studies and sometimes observable from occupational health studies.
3. Criteria Air Contaminants (CACs) emissions whose impacts are measurable by epidemiology studies and have significant measurable health effects with a predominance of public health burden related to mortality.

A health impact comparison is required if the different classes of emissions (note that occasionally an emission might have more than one of the above impacts) are to be combined in order to set priorities for emissions reduction with the objective of improving human health. Since this is the objective of the Health Prioritization Sub-Group of the NAICC-A National Framework for Petroleum Refinery Emissions Reduction, it is necessary to attempt a comparison and at the same time to be clear about its limitations so policy makers can determine its value. Conceptual and implementation limitations are discussed.

Understandably, scientists are hesitant to undertake comparisons of the risks of various types of emissions because of the wide variation in methods of estimating health effects and their associated uncertainties. For example, the health effects of most air toxics are estimated for purposes of public safety for "worst case" scenarios and include several factors of safety (uncertainty), while epidemiological studies for CACs attempt to provide a best estimate of the actual expected number of health effects to be observed in a given population. These health impact estimation issues are discussed below.

If public health is to be protected in the most expeditious way, i.e., by reducing the highest risks first, then it is necessary to attempt a comparison using a 'level playing field'. There are basically two ways to make this comparison, firstly informed decision makers can look at the priority rankings within the three classes and reach a decision based on their interpretation of the separate rankings; secondly, a common health metric can be used for all three classes and the priority rankings combined. The first approach has the advantage that any reasons for comparisons are explicit rather than being implicit within decisions made in the health metric methods. However, the complexity of the comparisons may mean that any unstructured evaluation of the three classes is open to bias. In HEIDI both comparisons are possible and the choice of approach is up to the user.

Inherent in the comparison of the three classes of emissions is the choice of the health impact measure. Currently there are three health impact measures widely used: QALY

(Quality Adjusted Life Years), DALY (Disability Adjusted Life Years), and WTP (Willingness to Pay). The recommended choice of DALY is primarily based on two considerations: firstly the use of DALYs by the World Health Organization and the associated rapid development of reliable and comparative parameters and methods, and secondly on the flexibility of DALYs to adapt to the decision principles desired by the users of HEIDI. For example, methods for calculating DALYs can be selected to broadly reflect the WTP metric or the QALY metric. Details of the rational for the choice of DALYs are given below.

2.0 Differences between the three classes of emissions?

It is helpful to understand the differences in methods used to estimate the health effects for each of the three classes of emissions. The differences between the three classes are directly linked to the history of each, the estimation methods used, and the risk management objective:

1. Carcinogen effects are estimated from studies that link an increase in cancer cases with an increase in exposure to a chemical. For some historical “high exposure” situations such as asbestos and benzene, use of epidemiological methods allowed direct observation of the link between exposure and cancer. For most other chemicals, a variety of toxicological animal studies are used to establish a dose-response function. The functions are usually extrapolated linearly into the low-dose ranges (i.e., it is assumed that there is no threshold for carcinogenic action) to obtain information for low exposures. This dose-response function is then used to find a “tolerably safe” or “de minimis” level of exposure for humans; typically an incidence of one in 1 million or one in 100,000 has been considered as a tolerable level of risk.
2. For non-carcinogenic air toxics, the public health objective is to determine under what exposure circumstances a chemical may be toxic to humans, and if it is deemed a significant risk, to set a maximum allowable ambient concentration of that chemical that corresponds to an exposure that is deemed “tolerable", “de minimis”, or "reasonably safe”. The estimation of a “tolerable dose” of the chemical is typically based on tests in animals and the determination of a dose that has no effect. The no-observed-effect level of dose (NOEL) is then reduced by a safety or uncertainty factor (typically in the range of 100 to 5000) to account for the most sensitive persons as well as to account for the possible difference between animals and people. Additional factors are also used to account for the use of a low-effect level rather than a no-effect level, or to account for limitations in the available toxicological data for the substance. The reduction may also include a factor for special protection of children and other groups in the population. The toxicity is related to specific health end points that are defined as harmful, and may include health effects such as irritation. In a few cases toxics may have negative impacts but also improve life expectancy. Generally morbidity effects are more common than mortality effects.
3. CACs impacts are average best estimates of the number of cases of mortality or morbidity per 100,000 population from a given exposure This is usually expressed as a Relative Risk (RR) for a unit of concentration, say an RR of 1.025 per 10 ug/m^3.

These estimates are made from either cross sectional studies or longitudinal studies of actual populations that are exposed to the ambient level of the emissions.

3.0 Issues and Concerns on the Comparison of Emission Classes in HEIDI II

In order to measure the health metric within each class and allow for the possibility of a comparison of the three classes of emissions there are a number of issues that need to be addressed. In general terms, these issues are:

1. How to combine and compare morbidity and mortality?
2. How to remove factors of safety from the methods, especially for toxics?
3. Are the uncertainties accounted for in a "comparable" way?

These issues interact in a complex way since there is uncertainty in the estimation of health effects, the exposure of people to emissions, the evaluation of health effects, and so forth. Table 1 (Newhook, 2004) compares a variety of attributes of these issues and sets the whole complexity of the comparison into context.

3.1 Comparing Mortality and Morbidity

Combining morbidity and mortality is a critical issue in combining health effects since class 1 (carcinogens) and class 3 (CACs) predominantly have mortality health impacts while class□2 (non-carcinogenic air toxics) has mainly morbidity impacts.

To compare mortality and morbidity, for any of the three classes of emissions, the first step is to establish a metric for mortality. This metric is usually expressed in terms such as years of life lost (YLL), or change in health adjusted life expectancy (HALE) attributable to a death from exposure to the chemical. For example, using this metric, late-in-life cancers or cancers that have a long latency period (such as cancers due to exposure to asbestos) will have relatively low measures of loss of life expectancy per case (i.e. < 10 years per case) relative to cancers such as radiation-induced leukemias which have a short latency and disproportionately impact children (i.e. >>10 years per case). The loss in life expectancy is calculated by comparing those with the disease to those without the disease in terms of the life table for each group (the difference being those who die from the exposure to the emission in question, prematurely relative to the life table) Those with the disease tend to die at a younger age and they have fewer years of life in comparison to those without the disease – the difference is estimated as the loss of life years. Table 2 gives some examples of loss of life expectancy, and the associated DALYs, due to mortality for a variety of cancers (Crettaz, 2002).

The next step in comparing mortality and morbidity is to estimate morbidity in terms of the metric loss of life years. The usual assumption is that this loss of Quality Adjusted Life Years (QALYs) or the accumulation of Disabled Adjusted Live Years (DALYs) is, for every year lived, a value between 0, or no effect, and 1, as equivalent to being dead. For example, a health outcome that results in an irreversible coma might be assigned a DALY of 1 or a loss of 1 life year for every year lived in the coma. (NOTE: the scale for

QALYs complements that of DALYs since one measures life lived and the other life lost).

Table 1 – Comparison of Attributes of Chemical Classes in Overall Estimation of Health Impacts

ATTRIBUTE	CHEMICAL CLASS			COMMENTS
	Carcinogenic air toxics	**Noncarcinogenic air toxics**	**CACs**	
Population in which critical effect observed	Relatively homogeneous groups of adult male workers or experimental animal models	Relatively homogeneous groups of adult male workers or experimental animal models	Relatively heterogeneous general population, including susceptible populations (the young, the old and/or compromised individuals)	Animal studies with PM indicate that there are vast differences in susceptibility between normal and compromised individuals, though it is hard based on available data to determine exactly how large this difference is. It is unclear how the measures of concentration response would be affected if studies of the general population, including susceptible subgroups, were available for carcinogenic and noncarcinogenic air toxics, instead of only for CACs. It is easy to hypothesize susceptible sub-groups for the air toxics that would not be revealed in currently available measures of exposure-response (e.g., renal toxicants affecting people with end-stage kidney disease)
Modelling of concentration-response	Assumed linear	Assumed threshold-like	Assumed linear	These assumptions are somewhat simplistic, though conventional
Nature of effect	Initiates new disease (casse- independence)	Initiates new disease (case-independence)	Exacerbation of pre-existing health conditions (case-additivity)	While it is difficult to conceive how some of the outcomes associated with chronic exposure to PM would reflect exacerbation of a pre-existing condition (e.g., lung cancer), others may well be such a result (e.g., cardiovascular hospital admissions). The acute effects of PM are widely regarded as the result of exacerbation of a pre-existing condition
Reliance on animal data for critical effect and C-R measure	Mostly human data, animal data used for some substances	Generally heavier reliance on animal data than for cancer	Human data by definition, due to reliance on epidemiological studies	Owens (2002) concludes that – "For most chemicals, there is not apparent means to convert the critical effects in animal studies into time of human deaths or length and severity of disability for a DALYs approach. The current judgement then is that DALYs are not technically feasible to give an overall human health score incorporating chemical toxicity until the quantitation problem is addressed."
'Goodness of fit' of DALYs to critical effect	Quite good	Less good	Very good	The DALYs rely heavily on epidemiological evidence and have a strong clinical focus. The conditions for which DALYs have been developed correspond fairly well to the outcomes of interest for cancer and for PM health effects, but less well for non-cancer effects, for which the toxicological endpoints often do not correspond clearly to a clinical condition. For non-carcinogenic air toxics, the simple DALY scheme gets around this by having few classes, for which it is fairly straightforward to classify an effect, but the DALY value is based on little or no data. The complex DALY scheme has more categories, but there is difficulty in relating the toxicological endpoint to a corresponding clinical condition
Geographical scale of concern for critical effect	Principally local	Principally local	Principally regional	May tend to underestimate health effects of PM, since presumably secondary PM formation would continue well outside of 25 km limit
Atmospheric chemistry	Estimates photodegradation	Estimates photodegradation	Estimates secondary formation based on conversion factors	Unclear whether photodegradation effects modeled within a 25 km radius can adequately capture wider regional air pollution impacts

Source: Newhook, R. (2004).

For morbidity, Table 2, shows the weights, W, which reflect the consensus WHO DALY estimates for the average proportional loss of normal ability for a year with the specified disease; these range from a reduction of 6.6% to 30.1% reduction in ability. For morbidity, this weight is combined with the estimated number of years that an incident case is affected before they are either cured or die to estimate the YLDp. The duration of the disability can be found for cancers from data in cancer registries.

Table 2 Disability Adjusted Life Years per Affected Person (DALYp) for Various Sites of Tumor, using WHO Data reported by Murray and Lopez (1996)

Type of Cancer	Disability Weight (DW)	Disability D(yr lost/ person)	Disability YLDp =DWxD	Death L (yr lost)	Death N(inc.)	YLLp=L/N (yr lost/inc	DALYp =YLDp+YLLp (yr lost/inc)
Mouth and Oropharynx	0.145	4.3	0.62	3.2E+06	1.1E+06	2.9	3.5
Oesophagus	0.217	1.7	0.37	3.4E+06	1.1E+06	8.9	9.3
Stomach	0.217	2.9	0.63	7.0E+06		6.5	7.2
Colon and Rectum	0.217	3.7	0.80	3.9E+06		3.9	4.7
Liver	0.239	1.6	0.38	6.3E+06	5.4E+06	11.6	12.0
Pancreas	0.301	1.2	0.37	1.5E+065	1.9E+05	7.9	8.3
Trachea, bronchus, lung	0.146	1.8	0.26	8.3E+06	1.1E+06	7.9	8.2
Melanoma	0.045	4.2	0.19	5.1E+06	1.7E+05	3.1	3.2
Breast	0.069	4.2	0.29	3.8E+06	1.1E+06	3.6	3.9
Cervix uteri	0.066	3.8	0.25	2.7E+06	4.5E+05	6.0	6.2
Corpus uteri	0.066	4.5	0.30	5.8E+05	3.1E+05	1.9	2.2
Ovary	0.081	3.4	0.28	1.3 E+06	2.0E+05	6.4	6.7
Prostate	0.113	4.2	0.47	1.1 E+06	6.8E+05	1.6	2.1
Bladder	0.085	4.2	0.36	9.8 E+05	4.6E+05	2.1	2.5
Lymphomas and myeloma	0.089	3.5	0.31	3.0E+06	4.2E+05	7.2	7.5
Leukemia	0.112	3.1	0.35	4.4E+06	3.1E+05	14.3	14.6
Cancers - terminal	0.809	n.a.		1.3 E+07	1.0E+06	13.0	13.0
Average							6.7

3.2 Factors of safety

For CACs these are direct estimates of premature mortality in the population, as well as morbidity as estimated by hospital emissions, incidence of asthma, etc. from health records. These estimates are best estimates of the average health impacts that actually occur in the population exposed. They do not need radical transformation or adjustment.

The concentration-response (C-R) parameter estimates are complicated by the age, sex, health condition, etc. of the population, the length of the exposure for different individuals, variations in exposure (e.g. indoor versus outdoor exposure), measurement of the concentration of the exposure, confounding factors such as smoking, low income, and so forth. So for even the most directly observable class of ambient pollutants, the criteria air pollutants, the resulting C-R estimates are uncertain but the expectation is that the average (best-fit) parameter values are usable.

Similarly, for carcinogenic air toxics, no major adjustment factors are present, beyond those already embedded in the statistical estimation of the C-R unit risk (UR) parameter

Air toxics, on the other hand, have usually been studied in order to produce a public health regulation for a 'safe' exposure level, by starting with an exposure with no observed effects and then introducing factors of safety or adjustments for a variety of factors including: uncertainly in animal studies, translation of results from animals to humans, and possibility of sensitive members of the population such as children. To compare CACs and carcinogenic air toxics it is necessary to discount these factors of uncertainty/safety.

4.0 Recommended Approach to Combining the Three Classes of Refinery Emissions

This section provides the rationale for selecting DALYs as defined by WHO as the recommended method for measuring health impacts. In following sections, DALYs are defined and the conceptual and computational issues and limitations are discussed. Finally particulars of the method as it is applied in this study are given on an emission by emission basis.

This study uses the WHO approach of DALYs (Disability Adjusted Life Years) for the reasons given below. It is also noted that others, faced with the same situation – the need to compare chemicals in the three classes for policy purposes – have proceeded with similar methods and using similar principles as those outlined in this document especially for the toxics class which is the most problematic for any health metric (de Hollander, 1999; Owens, 2002). To the extent possible, existing methods, principles, and parameter values from the literature are used.

4.1 General Reasons for the Selection of DALYs as Defined by WHO

1. WHO reviewed available methods for estimating health impacts using eight objectives including; “providing appropriate and balanced attention to the effects of non-fatal health outcomes on overall population health”, “informing debates on priorities for health service delivery and planning” and recommended the use of DALYs either with a zero time discount rate and uniform age weights or DALYs with a 3% time discount rate and non-uniform age weights (Murray, 1999). Since then WHO has proceeded to use DALYs for their Global Burden of Disease study to measure and evaluate in a comprehensive, mutually exclusive manner all health effects for a population. They are currently in the third round of refinements in estimates of parameters and review of implementation techniques and methods. For example, there are complete version 3 estimates for WHO region 3 (Canada and the US) for the year 2000. (http://www3.who.int/whosis/menu.cfm?path=evidence,burden,burden_estimates&language=english)
2. More recently, WHO has extended the use of DALYs to the environment and is beginning to address directly problems directly related to the comparison of refinery emissions (WHO, Pruss-Ustun, 2003). WHO considers the following environmental

risk factors; water for drinking, recreation or agriculture, food, wetlands, indoor and outdoor air, chemical substances, noise, and radiation.
(http://www.who.int/peh/burden/9241546204/9241546204toc.htm)

3. DALYs are compatible with HALEs (Health Adjusted Life Expectancy) since they use the same health data and weights. In fact, in 2001, DALE (Disability-Adjusted Life Expectancy) was renamed HALE by the WHO in response to feedback from member states (WHO, 2001, p118). This means that data collected for DALYs will not have to be modified in the future when more population health measures are introduced into health policy analysis and evaluation. Canada now uses HALEs to estimate health-adjusted life expectancy (Manuel, 2003) using a modified method for estimating DALYs
4. It seems desirable to estimate health effects directly and this is possible since it is the health impact that is of interest. Other alternatives such as Willingness to Pay (WTP) involve taking a dollar estimate of a health impact and translating this dollar value into an equivalent health effect. WTP methods also include non- health impacts on utilities and when it is desirable to separately consider health impacts and other policy issues then DALYs are the preferred choice. It is noted that conceptually QALYs and DALYs are equivalent measures; they both attempt to measure the social utility of health outcomes (Hammitt, 2002). However, it is noted that often QALY estimates are measured in a way that non health outcomes, including wealth, can impact the utility measured (Gold, 2002). While this was not the case for early estimates of DALYs, recent measurement techniques allow for decision makers to include these effects if they desire.
5. The morbidity health effects with DALYs are limited to a range of 0 to 1 in terms of comparison to the loss of one year due to mortality. This limitation of DALYs is an advantage since the estimation methods are subjective and depend on value judgements, especially in the aggregation of individual utilities to social preferences. In some cases, the use of standard gambles and other estimation methods when applied to measures such as WTP have resulted in very inconsistent estimates of health impacts.
6. WHO version 3.0 of DALY Burden of disease estimates are comprehensive and the decision maker can select particular versions of DALYs that suit their own judgements. Current methods for estimating DALY weights include: rating scales, standard gamble, time trade-offs, and person trade-offs. These can be done by either; individuals in the health state, health care providers, general public, or patients' families (WHO, 2001). While WHO are developing a multi-method protocol for combining these estimates, the decision-maker is free to select weights according to their own judgements. Choices include not only the measurement method but also the weighting of age preferences and the discounting of time. For example, a recent Canadian application of DALYs (Manuel, 2003), used HUI3 (Health Utility Index 3) to estimate DALY weights with no age weights and no time discounting.
7. The use of DALYs by WHO is limited to the main causes of morbidity and mortality and uses vital statistics, and a wealth of other health care data (WHO, 2003), especially for the US, Australia, Canada, and other countries. To incorporate additional causes such as the emissions from refineries, it is necessary to apportion existing health care

outcomes to these sub causes, or alternatively to make "relative" estimates of the DALYs. This is not a trivial task but given the overarching value associated with a standard approach to health status (i.e. WHO's approach to DALYs) a considerable literature is developing on making these estimates. For example, SETAC (Society for Environmental Toxicology and Chemistry) have addressed this issue (Owens, 2002) and have also provided some simplified methods which are used in this study for one of the DALY estimates. Crettaz (2002) and Pennington (2002) in a special issue of Risk Analysis illustrate the ongoing research for toxics and carcinogens, that will support the extension of the basic WHO population health estimates to other health priority studies such at this study on refinery emissions. De Hollander (1999) made DALY estimates for the Netherlands, for a number of environmental exposures including PM, Ozone, PAH, Benzene, Ethylene oxide, Vinyl Chloride, 1,2-Dichloroethane, Acylonitrite, Radon, passive smoke, lead, noise, and UV-A and B.

8. Last but not least is the issue of whose values should be used to evaluate health impacts. In keeping with the modern approach to risk management, the values of the decision-maker(s) is considered to be the most appropriate source of these values. In many of the methods, such as WTP the source of the values is "theoretical" values – for example, the principles of economic utility theory (Hofstetter, 2002). DALYs provide the most flexibility to build in the decision-maker(s) values in the health impact analysis.

4.2 Definition of DALY

Disability-Adjusted Life Years, was selected by WHO (Murray, 1996) as the preferred measure of health impacts that combine mortality and morbidity. The measure is related to health only (intent is to rule out effects of type of risk, wealth, etc. that modify the individual utility) and is intended to be a measure of society at large for purposes of public policy. The measure for society is constructed by the aggregation of individual measures of utilities of health status.

$$DALY = YLL + YLD \quad (1)$$

Where DALY is the disability-adjusted life years
YLL is the years of life lost due to premature mortality
YLD is the years lived with disability (morbidity)

$$YLD = DW \times L \quad (2)$$

Where YLD is the years lived with a disability
DW is the disability weight
L is the average duration of disability (years)

DALYs are a time measure of a health gap, the time lived with less than perfect health and the time lost due to death before a standard life expectancy (life expectancy at birth of 80 years for men and 82 years for women) (Pruss-Ustun, 2003). In many cases rather than the

standard life expectancy the life expectancy for the country considered is used, e.g. in Canada the age specific life expectancy is used.

Equations (1) and (2) are for an individual case. Usually DALYs are estimated for a total population, for an exposed population, and so forth. The equations are modified by inserting the number of cases. For a total population estimate, such as for the WHO burden of disease study, the total population is used and the equations are modified from cases (incidence) to prevalence in order to estimate DALYs.

Disability Weight (DW) is a weight that reflects the severity of the disease on a scale from 0 (equivalent to perfect health) to 1 (equivalent to dead). Depending on how these weights are determined they are called disability weights, QALY weights (complements of DW), health state valuations, health state preferences or health state utilities (Pruss-Ustun, 2003). DW do not represent the lived experience of any disability or health state, or imply any societal value for the person in a disability or health state, but rather quantify societal preferences for health states in relation to the societal ideal of good health (Pruss-Ustun, 2003). As a preference it means that society would be indifferent between any l person in the population living three years with a DW of .33 and any other person in the population dying one year prematurely, since the DALYs are the same.

The term disability is used broadly to refer to departures from good or ideal health in any important domain of health, including; mobility, self-care, participation in usual activities, pain and discomfort, anxiety and depression, and cognitive impairment (Pruss-Ustun, 2003). For example, HUI3 (Health Utility Index 3) asks people about the current status of their vision, hearing, speech, ambulation, dexterity, emotion, cognition, and pain (Hammitt, 2002; Manuel, 2003). HUI3 was constructed to reproduce individual utility measures for standard gambles, time trade-offs, and other methods to measure individual utilities (Torrance, 1995).

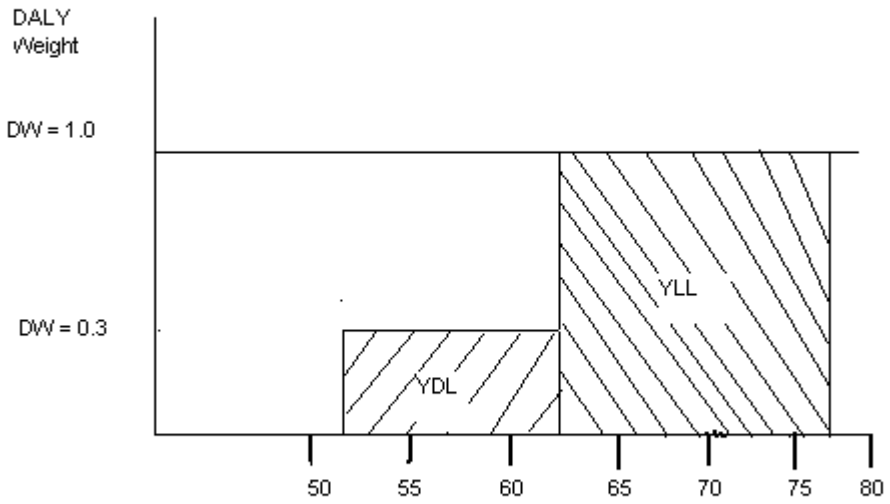
Estimates for the period 1998-2000 for life expectancy and HALE (Health Adjusted Life Expectancy) have been made by Health Canada (Manuel, 2003) and WHO (2001). Health Canada used HUI3, a utility based estimate of health-related quality of life while WHO used DALYs to estimate the utilities due to ill health. In both cases the utility estimates were in life years lost. The differences between the two estimates are small with WHO estimating the lost life years (Life Expectancy less HALE) for males as 7.7 years lost and females as 9.8 years. The Canadian estimate of lost life years for males was 8.7 years and females 10.4 years. This indicates that even with two quite different utility methods used to estimate DALYs or their equivalents, the resulting estimates are close but the HUI3 utility index gives utility estimates that are about 4 to 10% higher.

Figure 1 illustrates DALY for one case of a disease for an individual person in the 50-55 age group. The DALY weight for the cause/disease is 0.3 and the person has a premature death at age 62.5. At the onset of the disease at age 52.5 the life expectancy is 77.5.

In Figure 1 the estimated YDL is 0.3 x 10 years = 3 years. The YLL is 1 x 15 years = 15. The DALY is YDL + YLL or 18 years.

Time component of DALYs – DALYs can be discounted over time to reflect individual utility measurements that indicate people discount the future occurrence of health impacts similar to the way they discount future wealth (to accurately represent the value of future costs or benefits in today's dollars). If the decision maker decides to discount DALYs over time then WHO recommends a discount rate of 3%. For example, in Figure 1, considering only YLL discounting the years lost back to the age group 50-55 the value with discounting would be 6.65 since each year of death would be discounted back to age 52. Similarly YDL would be 2.59 rather than 3.0 and the total DALY would be 11.47 rather than 18. The calculation is illustrated in Table 3.

Figure 1 Example DALY for one incident or case



Age Weighting of DALYs - DALYs can be weighted by age, again to reflect the measurements of utility for health states that indicate people value the quality of life (or utility) more at middle age than when young or old. If the decision maker decides to use an age weighting, for example if their organization elsewhere uses WTP with a considerable age weighting adjustment, then it is recommended that the WHO recommendation for weights be used. This will ensure comparability with other DALY estimates of environmental exposure and health impacts. The recommended WHO weights vary from 0 at birth, to 1.52 at age 25, to 1.0 at age 55, and 0.3 at age 100. For the example in Table 2 with age weighting YLD changes from 3 to 2.9 and if both age and time are adjusted to 2.52. Similarly in Table 3 total DALYs are 18 with no adjustments, 11.47 adjusted for time, 13.67 adjusted for age, and 8.98 adjusted for time and age.

4.3 DALYs – General Conceptual Issues and Limitations

WHO set in motion a process to define the "standard" approach to measuring health impact (quantity and quality) for use in policy when they selected DALYs for their Burden of Disease study. This study was and is an ongoing attempt to comprehensively and mutually exclusively estimate the health impact of "causes". Causes are usually defined as specific diseases or health conditions such as HIV, myocardial infarctions, blindness, low birth weight, abortion, iron-deficiency anaemia, diarrhoeal episode, and so forth. The causes are rather eclectic but collectively are comprehensive and mutually exclusive, thus defining all health impacts in the world.

Currently WHO has about 130 "causes". For example, for cancers WHO identifies; mouth and oropharynx, oesophagus, stomach, colon and rectum, liver pancreas tracea, bronchus and lung, melanoma and other skin, breast, cervix uteri, corpus uteri, ovary, prostate, bladder, non-hodgkin lymphoma, hodgkin lymphoma, leukaemia, others. And these cancers are further broken down into time 4 time periods for analysis: Diagnosis/therapy; Waiting; Metastasis; Terminal.

Table 3 Example DALY Calculation for Figure 1.

						Parameters	Age b=	0.04	(toggle for age)
							k=	1	
							Discount r=	0.03	
AGE(yr)	DW	L(years)	DALY/yr	time discount	DALY/yr (Time disc)	age factor	DALY (age)	DALY (time,age)	
52	0.3	0.5	0.15	1.00	0.15	1.08	0.16	0.16	
53	0.3	1	0.3	0.97	0.29	1.05	0.32	0.31	
54	0.3	1	0.3	0.94	0.28	1.03	0.31	0.29	
55	0.3	1	0.3	0.91	0.27	1.01	0.30	0.28	
56	0.3	1	0.3	0.89	0.27	0.99	0.30	0.26	
57	0.3	1	0.3	0.86	0.26	0.97	0.29	0.25	
58	0.3	1	0.3	0.83	0.25	0.95	0.28	0.24	
59	0.3	1	0.3	0.81	0.24	0.92	0.28	0.22	
60	0.3	1	0.3	0.78	0.24	0.90	0.27	0.21	
61	0.3	1	0.3	0.76	0.23	0.88	0.26	0.20	
62	0.3	0.5	0.15	0.74	0.11	0.86	0.13	0.10	
		YLD = 3	YLD(time) 2.59	YLD(age) 2.90	YLD(t,a) 2.52				
62	0.6	0.5	0.5	0.74	0.37	0.86	0.43	0.32	
63	0.6	1	1	0.72	0.72	0.84	0.84	0.60	
64	0.6	1	1	0.69	0.69	0.82	0.82	0.57	
65	0.6	1	1	0.67	0.67	0.80	0.80	0.54	
66	0.6	1	1	0.65	0.65	0.78	0.78	0.51	
67	0.6	1	1	0.63	0.63	0.76	0.76	0.48	
68	0.6	1	1	0.61	0.61	0.74	0.74	0.46	

69	0.6	1	1	0.60	0.60	0.72	0.72	0.43
70	0.6	1	1	0.58	0.58	0.71	0.71	0.41
71	0.6	1	1	0.56	0.56	0.69	0.69	0.39
72	0.6	1	1	0.54	0.54	0.67	0.67	0.36
73	0.6	1	1	0.53	0.53	0.65	0.65	0.34
74	0.6	1	1	0.51	0.51	0.64	0.64	0.33
75	0.6	1	1	0.50	0.50	0.62	0.62	0.31
76	0.6	1	1	0.48	0.48	0.60	0.60	0.29
77	0.6	0.5	0.5	0.47	0.23	0.59	0.29	0.14

		YLL=	15	YLL(time)	8.88	YLL(age)	10.77	YLL(t,a)	6.47
	Total DALY=YLD+YLL		18		11.47	DALY (age)	13.67	DALY (t,a)	8.98

The process that WHO set in motion has resulted in an extensive and wide ranging debate on the fundamental issues in the selection of DALYs. This debate is too extensive to record here in any detail, rather the main issues and limitations are set out more or less in order of importance to decision makers. The process continues with extensive improvements and, of importance for this study, extensions to the use of DALYs for a variety of situations, anchoring these extensions to a "standardized" concept of health impact for policy.

Treatment of Life Expectancy and Time - Simple mortality measures, such as the deaths per 100,000 person-years, treat each death equally. The most important characteristic of DALYs is that they weight each death by the amount of the life lost. This issue is most important for comparisons with causes with late in life mortality, such as $PM_{2.5}$ health impacts, and causes with early in life mortality, such as traffic accidents. This issue is inherent in the use of DALYs and can only be modified slightly through the use of the age weight.

Mutually Exclusive allocation of morbidity and mortality to "Causes" – DALYs require the allocation of health impacts to one and only one causes. The original approach was disease oriented and extension of DALYs to issues of pollution with impacts on specific populations, requires careful examination of the allocation of health impacts to causes. A pollutant can impact more than one cause and any health impact can be from more than one cause. This is really a specification issue. Part of this issue is also referred to as the attribution issue and a technique called "causal web analysis" is proposed to minimize the estimation errors (Pruss-Ustun, 2003).

Health Preferences may depend on more than Health – DALYs have preferences (DW weights) that were defined to reflect only health status (recent measurement methods included in the WHO version 3 studies provide for including other than health factors). They ignore issues of risk perception (dread, familiarity, voluntary, etc.) as well as the socio-economic status of the population and the specific cultural environment (Reidpath, 2003). Methods such as WTP can overcome this difficulty, if that is desired, since they consider these to the extent that the monetary value developed incorporates these impacts and environmental situations. However, the differences are not large, for example, Hoffstetter (2002), found that the DALYs and WTP estimate for PM mortality differed by a factor of 2, mainly due to large variations in DALY weights associated with severe annoyance and sleep disturbance, causes which people assigned a large weight to in WTP. To a very limited extent the use of the age weight in calculating WHO DALYs reproduces the age effects observed in WTP studies.

Comorbidity – Many people have more than one disease. It is possible for a DALY to be considered for each disease separately and for these persons to have total YLD DALYs in excess of 1 DALY per year. WHO have developed a set of disease specific weights to counteract this effect (Mathers, 2003) This issue also has an impact on the relative priorities assigned to specific causes because of the WHO assumption of measuring DALYs against an ideal state of DW=1.0. It is possible that due to old age alone many people will not achieve an ideal health state and if this is so then the DW weights for older persons should be reduced. Also related to this issue is the difficulty of including in DALYs policy

preferences such as more emphasis on certain socio-economic groups in the population who have low health status.

Population Age Distribution and the Future of Health States – This is a general issue and is common to any life expectancy analysis. The future, if it is like the past, will see longer life expectancies and improved health care interventions that will reduce both morbidity and mortality. These effects will act to over estimate DALYs since they usually represent an extrapolation of the future with today's conditions. A related issue is the difference between the "stable" population distribution used for estimating life tables (e.g. age specific mortality estimates applied to an initial birth cohort) and the actual population age distribution. For many studies the latter is accounted for.

Uncertainty for Low DALY Weights - For many causes such as noise and insomnia, the estimated DALY weights are low but the impacted populations are high. This results in considerable variations in estimated DALYs. For example, Hofstetter (2002) found in a study of PM, ozone, lead, noise, and ozone depletion, in the Netherlands, that for non-mortality DALYs, noise annoyance and sleep disturbance accounted for _ of total YLD (morbidity) DALYs. Both had DW estimates of 0.1 and millions of cases.

DALY Weights – DALY weights (DW) are one of the most discussed issues of DALYs. In comparison to WTP Hammitt (2002), identifies many issues, mainly related to the issues of utility (are the weights representative of individual preferences?) and aggregation (are the individual preferences aggregated in a reasonable way?). The definition of DALYs sidesteps this issue by defining DALY weights as social preferences for purposes of health policy analysis. De Hollander (1999), concludes that "Uncertainty analysis shows that altering weights within the variance seen in most weighting exercises does not substantially affect the overall picture. Compared with the huge uncertainties that are often connected with health impact estimates, the effect of the possible variance in attributed weights appears rather small."

DALYs do not Reflect Health Costs - The costs of hospitals and other medical care are not included in the comparisons since only the health effects are estimated. For policy considerations the costs to the health care system may be significant and vary between causes to the extent that they should be considered. This would have to be done outside the DALY calculation and would depend on the risk management context selected by the decision-maker.

4.4 Approaches to Calculating DALYs

Introduction

There are several approaches to estimating DALYs and these are the result of the historical selection and use of health quality measures. This section is introductory and not intended to be comprehensive. The discussion is organized by specific approaches and issues after this introduction on the history of DALYs selection, measurement and use, which form the "environment" of their development, use and practicality of health quality definition and

measurement for policy purposes.

The purpose of the WHO Global Burden of Disease study is to guide policy and to direct resources towards the causes with the highest health risks. The study has been successful in the sense that the results of DALY estimates for each cause for each region of the world have been used in setting policies over the last few years, the methods and data are continuously being revised and improved, and the method has been extended to new initiatives such as the burden of environmental impacts.

For population DALY studies, such as the WHO Burden of Disease studies, DALYs can be calculated in two main ways: (for convenience **Names of Approach** is given)

1. **Comprehensive Incident, Mortality Data Approach -** Identify the incidence of a "cause" and calculate DALYs from data for each case or incident as illustrated in Figure 1. This requires data on the incidence of a cause, usually for an age group, the length of time each incidence lasts, any mortality due to the cause incident for each age group, estimates of the disability weights (DW), and life expectancy for those who die due to the cause. Causes with multiple episodes such as asthma attacks create some difficulties (de Hollander, 1999).
2. **Population Cause Prevalence and Mortality Data Approach** - By separating the DALY into two components as illustrated in Figure 1 and estimating mortality (YLL) as in approach 1, then estimating YLD (morbidity) from prevalence data. The prevalence of the cause, in the population under study, is estimated for each age group and then the number of persons with the cause are assigned DW years for the interval of the age group to estimate the YLD for the population. The data required is the prevalence for each cause in each age group, and for YLL as in approach 1, the mortality for each cause for each age group and the life expectancy of those who die from the cause in each age group. For example, Manuel (2003) used this approach for Canada's whole population but the results were limited to 10 "causes" (ischemic heart disease, stroke, lung cancer, colorectal cancer, breast cancer, melanoma, diabetes, COPD, osteoarthritis, and mental disorders). DW was estimated by HUI3 and surveys of the population. There was no age weighting or time discounting.

In order to calculate DALYs for the Global Burden of Disease study for whole populations of a region, WHO developed some analytical methods to estimate incidence from prevalence data and also to estimate mortality for age groups. In addition they developed protocols for assignment of deaths to causes, for the use of cancer registries to estimate parameters for DALYs, for estimating missing data for some countries based on other countries in a region, as well as other methods for providing the necessary data for approaches 1 and 2. For example, the program DisMod is a shareware software program developed at Harvard University that allows one to find a set of incidence rates by age that match observed prevalences, given estimates of remission rates and cause-specific mortality risk derived from population data or epidemiological studies. Some of these methods are relatively well developed.

For example all calculations of DALYs (YLD and YLL) presented in the Victoria, Australia, Burden of Disease Study: Morbidity, can be downloaded as self-extracting Zipped Excel and DisMod files. Available at http://www.dhs.vic.gov.au/phd/bod/daly.htm (The self-extracting Zip (.exe) files, each contain a number of Excel and DisMod files. After downloading, double-click the .exe file to extract the compressed files.) Also DisMod software can be downloaded from the Harvard University DisMod website.

If the DALY estimate is not for the whole population of a country or a region or if it is not possible to represent the desired exposure as one of the "causes" (it might be a subset of a cause, or one exposure might lead to a variety of causes, or an exposure might be in addition to other "background" exposures, and so forth) then it is necessary to consider other methods for estimating DALYs. The possible approaches in addition to 1 and 2 include:

3. **Proportional Allocation of WHO Regional BOD results -** For some studies it is possible to proportion out the WHO "cause" estimates for a region to provide the required DALY estimate. For example, the AMR-A, WHO region has a version 3 comprehensive set of estimates for the populations of Canada, the US, and Cuba available on the WHO web site. If the question of interest is, for example, the total chronic mortality health impact of all PM exposures on a specified population then using the methods of de Hollander (1999), it is possible to estimate the proportion of assignable lung cancer deaths and cardiopulmonary mortality deaths due to PM and then by also proportioning the population and assuming that the population is similar to the average in the region and that the concentration of PM is similar, an estimate can be made.

In the case of PM then if the following are known or accepted: (1) assumption of linear health effects for PM, (2) the estimate of average PM exposure for region AMR-A, and (3) the change in PM for the population of concern; then the Proportional Allocation approach can be extended to estimate this situation also.

One advantage of methods that utilize WHO regional data is that the WHO results are available with or without age weights and with or without time weighting.

For other situations where it is not possible to use any combination of approaches 1, 2, or 3, then other methods must be used. These can be identified as:

4. **Estimates from "Reliable" studies using similar assumptions** – If there are estimates from RVIM, or SETAC, based on similar populations and with a process that includes considerable peer review, then their results can be proportioned. This can only be done with care and a number of assumptions. The approach is not comprehensive since it is limited to substances in other studies. The approach also inevitably requires many assumptions about the similarity of populations, health profiles, validity of proportioning exposures (including knowing the background exposures) and so forth.

For example, de Hollander (1999) provides estimates of PM acute health morbidity effects in the Netherlands, with the total YLD (with no age weighting or time discounting) of 2692 years, with a 95% confidence range of 271 to 6808 DALYs. It might be possible to proportion these DALY estimates by population ratios and also the ratio of exposure to get a DALY estimate. The Netherlands estimate was based on data on hospital admissions for respiratory (DW=.64m L=.038 years) and cardiovascular (DW=.71 and L=.038), emergency room visits (DW=.51, L=.033), asthma attacks (DW=.22, L=.005), use of bronchodilators (DW=.22, L=.005), aggravated upper respiratory tract (DW=.05, L=.02), and aggravated lower respiratory tract (DW=.21, L=.04). The approach also needed to estimate the attributable prevalence by age group of each health measurement for acute PM.

5. **Innovative Case Approach**– A variety of assumptions are made and the estimate is based on estimated incident data for the situation. This method will also utilize data or parameters from approaches 1 through 4 usually to provide estimates for DW, L, mortality and life expectancy.
6. **Innovative Exposure Approach** – A variety of assumptions are made and the estimate is based on estimated change in exposure data for the situation. This method will also utilize data or parameters from approaches 1 through 4.

The Importance Principle

The approach used and the effort expended in making a DALY estimate is clearly limited by the data available. Within the available data there is still a need to consider the importance of the DALY estimate and the resources used to make that estimate. This is an analysis method cost-benefit issue. The analysis benefit can be estimated, in this case, as being proportional to the DALY estimate itself. A low DALY indicates a low health impact and a lower importance. Similarly a high DALY indicates a high importance and a justification for more effort and resources in the analysis.

In judging importance it is useful to have confidence limits for DALYs and if these are available then the comparison between high DALYs and low DALYs for substances would use the 5% and 95% confidence estimates respectively in making allocation decisions.

The importance principle is very practical since an iterative approach can be taken with each iteration providing a more complex and hopefully more accurate estimate of DALYs. In each iteration the remaining resources can be allocated according to the DALYs and also the prior expectation of improvement in accuracy. A Bayesian approach may be possible.

The approach is also practical since many of the DALY comparisons involve orders of magnitude differences. For example, de Hollander found that for chronic PM the DALY estimate was in excess of 100,000 years for the Netherlands while the acute PM estimate was about 3,000. Any available resources should be allocated first to the estimate of chronic DALYs. For example in HEIDI II the acute PM health effects are not estimated.

A Simplified Approach for Toxics (SETAC)

The simplified approach of SETAC for toxic effects are generally defined for a variety of end points which may not necessarily correspond to health effects end points for CACs and carcinogens. SETAC (Owen, 2002) faced this same difficulty and have the same objective as our study (i.e. characterizing chronic non-cancer toxicity with a view to establishing a screening indicator for organizing and aggregating information in order to provide meaningful direction for further policy analysis). Their solution as documented by Owens (2002) can be extended to provide a comparable estimate. It is noted that the SETAC approach case study covers many of the emissions on our list but there will be a need to provide estimates for missing data. The procedure proposed by SETAC is:

1. The procedure is a subjective scoring exercise, not a scientific or technical operation, but it does use the original toxicity data in an attempt to avoid hidden weighting and valuation schemes (e.g. ADI and RfD are not used) as a substitute for scientific characterization. When available, the toxicological ED05 or ED10 levels will be used to estimate toxicity. The method was developed for the case of an interrelated industrial system with environmental emissions as a focus and an objective of "identify and prioritize potentially important emissions and to facilitate risk assessment, including comparative risk assessment".

2. There were three classes of severity established: 1) irreversible/life shortening, 2) may be reversible, could be life shortening, and 3) generally reversible, generally not life shortening. The WHO uses general estimates of DALYs for these three broad classes of endpoints -- 6.7, .67, and .067 respectively (Pennington, 2002). For example, category 1 includes cancers and the DALY weighting of 6.7 is the average in Table 2 for all cancers considered by WHO.

3. Toxicity is estimated for the major organs targeted by the toxic effects. For example, Table 4 is reproduced from Owens (2000; Table 3 p.).

Table 4 Impact Categories for proposed classification scheme (after Owens, 2002)

Classification Category	Description and Examples of Toxicity
Systemic toxicity	Significant effects on the whole animal, Examples are excess mortality due to chemical or decrease in body weight gain during the study
Hepatotoxicity	Effects on the liver. Examples are an increase in the liver weight versus controls or adverse change in liver cell histopathology
Nephrotoxicity	Effects on the kidney. Increase in the kidney weight versus controls or adverse change in histopathology of kidney cells.
Neurotoxicity	Effects on the nervous system. Evidence of neurological dysfunction, e.g., tremors, decrease in brain weight, change in central or peripheral nervous system histopathology.
Reproductive	Effect on the reproductive capacity of the parental generation. May be both or either sex (male effects or female effects). Examples are changes in reproductive organ weights, estrus cycle, sperm counts, etc.
Teratogenic	Evidence of birth defects or malformations in offspring when maternal exposure occurs during pregnancy. Typically dosed after implementation is expected from mating time through pregnancy
Developmental	Evidence of effects on developmental process. Effects are not manifest immediately by later, such as in adulthood from exposures in utero or in early postnatal development.
Pulmonary and cardiac	Effects on either the lung or lung function or the heart or heart function
Immunotoxicity	Effects on the immune system. Examples are changes in defined immune response to infectious disease, inability to target malignant cells, etc.
Hemopoietic	Effects on bone marrow generation of blood cells and related cells. Change in circulating red blood cells (e.g., hemoglobin level or red blood cell number), platelets, or bone marrow pathology
Other Toxicities	Effects on organ or organ system not listed in the classification proposal.
Irritation, sensitivity, inflammation	Irritation or hyperplasia of epithelial or mucosal surface (e.g., nose, skin, fore stomach) in contact with chemical or invasive response by immune cells.
Questions	Difficulty in making a classification, e.g., several simultaneous effects at a dose where choices and judgements may be employed or the effect described is not clear.

5. Uncertainties

"Extreme caution is advocated when comparing the likelihood and potential consequence estimates across chemical emissions in a LCA study, particularly between noncancer and cancer effects results. These estimates provide preliminary, or screening-level insights only. While the presented framework for the calculation of LCA characterization factors allows for the consideration of nonlinear low-dose-response curves, mechanistic thresholds, and multiple background exposure concentrations, the availability of required data is limited in practice." (Pennington, 2002)

This quote highlights the uncertainties involved in making comparisons. It should be noted that in HEIDI 1 we selected the Mantel-Bryan formulation for the concentration-response curve, perhaps the most uncertain component of the overall process. This effectively deals with some of the disparities that surround the use of the concentration-response curves, but the basic impact of uncertainties on health impact comparisons is still valid.

Within the separate classes of air emissions, the usual assumption is that the biasing influences of uncertainties in producing over-estimates and under-estimates of actual health effects are broadly consistent across all of the substances under examination. This makes possible an ordinal-quality (rank-order) estimate of health effects, which is valid in a relative sense but not an absolute sense. Therefore, it is possible to set relative priorities within a class despite major uncertainties and various sources of bias in the estimation procedure. However, when comparing between the 3 basic classes of air emissions, even for relative rankings of health effects there is a larger possibility that the biases and uncertainties will introduce systematic errors in comparisons. This is the main argument for the difficulty of comparisons. It is noted that the uncertainties are typically several orders of magnitude.

One exception to this difficulty of interclass comparisons, is the comparison of class 2 (toxics) with class 1 (carcinogenic air toxics) for a small number of emissions, such as benzene where there is data available from both epidemiology studies and the usual C-R slope estimates from animal studies. For example, the workers in the Pliofilm factory who were exposed to very high levels of benzene and had a 3% increase in cancers can be compared to the typical estimate from the toxicological C-R slopes.

The structure of HEIDI II is to have, in module 2, an estimate of the number of cases for each substance and then in module 3 estimate the DALYs per case. This is proposed to be done by considering each substance in turn and then using one of the 6 approaches identified above estimating the DALYs through some set of assumptions. For substances where the importance is low, even for the complex DALY process the simple DALY estimates will be used. - there are only 29 substances to select simple and complex DALYs for. The simple DALYs for most substances will be pretty straightforward. As for the complex DALY's they will be available for most of the cancers and for the health outcomes that the CACs are associated with, and it will be a matter of selecting which is the closest fit from the WHO diseases for the non-cancer effects.

References

Crettaz, P. et al. 2002. Assessing Human Health Response in Life Cycle Assessment Using ED10s and DALYs: Part 1 - Cancer Effects, *Risk Analysis*, 22, 5:931-946.

De Hollander, A.E.M., Melse J.M., Lebret, E., and Kramers, P.G.N. 1999. An Aggregate Public Health Indicator to Represent the Impact of Multiple Environmental Exposures, *Epidemiology,* 10, 5:606-617.

Evans, J., Levy, J., Hammitt, J., Santos Burgoa, C., and Castellejos, M. 2002. Chapter 4, Health Benefits of Air Pollution Control, in (Eds.) L. T. Molina and M.J. Molina Air *Quality in the Mexico Megacity: An Integrated Assessment.* pp. 105-136. Dordrecht: Kluwer Academic Publishers.

Gold, M.R., Stevenson, D., and Fryback, D.G. 2002. HALYs and QALYs, and DALYs, Oh My: Similarities and Differences in Summary Measures of Population Health, *Annu. Rev. Public Health* 23:115-34.

Hammitt, J.K. 2002. QUALYs Versus WTP, *Risk Analysis*, 22, 5:985-1001.

Hofstteter, P and Hammitt J.K., 2002. Selecting Human Health Metrics for Environmental Decision-Support Tools. *Risk Analysis*. 22, 5:965-982.

Manuel, D., Wei Luo, C., Ugnat A.-M., and Mao, Y. 2003.Cause-deleted health-adjusted life expectancy of Canadians with selected chronic conditions, Chronic Diseases in Canada, 24, 4, Fall 2003, Ottawa.

Murray, C.J., and Lopez, A.D. 1996.The Global Burden of Disease, WHO, Geneva, also Harvard School of Public Health and World Bank.

Murray, C.J.L., Salomon, J.A., and Mathers, C. 1999. A Critical Examination of Summary Measures of Population Health, WHO Discussion Paper number 2 Discussion Paper No. 54, WHO Global Program on Evidence for Health Policy, Geneva. (http://www3.who.int/whosis/discussion_papers/discussion_papers.cfm?path).

Newhook, R. 2004. Personal communication.

Owens, J.W., 2002. Chemical Toxicity Indicators for Human Health: Case study for Classification of Chronic Noncancer Chemical Hazards in Life-Cycle Assessment, *Environmental Toxicology and Chemistry*, 21, 1:207-225.

Pennington, D. et al. 2002. Assessing Human Health Response in Life Cycle Assessment Using ED10s and DALYs: Part 2 - Noncancer Effects, *Risk Analysis*, 22, 5:947-963.

Reidpath, Daniel D., Allotey, P.A., Kouame, A., and Cummins, R.A. 2003. Measuring health in a vacuum: examining the disability weight of the DALY, *Health Policy and Planning* 18, 4:351-356.

Torrance, G., Furlong, W., Feeney, D., Boyle, M. 1995. Multi-Attribute preference functions – Health Utility Index, *Pharmacoeconomics*, 7:503-520

WHO, 2001.National Burden of Disease Studies: A Practical Guide, Edition 2.0, WHO Global Program on Evidence for Health Policy, Geneva. (http://www3.who.int/whosis/menu.cfm?path=evidence,burden,burden_gbd2000docs&language=english – select "guidelines")

WHO, Pruss-Ustun, A., Colin Mathers, Carlos Corvalan, and Alistair Woodward, 2003. Assessing the environmental burden of disease at national and local levels: Introduction and methods, Environmental Burden of Disease Series, No. 1. Geneva.

WHO, Colin Mathers et al., 2003. Global Burden of Disease in 2002: data sources, methods and results, Discussion Paper No. 54, WHO Global Program on Evidence for Health Policy, Geneva. (http://www3.who.int/whosis/discussion_papers/discussion_papers.cfm?path).