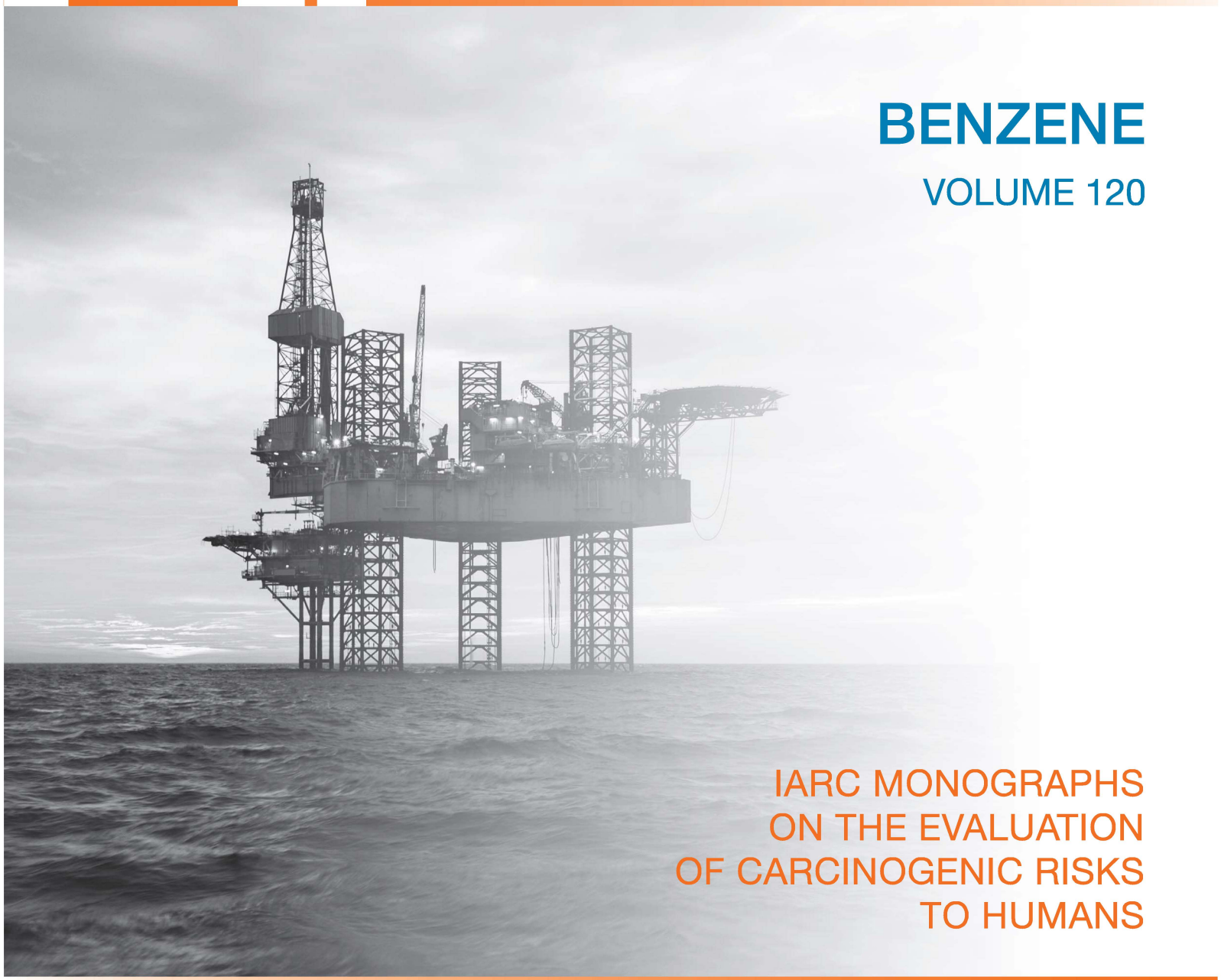
IARC MONOGRAPHS



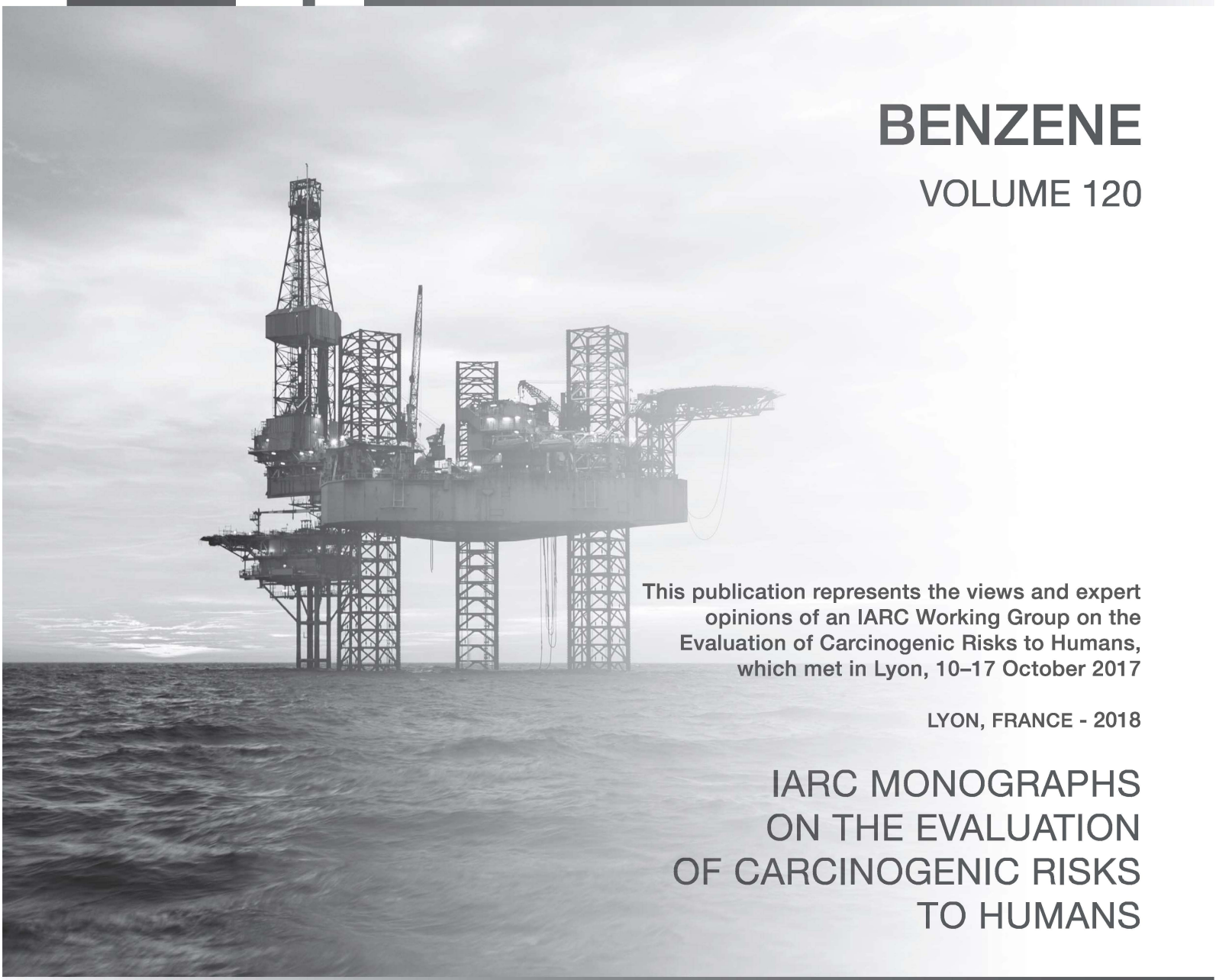
BENZENE

VOLUME 120

IARC MONOGRAPHS
ON THE EVALUATION
OF CARCINOGENIC RISKS
TO HUMANS

International Agency for Research on Cancer
World Health Organization

IARC MONOGRAPHS



BENZENE

VOLUME 120

This publication represents the views and expert opinions of an IARC Working Group on the Evaluation of Carcinogenic Risks to Humans, which met in Lyon, 10–17 October 2017

LYON, FRANCE - 2018

IARC MONOGRAPHS ON THE EVALUATION OF CARCINOGENIC RISKS TO HUMANS

International Agency for Research on Cancer
World Health Organization

IARC MONOGRAPHS

In 1969, the International Agency for Research on Cancer (IARC) initiated a programme on the evaluation of the carcinogenic risk of chemicals to humans involving the production of critically evaluated monographs on individual chemicals. The programme was subsequently expanded to include evaluations of carcinogenic risks associated with exposures to complex mixtures, lifestyle factors and biological and physical agents, as well as those in specific occupations. The objective of the programme is to elaborate and publish in the form of monographs critical reviews of data on carcinogenicity for agents to which humans are known to be exposed and on specific exposure situations; to evaluate these data in terms of human risk with the help of international working groups of experts in carcinogenesis and related fields; and to indicate where additional research efforts are needed. The lists of IARC evaluations are regularly updated and are available on the Internet at http://monographs.iarc.fr/.

This programme has been supported since 1982 by Cooperative Agreement U01 CA33193 with the United States National Cancer Institute, Department of Health and Human Services. Additional support has been provided since 1986 by the European Commission Directorate-General for Employment, Social Affairs, and Inclusion, initially by the Unit of Health, Safety and Hygiene at Work, and since 2014 by the European Union Programme for Employment and Social Innovation "EaSI" (2014–2020) (for further information please consult: http://ec.europa.eu/social/easi). Support has also been provided since 1992 by the United States National Institute of Environmental Health Sciences, Department of Health and Human Services. The contents of this volume are solely the responsibility of the Working Group and do not necessarily represent the official views of the United States National Cancer Institute, the United States National Institute of Environmental Health Sciences, the United States Department of Health and Human Services, or the European Commission.

Published by the International Agency for Research on Cancer,
150 cours Albert Thomas, 69372 Lyon Cedex 08, France

On-line publication, December 2018

Distributed by WHO Press, World Health Organization, 20 Avenue Appia, 1211 Geneva 27, Switzerland
(tel.: +41 22 791 3264; fax: +41 22 791 4857; email: bookorders@who.int).

Corrigenda to the IARC Monographs are published online at http://monographs.iarc.fr/ENG/Publications/corrigenda.php
To report an error, please contact: editimo@iarc.fr

Co-funded by the European Union

IARC Library Cataloguing in Publication Data

Benzene / IARC Working Group on the Evaluation of Carcinogenic Risks to Humans (2017: Lyon, France)

(IARC monographs on the evaluation of carcinogenic risks to humans ; volume 120)

1. Carcinogens 2. Benzene – adverse effects 3. Occupational Exposure – adverse effects 4. Risk Factors

I. International Agency for Research on Cancer II. Series

ISBN 978-92-832-0187-8 (NLM Classification: W1)
ISSN 1017-1606

CONTENTS

NOTE TO THE READER

The term ‘carcinogenic risk’ in the *IARC Monographs* series is taken to mean that an agent is capable of causing cancer. The *Monographs* evaluate cancer hazards, despite the historical presence of the word ‘risks’ in the title.

Inclusion of an agent in the *Monographs* does not imply that it is a carcinogen, only that the published data have been examined. Equally, the fact that an agent has not yet been evaluated in a *Monograph* does not mean that it is not carcinogenic. Similarly, identification of cancer sites with *sufficient evidence* or *limited evidence* in humans should not be viewed as precluding the possibility that an agent may cause cancer at other sites.

The evaluations of carcinogenic risk are made by international working groups of independent scientists and are qualitative in nature. No recommendation is given for regulation or legislation.

Anyone who is aware of published data that may alter the evaluation of the carcinogenic risk of an agent to humans is encouraged to make this information available to the IARC Monographs Group, International Agency for Research on Cancer, 150 cours Albert Thomas, 69372 Lyon Cedex 08, France, in order that the agent may be considered for re-evaluation by a future Working Group.

Although every effort is made to prepare the *Monographs* as accurately as possible, mistakes may occur. Readers are requested to communicate any errors to the IARC Monographs Group, so that corrections can be reported in future volumes.

LIST OF PARTICIPANTS

Members [1]

Roberta Andreoli

Department of Medicine and Surgery
University of Parma
Parma
Italy

Fiorella Belpoggi (Subgroup Chair, Cancer in Experimental Animals)

Research Department
Cesare Maltoni Cancer Research Centre
Ramazzini Institute
Bologna
Italy

Weihsueh A. Chiu (Subgroup Chair, Mechanisms)

Department of Veterinary Integrative Biosciences
College of Veterinary Medicine
Texas A&M University
College Station, TX
USA

John E. French

Michael Hooker Research Center
University of North Carolina at Chapel Hill
Chapel Hill, NC
USA

[1] Working Group Members and Invited Specialists serve in their individual capacities as scientists and not as representatives of their government or any organization with which they are affiliated. Affiliations are provided for identification purposes only. Invited Specialists do not serve as Meeting Chair or Subgroup Chair, draft text that pertains to the description or interpretation of cancer data, or participate in the evaluations.
Each participant was asked to disclose pertinent research, employment, and financial interests. Current financial interests and research and employment interests during the past 4 years or anticipated in the future are identified here. Minor pertinent interests are not listed and include stock valued at no more than US$ 1000 overall, grants that provide no more than 5% of the research budget of the expert's organization and that do not support the expert's research or position, and consulting or speaking on matters not before a court or government agency that does not exceed 2% of total professional time or compensation. All grants that support the expert's research or position and all consulting or speaking on behalf of an interested party on matters before a court or government agency are listed as significant pertinent interests.

Silvia Fustinoni

Department of Clinical Sciences and Community Health
University of Milan
Milan
Italy

Pascal Guénel

Center for Research in Epidemiology and Population Health (CESP)
INSERM
Villejuif
France

Yoko Hirabayashi

Division of Cellular and Molecular Toxicology
Center for Biological Safety and Research
National Institute of Health Sciences
Kawasaki
Japan

Nancy B. Hopf (Subgroup Chair, Exposure Data)

Institute for Work and Health (IST)
Epalinges-Lausanne
Switzerland

James Huff [did not attend]

National Institute of Environmental Health Sciences
Research Triangle Park, NC
USA

Jennifer Jinot [retired]

National Center for Environmental Assessment
United States Environmental Protection Agency
Washington, DC
USA

Jorunn Kirkeleit

Department of Global Public Health and Primary Care
University of Bergen
Bergen
Norway

Dong-Hee Koh

International St Mary's Hospital
Catholic Kwandong University
Gangneung-si
Republic of Korea

Igor Linhart

Department of Organic Chemistry
Faculty of Chemical Technology
University of Chemistry and Technology
Prague
Czechia

Ole Raaschou-Nielsen

Department of Nutrition, Genes and the Environment
Danish Cancer Society Research Center
Copenhagen
Denmark

David M. Reif

Center for Human Health and the Environment
Department of Biological Sciences
North Carolina State University
Raleigh, NC
USA

David B. Richardson

Department of Epidemiology
School of Public Health
University of North Carolina at Chapel Hill
Chapel Hill, NC
USA

David Ross

Department of Pharmaceutical Sciences
Skaggs School of Pharmacy and Pharmaceutical Sciences
University of Colorado
Aurora, CO
USA

Nathaniel Rothman

Occupational and Environmental Epidemiology Branch
Division of Cancer Epidemiology and Genetics
National Cancer Institute
Rockville, MD
USA

Keith Salazar

National Center for Environmental Assessment
United States Environmental Protection Agency
Washington, DC
USA

Andreas Seidler

Institute and Polyclinic for Work and Social Medicine
Dresden Technical University
Dresden
Germany

Sharon Silver

Centers for Disease Control and Prevention
National Institute for Occupational Safety and Health
Cincinnati, OH
USA

Elaine Symanski

Department of Epidemiology, Human Genetics and Environmental Sciences
University of Texas Health Science Center (UTHealth) School of Public Health
Houston, TX
USA

Reuben Thomas

J. David Gladstone Institutes
University of California San Francisco
San Francisco, CA
USA

Hiroyuki Tsuda

Nanotoxicology Project Laboratory
Nagoya City University
Nagoya
Japan

Roel Vermeulen (Overall Chair)

Division of Environmental Epidemiology
Institute for Risk Assessment Sciences (IRAS)
University of Utrecht
Utrecht
The Netherlands

Paolo Vineis (Subgroup Chair, Cancer in Humans)

Department of Environmental Epidemiology
Imperial College London
London
United Kingdom

Luoping Zhang

School of Public Health
University of California, Berkeley
Berkeley, CA
USA

Invited Specialist

Deborah Glass [2]

Monash Centre for Occupational and Environmental Health (MonCOEH)
Monash University
Melbourne
Australia

Representatives

Roberto Flores-Munguia [3]

Nutritional Sciences Research Group
Division of Cancer Prevention
National Cancer Institute
Bethesda, MD
USA

Ariane Leites Larentis [4]

Oswaldo Cruz Foundation (Fiocruz)
National School of Public Health Sergio Arouca (ENSP)
Research Center for Worker´s Health and Human Ecology (CESTEH)
Rio de Janeiro
Brazil

Marcia Sarpa de Campos Mello [5]

Occupational and Environmental Exposure and Cancer Unit
Prevention and Surveillance Coordination
National Cancer Institute (INCA)
Rio de Janeiro
Brazil

Eun Young Park [6]

National Cancer Center
Republic of Korea

[2] Deborah Glass receives significant research funding from the Australian Institute of Petroleum to Monash University.
[3] Roberto Flores-Munguia attended as a Representative of the National Cancer Institute, USA.
[4] Ariane Leites Larentis attended as a Representative of the Oswaldo Cruz Foundation, Brazil.
[5] Marcia Sarpa de Campos Mello attended as a Representative of the National Cancer Institute, Brazil.
[6] Eun Young Park attended as a Representative of the National Cancer Center, Republic of Korea.

Min Kyung Lim [7]

National Cancer Center
Republic of Korea

Observers [8]

Richard H. Adamson [9]

TPN Associates, L.L.C.
USA

Peter J. Boogaard [10]

The Netherlands

Alex Forrest [11]

International Association of Fire Fighters
Winnipeg, Manitoba
Canada

Tom K. Grimsrud [12]

Cancer Registry of Norway
Oslo
Norway

Sean Hays [13]

Summit Toxicology, L.L.P.
Bozeman, MT
USA

Peter F. Infante [14]

Peter F. Infante Consulting, L.L.C.
Falls Church, VA
USA

Veli-Pekka Parkkinen [15]

Department of Philosophy
University of Kent
Canterbury
United Kingdom

[7] Min Kyung Lim attended as a Representative of the National Cancer Center, Republic of Korea.
[8] Each Observer agreed to respect the Guidelines for Observers at *IARC Monographs* meetings. Observers did not serve as Meeting Chair or Subgroup Chair, draft any part of a *Monograph*, or participate in the evaluations. They also agreed not to contact participants before the meeting, not to lobby them at any time, not to send them written materials, and not to offer them meals or other favours. IARC asked and reminded Working Group Members to report any contact or attempt to influence that they may have encountered, either before or during the meeting.
[9] Richard H. Adamson attends as an Observer for the American Beverage Association.
[10] Peter J. Boogaard is employed by Shell International and is also a participant of the Global Employee Share Purchase Programme of Royal Dutch Shell. He attended as an Observer for the Aromatics Producers Association, Belgium.
[11] Alex Forrest attended as an Observer for the Manitoba Professional Firefighters Association, Canada.
[12] Tom K. Grimsrud attended as an Observer for the Cancer Registry of Norway.
[13] Sean Hays consulted for a confidential oil company and consults for a confidential tobacco company (amount not disclosed). His company, Summit Toxicology, serves in litigation involving petroleum products and benzene. He attended as an Observer for the American Petroleum Institute.
[14] Peter F. Infante has served as an expert witness for plaintiffs in litigation involving benzene. He attended as an Observer for Peter F. Infante Consulting, L.L.C.
[15] Veli-Pekka Parkkinen attended as an Observer for the University of Kent, United Kingdom.

Christian Wallmann [16]

School of European Culture and Languages
Cornwallis North West
University of Kent
Canterbury
United Kingdom

IARC/WHO Secretariat

Lamia Benbrahim-Tallaa *(Rapporteur, Mechanistic and Other Relevant Data)*
Véronique Bouvard *(Rapporteur, Exposure Data)*
Rafael Carel, Visiting Scientist, Ben Gurion University, Israel
Fatiha El Ghissassi *(Rapporteur, Mechanistic and Other Relevant Data)*
Yann Grosse *(Rapporteur, Cancer in Experimental Animals)*
Neela Guha *(Rapporteur, Cancer in Humans)*
Kathryn Guyton *(Rapporteur, Mechanistic and Other Relevant Data)*
Amy Hall
Rim Khlifi
Dana Loomis *(Responsible Officer, Group Head)*
Heidi Mattock *(Scientific Editor)*
Kurt Straif *(Section Head)*
Nadia Vilahur

Administrative Assistance

Marieke Dusenberg
Sandrine Egraz
Michel Javin
Helene Lorenzen-Augros
Lucy Shedden

Production Team

Elisabeth Elbers
Fiona Gould
Solène Quennehen

Post-Meeting Assistance

Elaine Rowan *(Technical Editor)*
Karen Müller *(Scientific Editor)*

Post-Meeting Scientific Assistance

Amy Hall
Mary Schubauer-Berigan
Eero Suonio

[16] Christian Wallmann attended as an Observer for the University of Kent, United Kingdom.

PREAMBLE

The Preamble to the *IARC Monographs* describes the objective and scope of the programme, the scientific principles and procedures used in developing a *Monograph*, the types of evidence considered and the scientific criteria that guide the evaluations. The Preamble should be consulted when reading a *Monograph* or list of evaluations.

A. GENERAL PRINCIPLES AND PROCEDURES

1. Background

Soon after IARC was established in 1965, it received frequent requests for advice on the carcinogenic risk of chemicals, including requests for lists of known and suspected human carcinogens. It was clear that it would not be a simple task to summarize adequately the complexity of the information that was available, and IARC began to consider means of obtaining international expert opinion on this topic. In 1970, the IARC Advisory Committee on Environmental Carcinogenesis recommended '... that a compendium on carcinogenic chemicals be prepared by experts. The biological activity and evaluation of practical importance to public health should be referenced and documented.' The IARC Governing Council adopted a resolution concerning the role of IARC in providing government authorities with expert, independent, scientific opinion on environmental carcinogenesis. As one means to that end, the Governing Council recommended that IARC should prepare monographs on the evaluation of carcinogenic risk of chemicals to man, which became the initial title of the series.

In the succeeding years, the scope of the programme broadened as *Monographs* were developed for groups of related chemicals, complex mixtures, occupational exposures, physical and biological agents and lifestyle factors. In 1988, the phrase 'of chemicals' was dropped from the title, which assumed its present form, *IARC Monographs on the Evaluation of Carcinogenic Risks to Humans*.

Through the *Monographs* programme, IARC seeks to identify the causes of human cancer. This is the first step in cancer prevention, which is needed as much today as when IARC was established. The global burden of cancer is high and continues to increase: the annual number of new cases was estimated at 10.1 million in 2000 and is expected to reach 15 million by 2020 (Stewart & Kleihues, 2003). With current trends in demographics and exposure, the cancer burden has been shifting from high-resource countries to low- and medium-resource countries. As a result of *Monographs* evaluations, national health agencies have been able, on scientific grounds, to take measures to reduce human exposure to carcinogens in the workplace and in the environment.

The criteria established in 1971 to evaluate carcinogenic risks to humans were adopted by the Working Groups whose deliberations resulted in the first 16 volumes of the *Monographs* series. Those criteria were subsequently updated by further ad hoc Advisory Groups (IARC, 1977, 1978, 1979, 1982, 1983, 1987, 1988, 1991; Vainio et al., 1992; IARC, 2005, 2006).

The Preamble is primarily a statement of scientific principles, rather than a specification of working procedures. The procedures through which a Working Group implements these principles are not specified in detail. They usually involve operations that have been established as being effective during previous *Monograph* meetings but remain, predominantly, the prerogative of each individual Working Group.

2. Objective and scope

The objective of the programme is to prepare, with the help of international Working Groups of experts, and to publish in the form of *Monographs*, critical reviews and evaluations of evidence on the carcinogenicity of a wide range of human exposures. The *Monographs* represent the first step in carcinogen risk assessment, which involves examination of all relevant information to assess the strength of the available evidence that an agent could alter the age-specific incidence of cancer in humans. The *Monographs* may also indicate where additional research efforts are needed, specifically when data immediately relevant to an evaluation are not available.

In this Preamble, the term 'agent' refers to any entity or circumstance that is subject to evaluation in a *Monograph*. As the scope of the programme has broadened, categories of agents now include specific chemicals, groups of related chemicals, complex mixtures, occupational or environmental exposures, cultural or behavioural practices, biological organisms and physical agents. This list of categories may expand as causation of, and susceptibility to, malignant disease become more fully understood.

A cancer 'hazard' is an agent that is capable of causing cancer under some circumstances, while a cancer 'risk' is an estimate of the carcinogenic effects expected from exposure to a cancer hazard. The *Monographs* are an exercise in evaluating cancer hazards, despite the historical presence of the word 'risks' in the title. The distinction between hazard and risk is important, and the *Monographs* identify cancer hazards even when risks are very low at current exposure levels, because new uses or unforeseen exposures could engender risks that are significantly higher.

In the *Monographs*, an agent is termed 'carcinogenic' if it is capable of increasing the incidence of malignant neoplasms, reducing their latency, or increasing their severity or multiplicity. The induction of benign neoplasms may in some circumstances (see Part B, Section 3a) contribute to the judgement that the agent is carcinogenic. The terms 'neoplasm' and 'tumour' are used interchangeably.

The Preamble continues the previous usage of the phrase 'strength of evidence' as a matter of historical continuity, although it should be understood that *Monographs* evaluations consider studies that support a finding of a cancer hazard as well as studies that do not.

Some epidemiological and experimental studies indicate that different agents may act at different stages in the carcinogenic process, and several different mechanisms may be involved. The aim of the *Monographs* has been, from their inception, to evaluate evidence of carcinogenicity at any stage in the carcinogenesis process, independently of the underlying mechanisms. Information on mechanisms may, however, be used in making the overall evaluation (IARC, 1991; Vainio et al., 1992; IARC, 2005, 2006; see also Part B, Sections 4 and 6). As mechanisms of carcinogenesis are elucidated, IARC convenes international scientific conferences to determine whether a broad-based consensus has emerged

on how specific mechanistic data can be used in an evaluation of human carcinogenicity. The results of such conferences are reported in IARC Scientific Publications, which, as long as they still reflect the current state of scientific knowledge, may guide subsequent Working Groups.

Although the *Monographs* have emphasized hazard identification, important issues may also involve dose–response assessment. In many cases, the same epidemiological and experimental studies used to evaluate a cancer hazard can also be used to estimate a dose–response relationship. A *Monograph* may undertake to estimate dose–response relationships within the range of the available epidemiological data, or it may compare the dose–response information from experimental and epidemiological studies. In some cases, a subsequent publication may be prepared by a separate Working Group with expertise in quantitative dose–response assessment.

The *Monographs* are used by national and international authorities to make risk assessments, formulate decisions concerning preventive measures, provide effective cancer control programmes and decide among alternative options for public health decisions. The evaluations of IARC Working Groups are scientific, qualitative judgements on the evidence for or against carcinogenicity provided by the available data. These evaluations represent only one part of the body of information on which public health decisions may be based. Public health options vary from one situation to another and from country to country and relate to many factors, including different socioeconomic and national priorities. Therefore, no recommendation is given with regard to regulation or legislation, which are the responsibility of individual governments or other international organizations.

3. Selection of agents for review

Agents are selected for review on the basis of two main criteria: (a) there is evidence of human exposure and (b) there is some evidence or suspicion of carcinogenicity. Mixed exposures may occur in occupational and environmental settings and as a result of individual and cultural habits (such as tobacco smoking and dietary practices). Chemical analogues and compounds with biological or physical characteristics similar to those of suspected carcinogens may also be considered, even in the absence of data on a possible carcinogenic effect in humans or experimental animals.

The scientific literature is surveyed for published data relevant to an assessment of carcinogenicity. Ad hoc Advisory Groups convened by IARC in 1984, 1989, 1991, 1993, 1998 and 2003 made recommendations as to which agents should be evaluated in the *Monographs* series. Recent recommendations are available on the *Monographs* programme web site (http://monographs.iarc.fr). IARC may schedule other agents for review as it becomes aware of new scientific information or as national health agencies identify an urgent public health need related to cancer.

As significant new data become available on an agent for which a *Monograph* exists, a re-evaluation may be made at a subsequent meeting, and a new *Monograph* published. In some cases it may be appropriate to review only the data published since a prior evaluation. This can be useful for updating a database, reviewing new data to resolve a previously open question or identifying new tumour sites associated with a carcinogenic agent. Major changes in an evaluation (e.g. a new classification in Group 1 or a determination that a mechanism does not operate in humans, see Part B, Section 6) are more appropriately addressed by a full review.

4. Data for the *Monographs*

Each *Monograph* reviews all pertinent epidemiological studies and cancer bioassays in experimental animals. Those judged inadequate or irrelevant to the evaluation may be cited but not summarized. If a group of similar studies is not reviewed, the reasons are indicated.

Mechanistic and other relevant data are also reviewed. A *Monograph* does not necessarily cite all the mechanistic literature concerning the agent being evaluated (see Part B, Section 4). Only those data considered by the Working Group to be relevant to making the evaluation are included.

With regard to epidemiological studies, cancer bioassays, and mechanistic and other relevant data, only reports that have been published or accepted for publication in the openly available scientific literature are reviewed. The same publication requirement applies to studies originating from IARC, including meta-analyses or pooled analyses commissioned by IARC in advance of a meeting (see Part B, Section 2c). Data from government agency reports that are publicly available are also considered. Exceptionally, doctoral theses and other material that are in their final form and publicly available may be reviewed.

Exposure data and other information on an agent under consideration are also reviewed. In the sections on chemical and physical properties, on analysis, on production and use and on occurrence, published and unpublished sources of information may be considered.

Inclusion of a study does not imply acceptance of the adequacy of the study design or of the analysis and interpretation of the results, and limitations are clearly outlined in square brackets at the end of each study description (see Part B). The reasons for not giving further consideration to an individual study also are indicated in the square brackets.

5. Meeting participants

Five categories of participant can be present at *Monograph* meetings.

(a) The Working Group

The Working Group is responsible for the critical reviews and evaluations that are developed during the meeting. The tasks of Working Group Members are: (i) to ascertain that all appropriate data have been collected; (ii) to select the data relevant for the evaluation on the basis of scientific merit; (iii) to prepare accurate summaries of the data to enable the reader to follow the reasoning of the Working Group; (iv) to evaluate the results of epidemiological and experimental studies on cancer; (v) to evaluate data relevant to the understanding of mechanisms of carcinogenesis; and (vi) to make an overall evaluation of the carcinogenicity of the exposure to humans. Working Group Members generally have published significant research related to the carcinogenicity of the agents being reviewed, and IARC uses literature searches to identify most experts. Working Group Members are selected on the basis of (a) knowledge and experience and (b) absence of real or apparent conflicts of interests. Consideration is also given to demographic diversity and balance of scientific findings and views.

(b) Invited Specialists

Invited Specialists are experts who also have critical knowledge and experience but have a real or apparent conflict of interests. These experts are invited when necessary to assist in the Working Group by contributing their unique knowledge and experience during subgroup and plenary discussions. They may also contribute text on non-influential issues in the section on exposure, such as a general description of data on production and use (see Part B, Section 1). Invited Specialists do not serve as meeting chair

or subgroup chair, draft text that pertains to the description or interpretation of cancer data, or participate in the evaluations.

(c) Representatives of national and international health agencies

Representatives of national and international health agencies often attend meetings because their agencies sponsor the programme or are interested in the subject of a meeting. Representatives do not serve as meeting chair or subgroup chair, draft any part of a *Monograph,* or participate in the evaluations.

(d) Observers with relevant scientific credentials

Observers with relevant scientific credentials may be admitted to a meeting by IARC in limited numbers. Attention will be given to achieving a balance of Observers from constituencies with differing perspectives. They are invited to observe the meeting and should not attempt to influence it. Observers do not serve as meeting chair or subgroup chair, draft any part of a *Monograph*, or participate in the evaluations. At the meeting, the meeting chair and subgroup chairs may grant Observers an opportunity to speak, generally after they have observed a discussion. Observers agree to respect the Guidelines for Observers at *IARC Monographs* meetings (available at http://monographs.iarc.fr).

(e) The IARC Secretariat

The IARC Secretariat consists of scientists who are designated by IARC and who have relevant expertise. They serve as rapporteurs and participate in all discussions. When requested by the meeting chair or subgroup chair, they may also draft text or prepare tables and analyses.

Before an invitation is extended, each potential participant, including the IARC Secretariat, completes the WHO Declaration of Interests to report financial interests, employment and consulting, and individual and institutional research support related to the subject of the meeting. IARC assesses these interests to determine whether there is a conflict that warrants some limitation on participation. The declarations are updated and reviewed again at the opening of the meeting. Interests related to the subject of the meeting are disclosed to the meeting participants and in the published volume (Cogliano et al., 2004).

The names and principal affiliations of participants are available on the *Monographs* programme web site (http://monographs.iarc.fr) approximately two months before each meeting. It is not acceptable for Observers or third parties to contact other participants before a meeting or to lobby them at any time. Meeting participants are asked to report all such contacts to IARC (Cogliano et al., 2005).

All participants are listed, with their principal affiliations, at the beginning of each volume. Each participant who is a Member of a Working Group serves as an individual scientist and not as a representative of any organization, government or industry.

6. Working procedures

A separate Working Group is responsible for developing each volume of *Monographs*. A volume contains one or more *Monographs*, which can cover either a single agent or several related agents. Approximately one year in advance of the meeting of a Working Group, the agents to be reviewed are announced on the *Monographs* programme web site (http://monographs.iarc.fr) and participants are selected by IARC staff in consultation with other experts. Subsequently, relevant biological and epidemiological data are collected by IARC from recognized sources of information on carcinogenesis, including data storage and retrieval systems such as PubMed. Meeting participants who are asked to prepare

preliminary working papers for specific sections are expected to supplement the IARC literature searches with their own searches.

Industrial associations, labour unions and other knowledgeable organizations may be asked to provide input to the sections on production and use, although this involvement is not required as a general rule. Information on production and trade is obtained from governmental, trade and market research publications and, in some cases, by direct contact with industries. Separate production data on some agents may not be available for a variety of reasons (e.g. not collected or made public in all producing countries, production is small). Information on uses may be obtained from published sources but is often complemented by direct contact with manufacturers. Efforts are made to supplement this information with data from other national and international sources.

Six months before the meeting, the material obtained is sent to meeting participants to prepare preliminary working papers. The working papers are compiled by IARC staff and sent, before the meeting, to Working Group Members and Invited Specialists for review.

The Working Group meets at IARC for seven to eight days to discuss and finalize the texts and to formulate the evaluations. The objectives of the meeting are peer review and consensus. During the first few days, four subgroups (covering exposure data, cancer in humans, cancer in experimental animals, and mechanistic and other relevant data) review the working papers, develop a joint subgroup draft and write summaries. Care is taken to ensure that each study summary is written or reviewed by someone not associated with the study being considered. During the last few days, the Working Group meets in plenary session to review the subgroup drafts and develop the evaluations. As a result, the entire volume is the joint product of the Working Group, and there are no individually authored sections.

IARC Working Groups strive to achieve a consensus evaluation. Consensus reflects broad agreement among Working Group Members, but not necessarily unanimity. The chair may elect to poll Working Group Members to determine the diversity of scientific opinion on issues where consensus is not readily apparent.

After the meeting, the master copy is verified by consulting the original literature, edited and prepared for publication. The aim is to publish the volume within six months of the Working Group meeting. A summary of the outcome is available on the *Monographs* programme web site soon after the meeting.

B. SCIENTIFIC REVIEW AND EVALUATION

The available studies are summarized by the Working Group, with particular regard to the qualitative aspects discussed below. In general, numerical findings are indicated as they appear in the original report; units are converted when necessary for easier comparison. The Working Group may conduct additional analyses of the published data and use them in their assessment of the evidence; the results of such supplementary analyses are given in square brackets. When an important aspect of a study that directly impinges on its interpretation should be brought to the attention of the reader, a Working Group comment is given in square brackets.

The scope of the *IARC Monographs* programme has expanded beyond chemicals to include complex mixtures, occupational exposures, physical and biological agents, lifestyle factors and other potentially carcinogenic exposures. Over time, the structure of a *Monograph* has evolved to include the following sections:

Exposure data

Studies of cancer in humans

Studies of cancer in experimental animals

Mechanistic and other relevant data

Summary

Evaluation and rationale

In addition, a section of General Remarks at the front of the volume discusses the reasons the agents were scheduled for evaluation and some key issues the Working Group encountered during the meeting.

This part of the Preamble discusses the types of evidence considered and summarized in each section of a *Monograph*, followed by the scientific criteria that guide the evaluations.

1. Exposure data

Each *Monograph* includes general information on the agent: this information may vary substantially between agents and must be adapted accordingly. Also included is information on production and use (when appropriate), methods of analysis and detection, occurrence, and sources and routes of human occupational and environmental exposures. Depending on the agent, regulations and guidelines for use may be presented.

(a) General information on the agent

For chemical agents, sections on chemical and physical data are included: the Chemical Abstracts Service Registry Number, the latest primary name and the IUPAC systematic name are recorded; other synonyms are given, but the list is not necessarily comprehensive. Information on chemical and physical properties that are relevant to identification, occurrence and biological activity is included. A description of technical products of chemicals includes trade names, relevant specifications and available information on composition and impurities. Some of the trade names given may be those of mixtures in which the agent being evaluated is only one of the ingredients.

For biological agents, taxonomy, structure and biology are described, and the degree of variability is indicated. Mode of replication, life cycle, target cells, persistence, latency, host response and clinical disease other than cancer are also presented.

For physical agents that are forms of radiation, energy and range of the radiation are included. For foreign bodies, fibres and respirable particles, size range and relative dimensions are indicated.

For agents such as mixtures, drugs or lifestyle factors, a description of the agent, including its composition, is given.

Whenever appropriate, other information, such as historical perspectives or the description of an industry or habit, may be included.

(b) Analysis and detection

An overview of methods of analysis and detection of the agent is presented, including their sensitivity, specificity and reproducibility. Methods widely used for regulatory purposes are emphasized. Methods for monitoring human exposure are also given. No critical evaluation or recommendation of any method is meant or implied.

(c) Production and use

The dates of first synthesis and of first commercial production of a chemical, mixture or other agent are provided when available; for agents that do not occur naturally, this information may allow a reasonable estimate to be made of the date before which no human exposure to the agent could have occurred. The dates of first reported occurrence of an exposure are also provided when available. In addition, methods of synthesis used in past and present commercial production and different methods of production,

which may give rise to different impurities, are described.

The countries where companies report production of the agent, and the number of companies in each country, are identified. Available data on production, international trade and uses are obtained for representative regions. It should not, however, be inferred that those areas or nations are necessarily the sole or major sources or users of the agent. Some identified uses may not be current or major applications, and the coverage is not necessarily comprehensive. In the case of drugs, mention of their therapeutic uses does not necessarily represent current practice nor does it imply judgement as to their therapeutic efficacy.

(d) Occurrence and exposure

Information on the occurrence of an agent in the environment is obtained from data derived from the monitoring and surveillance of levels in occupational environments, air, water, soil, plants, foods and animal and human tissues. When available, data on the generation, persistence and bioaccumulation of the agent are also included. Such data may be available from national databases.

Data that indicate the extent of past and present human exposure, the sources of exposure, the people most likely to be exposed and the factors that contribute to the exposure are reported. Information is presented on the range of human exposure, including occupational and environmental exposures. This includes relevant findings from both developed and developing countries. Some of these data are not distributed widely and may be available from government reports and other sources. In the case of mixtures, industries, occupations or processes, information is given about all agents known to be present. For processes, industries and occupations, a historical description is also given, noting variations in chemical composition, physical properties and levels of occupational exposure with date and place. For biological agents, the epidemiology of infection is described.

(e) Regulations and guidelines

Statements concerning regulations and guidelines (e.g. occupational exposure limits, maximal levels permitted in foods and water, pesticide registrations) are included, but they may not reflect the most recent situation, since such limits are continuously reviewed and modified. The absence of information on regulatory status for a country should not be taken to imply that that country does not have regulations with regard to the exposure. For biological agents, legislation and control, including vaccination and therapy, are described.

2. Studies of cancer in humans

This section includes all pertinent epidemiological studies (see Part A, Section 4). Studies of biomarkers are included when they are relevant to an evaluation of carcinogenicity to humans.

(a) Types of study considered

Several types of epidemiological study contribute to the assessment of carcinogenicity in humans — cohort studies, case–control studies, correlation (or ecological) studies and intervention studies. Rarely, results from randomized trials may be available. Case reports and case series of cancer in humans may also be reviewed.

Cohort and case–control studies relate individual exposures under study to the occurrence of cancer in individuals and provide an estimate of effect (such as relative risk) as the main measure of association. Intervention studies may provide strong evidence for making causal inferences, as exemplified by cessation of smoking and the subsequent decrease in risk for lung cancer.

In correlation studies, the units of investigation are usually whole populations (e.g. in

particular geographical areas or at particular times), and cancer frequency is related to a summary measure of the exposure of the population to the agent under study. In correlation studies, individual exposure is not documented, which renders this kind of study more prone to confounding. In some circumstances, however, correlation studies may be more informative than analytical study designs (see, for example, the *Monograph* on arsenic in drinking-water; IARC, 2004).

In some instances, case reports and case series have provided important information about the carcinogenicity of an agent. These types of study generally arise from a suspicion, based on clinical experience, that the concurrence of two events — that is, a particular exposure and occurrence of a cancer — has happened rather more frequently than would be expected by chance. Case reports and case series usually lack complete ascertainment of cases in any population, definition or enumeration of the population at risk and estimation of the expected number of cases in the absence of exposure.

The uncertainties that surround the interpretation of case reports, case series and correlation studies make them inadequate, except in rare instances, to form the sole basis for inferring a causal relationship. When taken together with case–control and cohort studies, however, these types of study may add materially to the judgement that a causal relationship exists.

Epidemiological studies of benign neoplasms, presumed preneoplastic lesions and other end-points thought to be relevant to cancer are also reviewed. They may, in some instances, strengthen inferences drawn from studies of cancer itself.

(b) Quality of studies considered

It is necessary to take into account the possible roles of bias, confounding and chance in the interpretation of epidemiological studies. Bias is the effect of factors in study design or execution that lead erroneously to a stronger or weaker association than in fact exists between an agent and disease. Confounding is a form of bias that occurs when the relationship with disease is made to appear stronger or weaker than it truly is as a result of an association between the apparent causal factor and another factor that is associated with either an increase or decrease in the incidence of the disease. The role of chance is related to biological variability and the influence of sample size on the precision of estimates of effect.

In evaluating the extent to which these factors have been minimized in an individual study, consideration is given to several aspects of design and analysis as described in the report of the study. For example, when suspicion of carcinogenicity arises largely from a single small study, careful consideration is given when interpreting subsequent studies that included these data in an enlarged population. Most of these considerations apply equally to case–control, cohort and correlation studies. Lack of clarity of any of these aspects in the reporting of a study can decrease its credibility and the weight given to it in the final evaluation of the exposure.

First, the study population, disease (or diseases) and exposure should have been well defined by the authors. Cases of disease in the study population should have been identified in a way that was independent of the exposure of interest, and exposure should have been assessed in a way that was not related to disease status.

Second, the authors should have taken into account — in the study design and analysis — other variables that can influence the risk of disease and may have been related to the exposure of interest. Potential confounding by such variables should have been dealt with either in the design of the study, such as by matching, or in the analysis, by statistical adjustment. In cohort studies, comparisons with local rates of disease may or may not be more appropriate than

those with national rates. Internal comparisons of frequency of disease among individuals at different levels of exposure are also desirable in cohort studies, since they minimize the potential for confounding related to the difference in risk factors between an external reference group and the study population.

Third, the authors should have reported the basic data on which the conclusions are founded, even if sophisticated statistical analyses were employed. At the very least, they should have given the numbers of exposed and unexposed cases and controls in a case–control study and the numbers of cases observed and expected in a cohort study. Further tabulations by time since exposure began and other temporal factors are also important. In a cohort study, data on all cancer sites and all causes of death should have been given, to reveal the possibility of reporting bias. In a case–control study, the effects of investigated factors other than the exposure of interest should have been reported.

Finally, the statistical methods used to obtain estimates of relative risk, absolute rates of cancer, confidence intervals and significance tests, and to adjust for confounding should have been clearly stated by the authors. These methods have been reviewed for case–control studies (Breslow & Day, 1980) and for cohort studies (Breslow & Day, 1987).

(c) Meta-analyses and pooled analyses

Independent epidemiological studies of the same agent may lead to results that are difficult to interpret. Combined analyses of data from multiple studies are a means of resolving this ambiguity, and well conducted analyses can be considered. There are two types of combined analysis. The first involves combining summary statistics such as relative risks from individual studies (meta-analysis) and the second involves a pooled analysis of the raw data from the individual studies (pooled analysis) (Greenland, 1998).

The advantages of combined analyses are increased precision due to increased sample size and the opportunity to explore potential confounders, interactions and modifying effects that may explain heterogeneity among studies in more detail. A disadvantage of combined analyses is the possible lack of compatibility of data from various studies due to differences in subject recruitment, procedures of data collection, methods of measurement and effects of unmeasured co-variates that may differ among studies. Despite these limitations, well conducted combined analyses may provide a firmer basis than individual studies for drawing conclusions about the potential carcinogenicity of agents.

IARC may commission a meta-analysis or pooled analysis that is pertinent to a particular *Monograph* (see Part A, Section 4). Additionally, as a means of gaining insight from the results of multiple individual studies, ad hoc calculations that combine data from different studies may be conducted by the Working Group during the course of a *Monograph* meeting. The results of such original calculations, which would be specified in the text by presentation in square brackets, might involve updates of previously conducted analyses that incorporate the results of more recent studies or de-novo analyses. Irrespective of the source of data for the meta-analyses and pooled analyses, it is important that the same criteria for data quality be applied as those that would be applied to individual studies and to ensure also that sources of heterogeneity between studies be taken into account.

(d) Temporal effects

Detailed analyses of both relative and absolute risks in relation to temporal variables, such as age at first exposure, time since first exposure, duration of exposure, cumulative exposure, peak exposure (when appropriate) and

time since cessation of exposure, are reviewed and summarized when available. Analyses of temporal relationships may be useful in making causal inferences. In addition, such analyses may suggest whether a carcinogen acts early or late in the process of carcinogenesis, although, at best, they allow only indirect inferences about mechanisms of carcinogenesis.

(e) Use of biomarkers in epidemiological studies

Biomarkers indicate molecular, cellular or other biological changes and are increasingly used in epidemiological studies for various purposes (IARC, 1991; Vainio et al., 1992; Toniolo et al., 1997; Vineis et al., 1999; Buffler et al., 2004). These may include evidence of exposure, of early effects, of cellular, tissue or organism responses, of individual susceptibility or host responses, and inference of a mechanism (see Part B, Section 4b). This is a rapidly evolving field that encompasses developments in genomics, epigenomics and other emerging technologies.

Molecular epidemiological data that identify associations between genetic polymorphisms and interindividual differences in susceptibility to the agent(s) being evaluated may contribute to the identification of carcinogenic hazards to humans. If the polymorphism has been demonstrated experimentally to modify the functional activity of the gene product in a manner that is consistent with increased susceptibility, these data may be useful in making causal inferences. Similarly, molecular epidemiological studies that measure cell functions, enzymes or metabolites that are thought to be the basis of susceptibility may provide evidence that reinforces biological plausibility. It should be noted, however, that when data on genetic susceptibility originate from multiple comparisons that arise from subgroup analyses, this can generate false-positive results and inconsistencies across studies, and such data therefore require careful evaluation. If the known phenotype of a genetic polymorphism can explain the carcinogenic mechanism of the agent being evaluated, data on this phenotype may be useful in making causal inferences.

(f) Criteria for causality

After the quality of individual epidemiological studies of cancer has been summarized and assessed, a judgement is made concerning the strength of evidence that the agent in question is carcinogenic to humans. In making its judgement, the Working Group considers several criteria for causality (Hill, 1965). A strong association (e.g. a large relative risk) is more likely to indicate causality than a weak association, although it is recognized that estimates of effect of small magnitude do not imply lack of causality and may be important if the disease or exposure is common. Associations that are replicated in several studies of the same design or that use different epidemiological approaches or under different circumstances of exposure are more likely to represent a causal relationship than isolated observations from single studies. If there are inconsistent results among investigations, possible reasons are sought (such as differences in exposure), and results of studies that are judged to be of high quality are given more weight than those of studies that are judged to be methodologically less sound.

If the risk increases with the exposure, this is considered to be a strong indication of causality, although the absence of a graded response is not necessarily evidence against a causal relationship. The demonstration of a decline in risk after cessation of or reduction in exposure in individuals or in whole populations also supports a causal interpretation of the findings.

Several scenarios may increase confidence in a causal relationship. On the one hand, an agent may be specific in causing tumours at one site or of one morphological type. On the other, carcinogenicity may be evident through the causation of

multiple tumour types. Temporality, precision of estimates of effect, biological plausibility and coherence of the overall database are considered. Data on biomarkers may be employed in an assessment of the biological plausibility of epidemiological observations.

Although rarely available, results from randomized trials that show different rates of cancer among exposed and unexposed individuals provide particularly strong evidence for causality.

When several epidemiological studies show little or no indication of an association between an exposure and cancer, a judgement may be made that, in the aggregate, they show evidence of lack of carcinogenicity. Such a judgement requires first that the studies meet, to a sufficient degree, the standards of design and analysis described above. Specifically, the possibility that bias, confounding or misclassification of exposure or outcome could explain the observed results should be considered and excluded with reasonable certainty. In addition, all studies that are judged to be methodologically sound should (a) be consistent with an estimate of effect of unity for any observed level of exposure, (b) when considered together, provide a pooled estimate of relative risk that is at or near to unity, and (c) have a narrow confidence interval, due to sufficient population size. Moreover, no individual study nor the pooled results of all the studies should show any consistent tendency that the relative risk of cancer increases with increasing level of exposure. It is important to note that evidence of lack of carcinogenicity obtained from several epidemiological studies can apply only to the type(s) of cancer studied, to the dose levels reported, and to the intervals between first exposure and disease onset observed in these studies. Experience with human cancer indicates that the period from first exposure to the development of clinical cancer is sometimes longer than 20 years; latent periods substantially shorter than 30 years cannot provide evidence for lack of carcinogenicity.

3. Studies of cancer in experimental animals

All known human carcinogens that have been studied adequately for carcinogenicity in experimental animals have produced positive results in one or more animal species (Wilbourn et al., 1986; Tomatis et al., 1989). For several agents (e.g. aflatoxins, diethylstilbestrol, solar radiation, vinyl chloride), carcinogenicity in experimental animals was established or highly suspected before epidemiological studies confirmed their carcinogenicity in humans (Vainio et al., 1995). Although this association cannot establish that all agents that cause cancer in experimental animals also cause cancer in humans, it is biologically plausible that agents for which there is *sufficient evidence of carcinogenicity* in experimental animals (see Part B, Section 6b) also present a carcinogenic hazard to humans. Accordingly, in the absence of additional scientific information, these agents are considered to pose a carcinogenic hazard to humans. Examples of additional scientific information are data that demonstrate that a given agent causes cancer in animals through a species-specific mechanism that does not operate in humans or data that demonstrate that the mechanism in experimental animals also operates in humans (see Part B, Section 6).

Consideration is given to all available long-term studies of cancer in experimental animals with the agent under review (see Part A, Section 4). In all experimental settings, the nature and extent of impurities or contaminants present in the agent being evaluated are given when available. Animal species, strain (including genetic background where applicable), sex, numbers per group, age at start of treatment, route of exposure, dose levels, duration of exposure, survival and information on tumours (incidence, latency, severity or multiplicity of neoplasms or preneoplastic lesions) are reported. Those studies in experimental animals that are judged to be irrelevant to the evaluation or judged to be inadequate

(e.g. too short a duration, too few animals, poor survival; see below) may be omitted. Guidelines for conducting long-term carcinogenicity experiments have been published (e.g. OECD, 2002).

Other studies considered may include: experiments in which the agent was administered in the presence of factors that modify carcinogenic effects (e.g. initiation–promotion studies, co-carcinogenicity studies and studies in genetically modified animals); studies in which the end-point was not cancer but a defined precancerous lesion; experiments on the carcinogenicity of known metabolites and derivatives; and studies of cancer in non-laboratory animals (e.g. livestock and companion animals) exposed to the agent.

For studies of mixtures, consideration is given to the possibility that changes in the physicochemical properties of the individual substances may occur during collection, storage, extraction, concentration and delivery. Another consideration is that chemical and toxicological interactions of components in a mixture may alter dose–response relationships. The relevance to human exposure of the test mixture administered in the animal experiment is also assessed. This may involve consideration of the following aspects of the mixture tested: (i) physical and chemical characteristics, (ii) identified constituents that may indicate the presence of a class of substances and (iii) the results of genetic toxicity and related tests.

The relevance of results obtained with an agent that is analogous (e.g. similar in structure or of a similar virus genus) to that being evaluated is also considered. Such results may provide biological and mechanistic information that is relevant to the understanding of the process of carcinogenesis in humans and may strengthen the biological plausibility that the agent being evaluated is carcinogenic to humans (see Part B, Section 2f).

(a) Qualitative aspects

An assessment of carcinogenicity involves several considerations of qualitative importance, including (i) the experimental conditions under which the test was performed, including route, schedule and duration of exposure, species, strain (including genetic background where applicable), sex, age and duration of follow-up; (ii) the consistency of the results, for example, across species and target organ(s); (iii) the spectrum of neoplastic response, from preneoplastic lesions and benign tumours to malignant neoplasms; and (iv) the possible role of modifying factors.

Considerations of importance in the interpretation and evaluation of a particular study include: (i) how clearly the agent was defined and, in the case of mixtures, how adequately the sample characterization was reported; (ii) whether the dose was monitored adequately, particularly in inhalation experiments; (iii) whether the doses, duration of treatment and route of exposure were appropriate; (iv) whether the survival of treated animals was similar to that of controls; (v) whether there were adequate numbers of animals per group; (vi) whether both male and female animals were used; (vii) whether animals were allocated randomly to groups; (viii) whether the duration of observation was adequate; and (ix) whether the data were reported and analysed adequately.

When benign tumours (a) occur together with and originate from the same cell type as malignant tumours in an organ or tissue in a particular study and (b) appear to represent a stage in the progression to malignancy, they are usually combined in the assessment of tumour incidence (Huff et al., 1989). The occurrence of lesions presumed to be preneoplastic may in certain instances aid in assessing the biological plausibility of any neoplastic response observed. If an agent induces only benign neoplasms that appear to be end-points that do not readily undergo transition to malignancy, the agent

should nevertheless be suspected of being carcinogenic and requires further investigation.

(b) Quantitative aspects

The probability that tumours will occur may depend on the species, sex, strain, genetic background and age of the animal, and on the dose, route, timing and duration of the exposure. Evidence of an increased incidence of neoplasms with increasing levels of exposure strengthens the inference of a causal association between the exposure and the development of neoplasms.

The form of the dose–response relationship can vary widely, depending on the particular agent under study and the target organ. Mechanisms such as induction of DNA damage or inhibition of repair, altered cell division and cell death rates and changes in intercellular communication are important determinants of dose–response relationships for some carcinogens. Since many chemicals require metabolic activation before being converted to their reactive intermediates, both metabolic and toxicokinetic aspects are important in determining the dose–response pattern. Saturation of steps such as absorption, activation, inactivation and elimination may produce nonlinearity in the dose–response relationship (Hoel et al., 1983; Gart et al., 1986), as could saturation of processes such as DNA repair. The dose–response relationship can also be affected by differences in survival among the treatment groups.

(c) Statistical analyses

Factors considered include the adequacy of the information given for each treatment group: (i) number of animals studied and number examined histologically, (ii) number of animals with a given tumour type and (iii) length of survival. The statistical methods used should be clearly stated and should be the generally accepted techniques refined for this purpose (Peto et al., 1980; Gart et al., 1986; Portier & Bailer, 1989; Bieler & Williams, 1993). The choice of the most appropriate statistical method requires consideration of whether or not there are differences in survival among the treatment groups; for example, reduced survival because of non-tumour-related mortality can preclude the occurrence of tumours later in life. When detailed information on survival is not available, comparisons of the proportions of tumour-bearing animals among the effective number of animals (alive at the time the first tumour was discovered) can be useful when significant differences in survival occur before tumours appear. The lethality of the tumour also requires consideration: for rapidly fatal tumours, the time of death provides an indication of the time of tumour onset and can be assessed using life-table methods; non-fatal or incidental tumours that do not affect survival can be assessed using methods such as the Mantel-Haenzel test for changes in tumour prevalence. Because tumour lethality is often difficult to determine, methods such as the Poly-K test that do not require such information can also be used. When results are available on the number and size of tumours seen in experimental animals (e.g. papillomas on mouse skin, liver tumours observed through nuclear magnetic resonance tomography), other more complicated statistical procedures may be needed (Sherman et al., 1994; Dunson et al., 2003).

Formal statistical methods have been developed to incorporate historical control data into the analysis of data from a given experiment. These methods assign an appropriate weight to historical and concurrent controls on the basis of the extent of between-study and within-study variability: less weight is given to historical controls when they show a high degree of variability, and greater weight when they show little variability. It is generally not appropriate to discount a tumour response that is significantly increased compared with concurrent controls by arguing that it falls within the range of

historical controls, particularly when historical controls show high between-study variability and are, thus, of little relevance to the current experiment. In analysing results for uncommon tumours, however, the analysis may be improved by considering historical control data, particularly when between-study variability is low. Historical controls should be selected to resemble the concurrent controls as closely as possible with respect to species, gender and strain, as well as other factors such as basal diet and general laboratory environment, which may affect tumour-response rates in control animals (Haseman et al., 1984; Fung et al., 1996; Greim et al., 2003).

Although meta-analyses and combined analyses are conducted less frequently for animal experiments than for epidemiological studies due to differences in animal strains, they can be useful aids in interpreting animal data when the experimental protocols are sufficiently similar.

4. Mechanistic and other relevant data

Mechanistic and other relevant data may provide evidence of carcinogenicity and also help in assessing the relevance and importance of findings of cancer in animals and in humans. The nature of the mechanistic and other relevant data depends on the biological activity of the agent being considered. The Working Group considers representative studies to give a concise description of the relevant data and issues that they consider to be important; thus, not every available study is cited. Relevant topics may include toxicokinetics, mechanisms of carcinogenesis, susceptible individuals, populations and life-stages, other relevant data and other adverse effects. When data on biomarkers are informative about the mechanisms of carcinogenesis, they are included in this section.

These topics are not mutually exclusive; thus, the same studies may be discussed in more than one subsection. For example, a mutation in a gene that codes for an enzyme that metabolizes the agent under study could be discussed in the subsections on toxicokinetics, mechanisms and individual susceptibility if it also exists as an inherited polymorphism.

(a) Toxicokinetic data

Toxicokinetics refers to the absorption, distribution, metabolism and elimination of agents in humans, experimental animals and, where relevant, cellular systems. Examples of kinetic factors that may affect dose–response relationships include uptake, deposition, biopersistence and half-life in tissues, protein binding, metabolic activation and detoxification. Studies that indicate the metabolic fate of the agent in humans and in experimental animals are summarized briefly, and comparisons of data from humans and animals are made when possible. Comparative information on the relationship between exposure and the dose that reaches the target site may be important for the extrapolation of hazards between species and in clarifying the role of in-vitro findings.

(b) Data on mechanisms of carcinogenesis

To provide focus, the Working Group attempts to identify the possible mechanisms by which the agent may increase the risk of cancer. For each possible mechanism, a representative selection of key data from humans and experimental systems is summarized. Attention is given to gaps in the data and to data that suggests that more than one mechanism may be operating. The relevance of the mechanism to humans is discussed, in particular, when mechanistic data are derived from experimental model systems. Changes in the affected organs, tissues or cells

can be divided into three non-exclusive levels as described below.

(i) Changes in physiology

Physiological changes refer to exposure-related modifications to the physiology and/or response of cells, tissues and organs. Examples of potentially adverse physiological changes include mitogenesis, compensatory cell division, escape from apoptosis and/or senescence, presence of inflammation, hyperplasia, metaplasia and/or preneoplasia, angiogenesis, alterations in cellular adhesion, changes in steroidal hormones and changes in immune surveillance.

(ii) Functional changes at the cellular level

Functional changes refer to exposure-related alterations in the signalling pathways used by cells to manage critical processes that are related to increased risk for cancer. Examples of functional changes include modified activities of enzymes involved in the metabolism of xenobiotics, alterations in the expression of key genes that regulate DNA repair, alterations in cyclin-dependent kinases that govern cell cycle progression, changes in the patterns of post-translational modifications of proteins, changes in regulatory factors that alter apoptotic rates, changes in the secretion of factors related to the stimulation of DNA replication and transcription and changes in gap–junction-mediated intercellular communication.

(iii) Changes at the molecular level

Molecular changes refer to exposure-related changes in key cellular structures at the molecular level, including, in particular, genotoxicity. Examples of molecular changes include formation of DNA adducts and DNA strand breaks, mutations in genes, chromosomal aberrations, aneuploidy and changes in DNA methylation patterns. Greater emphasis is given to irreversible effects.

The use of mechanistic data in the identification of a carcinogenic hazard is specific to the mechanism being addressed and is not readily described for every possible level and mechanism discussed above.

Genotoxicity data are discussed here to illustrate the key issues involved in the evaluation of mechanistic data.

Tests for genetic and related effects are described in view of the relevance of gene mutation and chromosomal aberration/aneuploidy to carcinogenesis (Vainio et al., 1992; McGregor et al., 1999). The adequacy of the reporting of sample characterization is considered and, when necessary, commented upon; with regard to complex mixtures, such comments are similar to those described for animal carcinogenicity tests. The available data are interpreted critically according to the end-points detected, which may include DNA damage, gene mutation, sister chromatid exchange, micronucleus formation, chromosomal aberrations and aneuploidy. The concentrations employed are given, and mention is made of whether the use of an exogenous metabolic system in vitro affected the test result. These data are listed in tabular form by phylogenetic classification.

Positive results in tests using prokaryotes, lower eukaryotes, insects, plants and cultured mammalian cells suggest that genetic and related effects could occur in mammals. Results from such tests may also give information on the types of genetic effect produced and on the involvement of metabolic activation. Some end-points described are clearly genetic in nature (e.g. gene mutations), while others are associated with genetic effects (e.g. unscheduled DNA synthesis). In-vitro tests for tumour promotion, cell transformation and gap–junction intercellular communication may be sensitive to changes that are not necessarily the result of genetic alterations but that may have specific relevance to the process of carcinogenesis. Critical appraisals of these tests

have been published (Montesano et al., 1986; McGregor et al., 1999).

Genetic or other activity manifest in humans and experimental mammals is regarded to be of greater relevance than that in other organisms. The demonstration that an agent can induce gene and chromosomal mutations in mammals in vivo indicates that it may have carcinogenic activity. Negative results in tests for mutagenicity in selected tissues from animals treated in vivo provide less weight, partly because they do not exclude the possibility of an effect in tissues other than those examined. Moreover, negative results in short-term tests with genetic end-points cannot be considered to provide evidence that rules out the carcinogenicity of agents that act through other mechanisms (e.g. receptor-mediated effects, cellular toxicity with regenerative cell division, peroxisome proliferation) (Vainio et al., 1992). Factors that may give misleading results in short-term tests have been discussed in detail elsewhere (Montesano et al., 1986; McGregor et al., 1999).

When there is evidence that an agent acts by a specific mechanism that does not involve genotoxicity (e.g. hormonal dysregulation, immune suppression, and formation of calculi and other deposits that cause chronic irritation), that evidence is presented and reviewed critically in the context of rigorous criteria for the operation of that mechanism in carcinogenesis (e.g. Capen et al., 1999).

For biological agents such as viruses, bacteria and parasites, other data relevant to carcinogenicity may include descriptions of the pathology of infection, integration and expression of viruses, and genetic alterations seen in human tumours. Other observations that might comprise cellular and tissue responses to infection, immune response and the presence of tumour markers are also considered.

For physical agents that are forms of radiation, other data relevant to carcinogenicity may include descriptions of damaging effects at the physiological, cellular and molecular level, as for chemical agents, and descriptions of how these effects occur. 'Physical agents' may also be considered to comprise foreign bodies, such as surgical implants of various kinds, and poorly soluble fibres, dusts and particles of various sizes, the pathogenic effects of which are a result of their physical presence in tissues or body cavities. Other relevant data for such materials may include characterization of cellular, tissue and physiological reactions to these materials and descriptions of pathological conditions other than neoplasia with which they may be associated.

(c) Other data relevant to mechanisms

A description is provided of any structure–activity relationships that may be relevant to an evaluation of the carcinogenicity of an agent, the toxicological implications of the physical and chemical properties, and any other data relevant to the evaluation that are not included elsewhere.

High-output data, such as those derived from gene expression microarrays, and high-throughput data, such as those that result from testing hundreds of agents for a single end-point, pose a unique problem for the use of mechanistic data in the evaluation of a carcinogenic hazard. In the case of high-output data, there is the possibility to overinterpret changes in individual end-points (e.g. changes in expression in one gene) without considering the consistency of that finding in the broader context of the other end-points (e.g. other genes with linked transcriptional control). High-output data can be used in assessing mechanisms, but all end-points measured in a single experiment need to be considered in the proper context. For high-throughput data, where the number of observations far exceeds the number of end-points measured, their utility for identifying common mechanisms across multiple agents is enhanced. These data can be used to identify mechanisms that not only seem

plausible, but also have a consistent pattern of carcinogenic response across entire classes of related compounds.

(d) Susceptibility data

Individuals, populations and life-stages may have greater or lesser susceptibility to an agent, based on toxicokinetics, mechanisms of carcinogenesis and other factors. Examples of host and genetic factors that affect individual susceptibility include sex, genetic polymorphisms of genes involved in the metabolism of the agent under evaluation, differences in metabolic capacity due to life-stage or the presence of disease, differences in DNA repair capacity, competition for or alteration of metabolic capacity by medications or other chemical exposures, pre-existing hormonal imbalance that is exacerbated by a chemical exposure, a suppressed immune system, periods of higher-than-usual tissue growth or regeneration and genetic polymorphisms that lead to differences in behaviour (e.g. addiction). Such data can substantially increase the strength of the evidence from epidemiological data and enhance the linkage of in-vivo and in-vitro laboratory studies to humans.

(e) Data on other adverse effects

Data on acute, subchronic and chronic adverse effects relevant to the cancer evaluation are summarized. Adverse effects that confirm distribution and biological effects at the sites of tumour development, or alterations in physiology that could lead to tumour development, are emphasized. Effects on reproduction, embryonic and fetal survival and development are summarized briefly. The adequacy of epidemiological studies of reproductive outcome and genetic and related effects in humans is judged by the same criteria as those applied to epidemiological studies of cancer, but fewer details are given.

5. Summary

This section is a summary of data presented in the preceding sections. Summaries can be found on the *Monographs* programme web site (http://monographs.iarc.fr).

(a) Exposure data

Data are summarized, as appropriate, on the basis of elements such as production, use, occurrence and exposure levels in the workplace and environment and measurements in human tissues and body fluids. Quantitative data and time trends are given to compare exposures in different occupations and environmental settings. Exposure to biological agents is described in terms of transmission, prevalence and persistence of infection.

(b) Cancer in humans

Results of epidemiological studies pertinent to an assessment of human carcinogenicity are summarized. When relevant, case reports and correlation studies are also summarized. The target organ(s) or tissue(s) in which an increase in cancer was observed is identified. Dose–response and other quantitative data may be summarized when available.

(c) Cancer in experimental animals

Data relevant to an evaluation of carcinogenicity in animals are summarized. For each animal species, study design and route of administration, it is stated whether an increased incidence, reduced latency, or increased severity or multiplicity of neoplasms or preneoplastic lesions were observed, and the tumour sites are indicated. If the agent produced tumours after prenatal exposure or in single-dose experiments, this is also mentioned. Negative findings, inverse relationships, dose–response and other quantitative data are also summarized.

(d) Mechanistic and other relevant data

Data relevant to the toxicokinetics (absorption, distribution, metabolism, elimination) and the possible mechanism(s) of carcinogenesis (e.g. genetic toxicity, epigenetic effects) are summarized. In addition, information on susceptible individuals, populations and life-stages is summarized. This section also reports on other toxic effects, including reproductive and developmental effects, as well as additional relevant data that are considered to be important.

6. Evaluation and rationale

Evaluations of the strength of the evidence for carcinogenicity arising from human and experimental animal data are made, using standard terms. The strength of the mechanistic evidence is also characterized.

It is recognized that the criteria for these evaluations, described below, cannot encompass all of the factors that may be relevant to an evaluation of carcinogenicity. In considering all of the relevant scientific data, the Working Group may assign the agent to a higher or lower category than a strict interpretation of these criteria would indicate.

These categories refer only to the strength of the evidence that an exposure is carcinogenic and not to the extent of its carcinogenic activity (potency). A classification may change as new information becomes available.

An evaluation of the degree of evidence is limited to the materials tested, as defined physically, chemically or biologically. When the agents evaluated are considered by the Working Group to be sufficiently closely related, they may be grouped together for the purpose of a single evaluation of the degree of evidence.

(a) Carcinogenicity in humans

The evidence relevant to carcinogenicity from studies in humans is classified into one of the following categories:

***Sufficient evidence of carcinogenicity*:**

The Working Group considers that a causal relationship has been established between exposure to the agent and human cancer. That is, a positive relationship has been observed between the exposure and cancer in studies in which chance, bias and confounding could be ruled out with reasonable confidence. A statement that there is *sufficient evidence* is followed by a separate sentence that identifies the target organ(s) or tissue(s) where an increased risk of cancer was observed in humans. Identification of a specific target organ or tissue does not preclude the possibility that the agent may cause cancer at other sites.

***Limited evidence of carcinogenicity*:**

A positive association has been observed between exposure to the agent and cancer for which a causal interpretation is considered by the Working Group to be credible, but chance, bias or confounding could not be ruled out with reasonable confidence.

***Inadequate evidence of carcinogenicity*:**

The available studies are of insufficient quality, consistency or statistical power to permit a conclusion regarding the presence or absence of a causal association between exposure and cancer, or no data on cancer in humans are available.

***Evidence suggesting lack of carcinogenicity*:**

There are several adequate studies covering the full range of levels of exposure that humans are known to encounter, which are mutually consistent in not showing a positive association between exposure to the agent and any studied cancer at any observed level of exposure. The results from these studies alone or combined should have narrow confidence intervals with an upper limit close to the null value (e.g. a relative

risk of 1.0). Bias and confounding should be ruled out with reasonable confidence, and the studies should have an adequate length of follow-up. A conclusion of *evidence suggesting lack of carcinogenicity* is inevitably limited to the cancer sites, conditions and levels of exposure, and length of observation covered by the available studies. In addition, the possibility of a very small risk at the levels of exposure studied can never be excluded.

In some instances, the above categories may be used to classify the degree of evidence related to carcinogenicity in specific organs or tissues.

When the available epidemiological studies pertain to a mixture, process, occupation or industry, the Working Group seeks to identify the specific agent considered most likely to be responsible for any excess risk. The evaluation is focused as narrowly as the available data on exposure and other aspects permit.

(b) Carcinogenicity in experimental animals

Carcinogenicity in experimental animals can be evaluated using conventional bioassays, bioassays that employ genetically modified animals, and other in-vivo bioassays that focus on one or more of the critical stages of carcinogenesis. In the absence of data from conventional long-term bioassays or from assays with neoplasia as the end-point, consistently positive results in several models that address several stages in the multistage process of carcinogenesis should be considered in evaluating the degree of evidence of carcinogenicity in experimental animals.

The evidence relevant to carcinogenicity in experimental animals is classified into one of the following categories:

***Sufficient evidence of carcinogenicity*:**

The Working Group considers that a causal relationship has been established between the agent and an increased incidence of malignant neoplasms or of an appropriate combination of benign and malignant neoplasms in (a) two or more species of animals or (b) two or more independent studies in one species carried out at different times or in different laboratories or under different protocols. An increased incidence of tumours in both sexes of a single species in a well conducted study, ideally conducted under Good Laboratory Practices, can also provide *sufficient evidence*.

A single study in one species and sex might be considered to provide *sufficient evidence of carcinogenicity* when malignant neoplasms occur to an unusual degree with regard to incidence, site, type of tumour or age at onset, or when there are strong findings of tumours at multiple sites.

***Limited evidence of carcinogenicity*:**

The data suggest a carcinogenic effect but are limited for making a definitive evaluation because, e.g. (a) the evidence of carcinogenicity is restricted to a single experiment; (b) there are unresolved questions regarding the adequacy of the design, conduct or interpretation of the studies; (c) the agent increases the incidence only of benign neoplasms or lesions of uncertain neoplastic potential; or (d) the evidence of carcinogenicity is restricted to studies that demonstrate only promoting activity in a narrow range of tissues or organs.

***Inadequate evidence of carcinogenicity*:**

The studies cannot be interpreted as showing either the presence or absence of a carcinogenic effect because of major qualitative or quantitative limitations, or no data on cancer in experimental animals are available.

***Evidence suggesting lack of carcinogenicity*:**

Adequate studies involving at least two species are available which show that, within the limits of the tests used, the agent is not carcinogenic. A conclusion of *evidence suggesting lack of carcinogenicity* is inevitably limited to the species, tumour sites, age at exposure, and conditions and levels of exposure studied.

(c) Mechanistic and other relevant data

Mechanistic and other evidence judged to be relevant to an evaluation of carcinogenicity and of sufficient importance to affect the overall evaluation is highlighted. This may include data on preneoplastic lesions, tumour pathology, genetic and related effects, structure–activity relationships, metabolism and toxicokinetics, physicochemical parameters and analogous biological agents.

The strength of the evidence that any carcinogenic effect observed is due to a particular mechanism is evaluated, using terms such as 'weak', 'moderate' or 'strong'. The Working Group then assesses whether that particular mechanism is likely to be operative in humans. The strongest indications that a particular mechanism operates in humans derive from data on humans or biological specimens obtained from exposed humans. The data may be considered to be especially relevant if they show that the agent in question has caused changes in exposed humans that are on the causal pathway to carcinogenesis. Such data may, however, never become available, because it is at least conceivable that certain compounds may be kept from human use solely on the basis of evidence of their toxicity and/or carcinogenicity in experimental systems.

The conclusion that a mechanism operates in experimental animals is strengthened by findings of consistent results in different experimental systems, by the demonstration of biological plausibility and by coherence of the overall database. Strong support can be obtained from studies that challenge the hypothesized mechanism experimentally, by demonstrating that the suppression of key mechanistic processes leads to the suppression of tumour development. The Working Group considers whether multiple mechanisms might contribute to tumour development, whether different mechanisms might operate in different dose ranges, whether separate mechanisms might operate in humans and experimental animals and whether a unique mechanism might operate in a susceptible group. The possible contribution of alternative mechanisms must be considered before concluding that tumours observed in experimental animals are not relevant to humans. An uneven level of experimental support for different mechanisms may reflect that disproportionate resources have been focused on investigating a favoured mechanism.

For complex exposures, including occupational and industrial exposures, the chemical composition and the potential contribution of carcinogens known to be present are considered by the Working Group in its overall evaluation of human carcinogenicity. The Working Group also determines the extent to which the materials tested in experimental systems are related to those to which humans are exposed.

(d) Overall evaluation

Finally, the body of evidence is considered as a whole, to reach an overall evaluation of the carcinogenicity of the agent to humans.

An evaluation may be made for a group of agents that have been evaluated by the Working Group. In addition, when supporting data indicate that other related agents, for which there is no direct evidence of their capacity to induce cancer in humans or in animals, may also be carcinogenic, a statement describing the rationale for this conclusion is added to the evaluation narrative; an additional evaluation may be made for this broader group of agents if the strength of the evidence warrants it.

The agent is described according to the wording of one of the following categories, and the designated group is given. The categorization of an agent is a matter of scientific judgement that reflects the strength of the evidence derived from studies in humans and in experimental animals and from mechanistic and other relevant data.

Group 1: The agent is carcinogenic to humans.

This category is used when there is *sufficient evidence of carcinogenicity* in humans. Exceptionally, an agent may be placed in this category when evidence of carcinogenicity in humans is less than *sufficient* but there is *sufficient evidence of carcinogenicity* in experimental animals and strong evidence in exposed humans that the agent acts through a relevant mechanism of carcinogenicity.

Group 2.

This category includes agents for which, at one extreme, the degree of evidence of carcinogenicity in humans is almost *sufficient*, as well as those for which, at the other extreme, there are no human data but for which there is evidence of carcinogenicity in experimental animals. Agents are assigned to either Group 2A (*probably carcinogenic to humans*) or Group 2B (*possibly carcinogenic to humans*) on the basis of epidemiological and experimental evidence of carcinogenicity and mechanistic and other relevant data. The terms *probably carcinogenic* and *possibly carcinogenic* have no quantitative significance and are used simply as descriptors of different levels of evidence of human carcinogenicity, with *probably carcinogenic* signifying a higher level of evidence than *possibly carcinogenic*.

Group 2A: The agent is probably carcinogenic to humans.

This category is used when there is *limited evidence of carcinogenicity* in humans and *sufficient evidence of carcinogenicity* in experimental animals. In some cases, an agent may be classified in this category when there is *inadequate evidence of carcinogenicity* in humans and *sufficient evidence of carcinogenicity* in experimental animals and strong evidence that the carcinogenesis is mediated by a mechanism that also operates in humans. Exceptionally, an agent may be classified in this category solely on the basis of *limited evidence of carcinogenicity* in humans. An agent may be assigned to this category if it clearly belongs, based on mechanistic considerations, to a class of agents for which one or more members have been classified in Group 1 or Group 2A.

Group 2B: The agent is possibly carcinogenic to humans.

This category is used for agents for which there is *limited evidence of carcinogenicity* in humans and less than *sufficient evidence of carcinogenicity* in experimental animals. It may also be used when there is *inadequate evidence of carcinogenicity* in humans but there is *sufficient evidence of carcinogenicity* in experimental animals. In some instances, an agent for which there is *inadequate evidence of carcinogenicity* in humans and less than *sufficient evidence of carcinogenicity* in experimental animals together with supporting evidence from mechanistic and other relevant data may be placed in this group. An agent may be classified in this category solely on the basis of strong evidence from mechanistic and other relevant data.

Group 3: The agent is not classifiable as to its carcinogenicity to humans.

This category is used most commonly for agents for which the evidence of carcinogenicity is *inadequate* in humans and *inadequate* or *limited* in experimental animals.

Exceptionally, agents for which the evidence of carcinogenicity is *inadequate* in humans but *sufficient* in experimental animals may be placed in this category when there is strong evidence that the mechanism of carcinogenicity in experimental animals does not operate in humans.

Agents that do not fall into any other group are also placed in this category.

An evaluation in Group 3 is not a determination of non-carcinogenicity or overall safety. It often means that further research is needed,

especially when exposures are widespread or the cancer data are consistent with differing interpretations.

Group 4: The agent is probably not carcinogenic to humans.

This category is used for agents for which there is *evidence suggesting lack of carcinogenicity* in humans and in experimental animals. In some instances, agents for which there is *inadequate evidence of carcinogenicity* in humans but *evidence suggesting lack of carcinogenicity* in experimental animals, consistently and strongly supported by a broad range of mechanistic and other relevant data, may be classified in this group.

(e) Rationale

The reasoning that the Working Group used to reach its evaluation is presented and discussed. This section integrates the major findings from studies of cancer in humans, studies of cancer in experimental animals, and mechanistic and other relevant data. It includes concise statements of the principal line(s) of argument that emerged, the conclusions of the Working Group on the strength of the evidence for each group of studies, citations to indicate which studies were pivotal to these conclusions, and an explanation of the reasoning of the Working Group in weighing data and making evaluations. When there are significant differences of scientific interpretation among Working Group Members, a brief summary of the alternative interpretations is provided, together with their scientific rationale and an indication of the relative degree of support for each alternative.

References

Bieler GS, Williams RL (1993). Ratio estimates, the delta method, and quantal response tests for increased carcinogenicity. *Biometrics*, 49:793–801. doi:10.2307/2532200 PMID:8241374

Breslow NE, Day NE (1980). Statistical methods in cancer research. Volume I - The analysis of case-control studies. *IARC Sci Publ*, 32:5–338. PMID:7216345

Breslow NE, Day NE (1987). Statistical methods in cancer research. Volume II–The design and analysis of cohort studies. *IARC Sci Publ*, 82:1–406. PMID:3329634

Buffler P, Rice J, Baan R et al. (2004). Workshop on mechanisms of carcinogenesis: contributions of molecular epidemiology. Lyon, 14–17 November 2001. Workshop report. *IARC Sci Publ*, 157:1–27. PMID:15055286

Capen CC, Dybing E, Rice JM, Wilbourn JD (1999). Species differences in thyroid, kidney and urinary bladder carcinogenesis. Proceedings of a consensus conference. Lyon, France, 3–7 November 1997. *IARC Sci Publ*, 147:1–225. PMID:10627184

Cogliano V, Baan R, Straif K et al. (2005). Transparency in IARC Monographs. *Lancet Oncol*, 6:747. doi:10.1016/S1470-2045(05)70380-6

Cogliano VJ, Baan RA, Straif K et al. (2004). The science and practice of carcinogen identification and evaluation. *Environ Health Perspect*, 112:1269–1274. doi:10.1289/ehp.6950 PMID:15345338

Dunson DB, Chen Z, Harry J (2003). A Bayesian approach for joint modeling of cluster size and subunit-specific outcomes. *Biometrics*, 59:521–530. doi:10.1111/1541-0420.00062 PMID:14601753

Fung KY, Krewski D, Smythe RT (1996). A comparison of tests for trend with historical controls in carcinogen bioassay. *Can J Stat*, 24:431–454. doi:10.2307/3315326

Gart JJ, Krewski D, Lee PN et al. (1986). Statistical methods in cancer research. Volume III–The design and analysis of long-term animal experiments. *IARC Sci Publ*, 79:1–219. PMID:3301661

Greenland S (1998). Meta-analysis. In: Rothman KJ, Greenland S, editors. Modern epidemiology. Philadelphia: Lippincott Williams & Wilkins, pp. 643–673.

Greim H, Gelbke H-P, Reuter U et al. (2003). Evaluation of historical control data in carcinogenicity studies. *Hum Exp Toxicol*, 22:541–549. doi:10.1191/0960327103ht394oa PMID:14655720

Haseman JK, Huff J, Boorman GA (1984). Use of historical control data in carcinogenicity studies in rodents. *Toxicol Pathol*, 12:126–135. doi:10.1177/019262338401200203 PMID:11478313

Hill AB (1965). The environment and disease: Association or causation? *Proc R Soc Med*, 58:295–300. PMID:14283879

Hoel DG, Kaplan NL, Anderson MW (1983). Implication of nonlinear kinetics on risk estimation in carcinogenesis. *Science*, 219:1032–1037. doi:10.1126/science.6823565 PMID:6823565

Huff JE, Eustis SL, Haseman JK (1989). Occurrence and relevance of chemically induced benign neoplasms in long-term carcinogenicity studies. *Cancer Metastasis Rev*, 8:1–22. doi:10.1007/BF00047055 PMID:2667783

IARC (1977). IARC Monographs Programme on the Evaluation of the Carcinogenic Risk of Chemicals to Humans. Preamble (IARC Intern Tech Rep No. 77/002).

IARC (1978). Chemicals with sufficient evidence of carcinogenicity in experimental animals – IARC Monographs Volumes 1–17 (IARC Intern Tech Rep No. 78/003).

IARC (1979). Criteria to select chemicals for IARC Monographs (IARC Intern Tech Rep No. 79/003).

IARC (1982). Chemicals, industrial processes and industries associated with cancer in humans (IARC Monographs, volumes 1 to 29). *IARC Monogr Eval Carcinog Risk Chem Hum Suppl*, 4:1–292.

IARC (1983). Approaches to classifying chemical carcinogens according to mechanism of action (IARC Intern Tech Rep No. 83/001).

IARC (1987). Overall evaluations of carcinogenicity: an updating of IARC Monographs volumes 1 to 42. *IARC Monogr Eval Carcinog Risks Hum Suppl*, 7:1–440. PMID:3482203

IARC (1988). Report of an IARC Working Group to Review the Approaches and Processes Used to Evaluate the Carcinogenicity of Mixtures and Groups of Chemicals (IARC Intern Tech Rep No. 88/002).

IARC (1991). A consensus report of an IARC Monographs Working Group on the Use of Mechanisms of Carcinogenesis in Risk Identification (IARC Intern Tech Rep No. 91/002).

IARC (2004). Some drinking-water disinfectants and contaminants, including arsenic. *IARC Monogr Eval Carcinog Risks Hum*, 84:1–477. PMID:15645577

IARC (2005). Report of the Advisory Group to Recommend Updates to the Preamble to the IARC Monographs (IARC Intern Rep No. 05/001).

IARC (2006). Report of the Advisory Group to Review the Amended Preamble to the IARC Monographs (IARC Intern Rep No. 06/001).

McGregor DB, Rice JM, Venitt S (1999). The use of short-and medium-term tests for carcinogens and data on genetic effects in carcinogenic hazard evaluation. Consensus report. *IARC Sci Publ*, 146:1–18. PMID:10353381

Montesano R, Bartsch H, Vainio H et al., editors (1986). Long-term and short-term assays for carcinogenesis—a critical appraisal. *IARC Sci Publ*, 83:1–564. PMID:3623675

OECD (2002). Guidance notes for analysis and evaluation of chronic toxicity and carcinogenicity studies (Series on Testing and Assessment No. 35), Paris: OECD.

Peto R, Pike MC, Day NE et al. (1980). Guidelines for simple, sensitive significance tests for carcinogenic effects in long-term animal experiments. *IARC Monogr Eval Carcinog Risk Chem Hum Suppl*, 2:Suppl: 311–426. PMID:6935185

Portier CJ, Bailer AJ (1989). Testing for increased carcinogenicity using a survival-adjusted quantal response test. *Fundam Appl Toxicol*, 12:731–737. doi:10.1016/0272-0590(89)90004-3 PMID:2744275

Sherman CD, Portier CJ, Kopp-Schneider A (1994). Multistage models of carcinogenesis: an approximation for the size and number distribution of late-stage clones. *Risk Anal*, 14:1039–1048. doi:10.1111/j.1539-6924.1994.tb00074.x PMID:7846311

Stewart BW, Kleihues P, editors (2003). World cancer report. Lyon: IARC.

Tomatis L, Aitio A, Wilbourn J, Shuker L (1989). Human carcinogens so far identified. *Jpn J Cancer Res*, 80:795–807. doi:10.1111/j.1349-7006.1989.tb01717.x PMID:2513295

Toniolo P, Boffetta P, Shuker DEG et al. (1997). Proceedings of the workshop on application of biomarkers to cancer epidemiology. Lyon, France, 20–23 February 1996. *IARC Sci Publ*, 142:1–318. PMID:9410826

Vainio H, Magee P, McGregor D, McMichael A (1992). Mechanisms of carcinogenesis in risk identification. IARC Working Group Meeting. Lyon, 11–18 June 1991. *IARC Sci Publ*, 116:1–608. PMID:1428077

Vainio H, Wilbourn JD, Sasco AJ et al. (1995). [Identification of human carcinogenic risks in IARC monographs] *Bull Cancer*, 82:339–348. PMID:7626841

Vineis P, Malats N, Lang M et al., editors (1999). Metabolic polymorphisms and susceptibility to cancer. *IARC Sci Publ*, 148:1–510. PMID:10493243

Wilbourn J, Haroun L, Heseltine E et al. (1986). Response of experimental animals to human carcinogens: an analysis based upon the IARC Monographs programme. *Carcinogenesis*, 7:1853–1863. doi:10.1093/carcin/7.11.1853 PMID:3769134

GENERAL REMARKS

This one-hundred-and-twentieth volume of the *IARC Monographs* presents an evaluation of the carcinogenic hazard to humans of exposure to benzene.

The conclusions of this volume represent the sixth evaluation of the carcinogenicity of benzene by an *IARC Monographs* Working Group. Successive evaluations published in Volumes 7 (IARC, 1974), 29 (IARC, 1982), and 100F (IARC, 2012) and Supplements 1 (IARC, 1979) and 7 (IARC, 1987) considered progressively larger and more complex volumes of data and yielded consistent, yet steadily broader and more compelling, conclusions about the carcinogenicity of benzene.

The available data were sparse at the time of the first evaluation (IARC, 1974). The Working Group determined that the available evidence from studies of experimental animals did not permit a conclusion to be drawn, but found suggestive evidence from epidemiological case reports and one case–control study that benzene causes leukaemia in humans; the current system of formal classifications of evidence had not yet been introduced at that time (it was introduced in Volume 17). Benzene was reviewed again in Supplement 1, which updated Volumes 1–20. With formal classifications then in place, the evidence in experimental animals was found to be *inadequate*, and the human epidemiological evidence, now supplemented by several occupational cohort studies and case–control studies in addition to case reports, was found to be *sufficient*. In the overall evaluation, benzene was found to be *carcinogenic to humans* (Group 1), a finding that has stood since that time.

Additional data had become available when benzene was reviewed again for Volume 29 (IARC, 1982). The Working Group now found the evidence in experimental animals to be *limited* and concluded that the modestly expanded epidemiological evidence established a causal relationship between exposure to benzene and development of acute myeloid leukaemia.

With further growth of the database during the 1980s, the evidence in experimental animals was found to be *sufficient* when benzene was evaluated again for Supplement 7 (IARC, 1987). Although mechanistic evidence was not yet formally incorporated into overall evaluations at that time, induction of chromosomal aberrations in exposed humans and of chromosomal aberrations and micronuclei in rodents was also noted in the summary report.

The volume of evidence had grown substantially larger and more complex by 2009, when the evaluation of benzene was updated for Volume 100F (IARC, 2012). The Working Group confirmed the previous findings of *sufficient evidence* of carcinogenicity in humans and experimental animals and, for the first time, presented *strong evidence* of multiple genotoxic effects based on a review of extensive mechanistic data. In humans, the Working Group concluded

that benzene causes acute myeloid leukaemia/acute non-lymphocytic leukaemia (both terms were used in epidemiological studies reviewed in that volume) and found *limited* evidence that benzene causes acute lymphocytic leukaemia, chronic lymphocytic leukaemia, non-Hodgkin lymphoma, and multiple myeloma.

The current evaluation was undertaken with two principal goals: (i) to incorporate new epidemiological and experimental evidence, including a large number of mechanistic studies in exposed humans, and (ii) to assess quantitative exposure–response relationships of exposure to benzene with both human cancer risks and relevant biological end-points in exposed humans. Such quantitative evaluations were recommended as an adjunct to future *Monographs* by an Advisory Group on quantitative risk characterization (IARC, 2014).

In the current evaluation, the Working Group again confirmed the carcinogenicity of benzene based on *sufficient evidence* of carcinogenicity in humans, *sufficient evidence* of carcinogenicity in experimental animals, and *strong* mechanistic evidence. The Working Group's evaluation of the accumulated evidence from human epidemiological studies focused on studies in which occupational or environmental exposure to benzene was specifically identified. The findings fully supported the previous conclusion that benzene causes acute non-lymphocytic leukaemia – including acute myeloid leukaemia – in adults, as well as the previous observations of *limited evidence* for chronic lymphocytic leukaemia, non-Hodgkin lymphoma, and multiple myeloma. On the basis of new data available since the last review, the Working Group also found *limited* evidence that benzene causes chronic myeloid leukaemia and lung cancer, and acute myeloid leukaemia in children. The Working Group's review of the large body of mechanistic studies took into account the key characteristics of carcinogens (Smith et al., 2016). The Working Group affirmed the *strong evidence* that benzene is genotoxic, and found that it also exhibits many other key characteristics of carcinogens, including in exposed humans. In particular, benzene is metabolically activated to electrophilic metabolites; induces oxidative stress and associated oxidative damage to DNA; is genotoxic; alters DNA repair or causes genomic instability; is immunosuppressive; alters cell proliferation, cell death, or nutrient supply; and modulates receptor-mediated effects.

The evidence reviewed for this evaluation, the Working Group's conclusions, and their analysis of exposure–response relationships are detailed in this volume. A summary of the key findings has appeared in *The Lancet Oncology* (Loomis et al., 2017).

References

IARC (1974). Some anti-thyroid and related substances, nitrofurans and industrial chemical. *IARC Monogr Eval Carcinog Risk Chem Man*, 7:1–326. Available from: http://publications.iarc.fr/25

IARC (1979). Chemicals and industrial processes associated with cancer in humans. IARC Monographs, volumes 1 to 20. *IARC Monogr Eval Carcinog Risk Chem Hum Suppl*, 1:1–71. Available from: http://publications.iarc.fr/133 PMID:296141

IARC (1982). Some industrial chemicals and dyestuffs. *IARC Monogr Eval Carcinog Risk Chem Hum*, 29:1––398. Available from: http://publications.iarc.fr/47 PMID:6957379

IARC (1987). Overall evaluations of carcinogenicity: an updating of IARC Monographs volumes 1 to 42. *IARC Monogr Eval Carcinog Risks Hum Suppl*, 7:1–440. Available from: http://publications.iarc.fr/139 PMID:3482203

IARC (2012). Chemical agents and related occupations. *IARC Monogr Eval Carcinog Risks Hum*, 100F:1–599. Available from: http://publications.iarc.fr/123 PMID:23189753

IARC (2014). Report of the IARC Advisory Group to Recommend on Quantitative Risk Characterization (IARC Internal Report 14/001). Available from: https://monographs.iarc.fr/wp-content/uploads/2018/06/14-001.pdf.

Loomis D, Guyton KZ, Grosse Y, El Ghissassi F, Bouvard V, Benbrahim-Tallaa L, et al.; International Agency for Research on Cancer Monograph Working Group (2017). Carcinogenicity of benzene. *Lancet Oncol*, 18(12):1574–5. doi:10.1016/S1470-2045(17)30832-X PMID:29107678

Smith MT, Guyton KZ, Gibbons CF, Fritz JM, Portier CJ, Rusyn I, et al. (2016). Key characteristics of carcinogens as a basis for organizing data on mechanisms of carcinogenesis. *Environ Health Perspect*, 124(6):713–21. doi:10.1289/ehp.1509912 PMID:26600562

1. EXPOSURE DATA

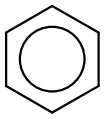
1.1 Identification of the agent

1.1.1 Nomenclature

Chem. Abstr. Serv. Reg. No.: 71-43-2
Primary name: benzene
IUPAC systematic name: benzene

1.1.2 Structural and molecular formulae, and relative molecular mass

Structural formula:



From O'Neil (2006) and Lide (2008)

Molecular formula: C_6H_6
Relative molecular mass: 78.1

1.1.3 Chemical and physical properties of the pure substance

From HSDB (2018)

Description: clear, colourless, volatile, highly flammable liquid

Boiling point: 80.1 °C

Melting point: 5.558 °C

Density: 0.8756 g/cm^3

Refractive index: 1.5011 at 20 °C

Solubility: slightly soluble in water (1.8 g/L at 25 °C); miscible with acetic acid, acetone, chloroform, ethyl ether, and ethanol

Viscosity: 0.604 mPa at 25 °C

Vapour pressure: 94.8 mmHg at 25 °C

Stability: benzene is a very stable molecule due to its aromaticity, that is, the delocalization of pi electrons in the benzene molecule creating a resonance; catalysts are often needed to make benzene undergo a chemical reaction; benzene is volatile with a boiling point of 80 °C, and is highly flammable

Flash point: –11.1 °C

Octanol/water partition coefficient: log K_{ow}, 2.13; conversion factor (20 °C, 101 kPa): 1 ppm = 3.19 mg/m^3.

1.1.4 Technical products and impurities

The impurities found in commercial products are toluene, xylene, phenol, thiophene, carbon disulfide, acetylnitrile, and pyridine. Thiophene-free benzene has been specially treated to avoid destroying the catalysts used in reactions with benzene. Refined nitration-grade benzene is free of hydrogen sulfide and sulfur dioxide (HSDB, 2018).

1.2 Production and use

1.2.1 Production

(a) Production process

Benzene was first isolated by Faraday in 1825 from a liquid condensed by compressing oil gas; Mitscherlich first synthesized it in 1833 by distilling benzoic acid with lime. Benzene was first recovered commercially from light oil derived from coal tar in 1849, and from petroleum in 1941 (IARC, 1982).

Benzene can be produced in several ways. One method is by catalytic reforming, which involves the dehydrogenation of cycloparaffins, dehydroisomerization of alkyl cyclopentanes, and the cyclization and subsequent dehydrogenation of paraffins. The feed to the catalytic reformer (platinum-rhenium on an alumina support of high surface area) for benzene is thermally cracked naphtha cut at 71–104 °C. The benzene product is most often recovered from the reformate by solvent extraction techniques (Fruscella, 2002).

Benzene can also be prepared by cracking, a multistep process where crude oil is heated, steam is added, and the gaseous mixture is then briefly passed through a furnace at temperatures of 700–900 °C. The dissolved compounds undergo fractional distillation, which separates out the different components, including benzene (Fruscella, 2002).

Alternatively, benzene can be prepared from toluene by hydrodealkylation. In the presence of a catalyst (chromium, molybdenum, and/or platinum), toluene and hydrogen are compressed to pressures of 20–60 atmospheres and the mixture is heated to temperatures of 500–660 °C. This reaction converts the mixture to benzene and methane, and benzene is separated out by distillation (Fruscella, 2002).

(b) Production volume

Benzene is listed as a high production volume chemical by the Organisation for Economic Co-operation and Development (OECD, 2009). In 2012, global benzene production was approximately 42.9 million tonnes. In the USA, production volumes during 1986–2002 were more than 1 billion pounds [> 450 000 tonnes] (HSDB, 2018). In order of volume produced, the five countries producing the greatest quantities of benzene in 2012 were China, the USA, the Republic of Korea, Japan, and Germany (Merchant Research & Consulting Ltd, 2014). In 2014, the industry reported benzene production and consumption in western Europe (Germany, Belgium, France, Italy, Luxembourg, the Netherlands, Denmark, Ireland, the United Kingdom, Greece, Spain, Portugal, Austria, Finland, and Sweden – the EU-15 – plus Norway and Switzerland) of 6.7 and 7.5 million tonnes, respectively (PetroChemicals Europe, 2015).

The use of benzene for the production of ethylbenzene, cumene, cyclohexane, and nitrobenzene accounts for 90% of annual benzene consumption. In order of volume consumed, China, the USA, and western Europe consume about half of the total benzene produced (IHS Markit, 2017).

The United States Environmental Protection Agency (EPA) report published in February 2017 (Report No. 17-P-0249) reports a total benzene consumption of 57 701 737 237 gallons (equivalent to 1.9×10^8 tonnes; 1 gallon = 3.7858 L, benzene density of 0.879 g/cm^3) for 84 facilities in the USA in 2014 (EPA, 2017).

1.2.2 Uses

Historically, benzene was used as a degreaser of metals, a solvent for organic materials, a starting and intermediate material in the chemical and drug industries (e.g. to manufacture rubbers, lubricants, dyes, detergents, and pesticides), and an additive to unleaded gasoline

(ATSDR, 2007; Williams et al., 2008; NTP, 2016). Benzene use has diminished since its carcinogenic properties became widely publicized (IARC, 1982); however, some countries have continued to use benzene in specific products such as glue (Vermeulen et al., 2004).

Benzene occurs naturally in petroleum products (e.g. crude oil and gasoline), and is also added to unleaded gasoline for its octane-enhancing and anti-knock properties. Typically, the concentration of benzene in these fuels is 1–2% by volume (ATSDR, 2007). Benzene concentration in fuels sold in the European Union must be less than 1.0% by volume (European Commission, 2009).

The percentage of benzene in gasoline has varied with the refinery and time period from which it originated. Until 1931, the benzene content of the gasoline imported into the United Kingdom was 1% v/v (Lewis et al., 1997). In 1971, Parkinson reported that gasoline in the United Kingdom contained 2.8–5.8% benzene v/v (Parkinson, 1971). In Canada in the 1970s and the 1980s, benzene content in fuel was reported as 0.7–3.7% (Armstrong et al., 1996); in Australia, benzene content of 1–5% by weight during 1950–1990 was reported (Glass et al., 2000).

Gasoline can be enriched with benzene by adding benzene-toluene-xylene, which is generated during coke making. Where necessary, side-stream petroleum is added to adjust the octane rating; for example, reformate includes 5–12% benzene (Glass et al., 2000). Before 1950, a small proportion of gasoline enriched with benzene sold in the United Kingdom included up to 36% benzene (Lewis et al., 1997). Gasoline enriched with benzene included up to approximately 10% benzene in Canada during 1914–1938 (Armstrong et al., 1996) and in Australia until around 1970 (Glass et al., 2000).

The primary use of benzene today is in the manufacture of organic chemicals. In Europe, benzene is mainly used to make styrene, phenol, cyclohexane, aniline, maleic anhydride, alkylbenzenes, and chlorobenzenes. It is an intermediate in the production of anthraquinone, hydroquinone, benzene hexachloride, benzene sulfonic acid, and other products used in drugs, dyes, insecticides, and plastics (ICIS, 2010). In the USA, the primary use of benzene is in the production of ethylbenzene, accounting for 52% of the total benzene demand in 2008. Most ethylbenzene is consumed in the manufacture of styrene, which is used in turn in polystyrene and various styrene copolymers, latexes, and resins. The second-largest use of benzene in the USA (accounting for 22% of demand) is in the manufacture of cumene (isopropylbenzene), nearly all of which is consumed in phenol production. Benzene is also used to make chemical intermediates, including cyclohexane, used in making certain nylon monomers (15%); nitrobenzene, an intermediate for aniline and other products (7%); alkylbenzene, used in detergents (2%); chlorobenzenes, used in engineering polymers (1%); and miscellaneous other uses (1%) (Kirschner, 2009).

1.3 Measurement and analysis

1.3.1 Detection and quantification

Common standard methods to assay benzene in air are presented in Table 1.1, along with selected methods for measuring some biomarkers of exposure in urine.

Assays to monitor benzene in air were first developed to measure air concentration in the workplace, including personal exposure of workers, and to assess compliance with occupational limits. Typically, to measure 8-hour exposure, air is pumped through cartridges containing charcoal or other suitable sorbents for the duration of the entire work shift. In the laboratory, benzene is desorbed from sorbent using solvents such as carbon disulfide (NIOSH, 2003, method 1501) or high-temperature thermal desorption (NIOSH, 1996, method 2549), and analysed with either a gas chromatograph equipped with a flame ionization

Table 1.1 Representative methods for the analysis of benzene in air and its main urinary biomarkers

Sample matrix	Analyte	Assay procedure	Limit of detection	Reference
Air	Benzene	Pumping air through solid sorbent tube, solvent desorption, and GC-FID	0.5 µg/sample (sample volume 5–30 L)	NIOSH (2003), method 1501
	Benzene	Pumping air through solid sorbent tube, thermal desorption, and GC-MS	100 ng per tube or less (sample volume 1–6 L)	NIOSH (1996), method 2549
	Benzene	Real-time monitor with FTIR detector	0.32 ppm for a 10 m absorption pathlength	NIOSH(2002), method 3800
	Benzene	Portable GC-PID	0.02 ppm	NIOSH (1994), method 3700
	Benzene	Passive sampling [with solid sorbent device], solvent/thermal desorption, and GC-MS	Variable depending on geometry of sampler and sampling time	EPA (2014)
Urine	t,t-MA	HPLC-UV analysis	5 µg/L	Lee et al. (2005)
	SPMA	SPE LC-MS/MS analysis	0.2 µg/L	NIOSH (2014), method 8326
	Benzene	HS GC-MS analysis	0.025 µg/L	Fustinoni et al. (1999)

FTIR, Fourier-transform infrared spectroscopy; FID, flame ionization detection; GC, gas chromatography; HPLC-UV, high-pressure liquid chromatography, ultraviolet spectroscopy; HS, head space; LC-MS/MS, liquid chromatography, tandem mass spectrometry; MS, mass spectrometry; PID, photoionization detector; SPE, solid phase extraction; SPMA, *S*-phenylmercapturic acid; t,t-MA, *trans*,*trans*-muconic acid

detector (NIOSH, 2003, method 1501) or a mass spectrometer (NIOSH, 1996, method 2549). As an alternative, passive samplers do not need a pump and allow benzene sampling via air diffusion through them; see EPA (2014) for a review of different assays using passive samplers for the determination of volatile organic compounds, including benzene. The sensitivity of both active and passive assays depends on sample volume, desorption method, and instrumental analysis; a higher sampling volume, the use of thermal desorption, and detection by mass spectrometer are associated with greater sensitivity (detection by mass spectrometer also offers high specificity). The design determines the sampling rate for passive samplers; radial geometry warrants a high flow rate and therefore larger sampling volume over a specific sampling time (Cocheo et al., 2000).

A real-time monitor can be used to check for benzene leaks and to measure short-term exposure, especially during critical operations, allowing the simultaneous sampling of air and detection of benzene. Benzene can be separated from other chemicals by portable gas chromatography and detected by photoionization detector (NIOSH, 1994, method 3700), or can be measured by extractive Fourier-transform infrared spectrometry (NIOSH, 2002, method 3800).

The alternative method of measuring benzene exposure by biomonitoring dates to the 1980s (Lauwerys, 1983); the first biomarkers, such as phenol, have been progressively abandoned in favour of biomarkers that are less abundant but more specific. The currently recommended biomarkers for assessment of benzene exposure in the workplace include urinary *trans*,*trans*-muconic acid (t,t-MA), urinary *S*-phenylmercapturic acid (SPMA), and urinary benzene (INRS, 2017).

t,t-MA is a urinary metabolite of benzene accounting for about 4% of the absorbed dose. Formed and excreted in urine with rapid kinetics with a half-life of about 5 hours (Boogaard & van

Sittert, 1995), it is useful for assessment of recent exposure. It is measured using high-performance liquid chromatography with an ultraviolet detector (Lee et al., 2005), and standardized assays are present on the market. Its limitation is poor specificity, as t,t-MA is also produced by the metabolism of the preservative sorbic acid or sorbates contained in food and beverages (Ruppert et al., 1997; Weaver et al., 2000). t,t-MA is recommended when exposure is higher than 0.2 ppm (Kim et al., 2006a), depending on the amount of sorbic acid preservatives in the diet.

SPMA is a urinary metabolite of benzene accounting for less than 1% of the absorbed dose; it is formed and excreted in urine with rapid kinetics (half-life of ~9 hours; Boogaard & van Sittert, 1995). SPMA in urine is a specific biomarker, and is assayed using solid phase extraction followed by liquid chromatography coupled with tandem mass spectrometry (NIOSH, 2014, method 8326). The limitations of the use of this biomarker are the few standardized assays available and the high cost of the equipment to perform the assay. The variability associated with genetic polymorphism of glutathione *S*-transferase enzymes also affects urinary levels of SPMA (see Section 4.1).

Unmetabolized benzene is excreted in urine in a tiny proportion (< 0.1%) and with rapid kinetics (a half-life of a few hours). It is a specific biomarker, being uniquely indicative of exposure to benzene. It is assayed using online headspace sampling followed by gas chromatography or mass spectrometry (Fustinoni et al., 1999). A limitation in the use of urinary unmetabolized benzene is the lack of standardized assays; in addition, the volatility of benzene in urine may cause the loss of the analyte if no precautions are taken during sampling and in the storage of samples.

Both SPMA and urinary benzene are currently the biomarkers of choice to assess exposure to benzene in studies involving the general population (Fustinoni et al., 2005; Lovreglio et al., 2011; Andreoli et al., 2015).

1.3.2 Assessment of occupational exposure in epidemiological studies

A variety of exposure assessment methods have been used in epidemiological studies of workers potentially exposed to benzene; methods are summarized in the following sections. Additional details on exposure assessment methods used in key epidemiological studies evaluated by the Working Group are provided in Section 1.6.

(a) Occupational cohorts compared with the general population

Many early studies of chemical and petroleum industry workers compared mortality and cancer incidence in the workers and in the general population (e.g. Decouflé et al., 1983; Consonni et al., 1999; Divine et al., 1999; Koh et al., 2014) in terms of either standardized mortality ratios and/or standardized incidence ratios. Benzene was known to be present at such facilities, but benzene exposure estimates were not provided and benzene may not have been specifically mentioned in such studies. Where benzene is mentioned, the metrics are usually expressed as exposed/not exposed, sometimes with the duration or era of the exposed job included. In all cases, there could have been individuals occupationally exposed to benzene in the general population (comparison group).

(b) Expert assessment using interviews, personal questionnaires, or job-specific modules

In occupational studies, some investigators have classified workers with respect to benzene exposure from questionnaires, including those that probe for specific determinants of exposure, such as job-specific modules (e.g. Reid et al., 2011). Benzene exposure may be categorized

semiquantitatively, for example, "no exposure" versus "probable exposure", or "high" versus "medium" versus "low" exposure (e.g. Adegoke et al., 2003; Black et al., 2004; Miligi et al., 2006; Krishnadasan et al., 2007; Seidler et al., 2007). The interpretation of such exposure categories varies from one study to another, depending on the era, country, and industry sectors evaluated, for example.

In population-based studies, exposure must be assessed across a range of occupations and industries by evaluating the type and duration of jobs reported by study participants.

(c) Expert assessment using job characteristics with no individual-level measurements

In some studies, experts classify workers within certain employment start-date periods, industry sectors, and/or job or task categories as exposed or not exposed to benzene (e.g. Koh et al., 2011; Linet et al., 2015). These experts are usually from the specific facility, or at least from the industry sector, and are often occupational hygienists. In most studies the exposure groupings appeared to be performed before case identification, for example in cohort studies, or the assessors were case-blind for case–control studies. This methodology can be used for cohort studies (Infante et al., 1977; Wong, 1987a; Koh et al., 2011), or in case–control studies (e.g. Wong et al., 2006). Duration of exposure is a common metric in these types of studies, and provides a semiquantitative dimension to the exposure assessment. The metrics commonly used in these analyses are exposure category (where provided) and duration of exposed job. Broad exposure groupings were based on employment structure in several studies, for example hourly (potentially higher risk of exposure) versus salaried (potentially lower risk of exposure) workers (e.g. Wen et al., 1983; Wongsrichanalai et al., 1989; Honda et al., 1995). Some similar exposure assessments have a semiquantitative element, for example providing an exposure dimension of high, medium, or low for the work area (McMichael et al., 1975; Rushton & Alderson, 1981).

(d) Exposure assessment using quantitative measurements grouped by job characteristics

The strongest exposure estimates are those where measured benzene exposure data from relevant facilities were attributed by experts to individual job titles or work areas (e.g. Dosemeci et al., 1994). Exposure data may have been collected on an industry- or cohort-wide basis and then applied to specific individual participants, notably in nested case–control studies. This methodology has been applied in China in population-based case–control studies (Bassig et al., 2015), where measured exposure data from many industries has been available since the 1950s (e.g. Wong et al., 2010; Friesen et al., 2012).

There will be some imprecision in the application of a (usually) limited number of data points to other individuals, perhaps employed at other facilities or over different timeframes. Exposure may vary between facilities, between workers, and between days for the same worker, regardless of how average exposure data are assigned. It is important to ensure that the measurement data are representative of usual exposure (normal working circumstances), and include jobs for which lower and higher levels of exposure have been measured. The exposure estimates are quantitative and usually expressed as averaged mean benzene intensity (ppm or mg/m^3) or cumulative exposure (ppm-years or (mg/m^3)-years). The exposure grouping may take into account measured exposure data from multiple sites across a range of industry sectors (e.g. Portengen et al., 2016).

Data on personal exposure to benzene were not usually available before 1970, so extrapolations back in time may be needed. Exposure modifiers, for example, historical changes in work processes, percentage of benzene in petrol, or the presence of ventilation, may have been

used to estimate exposure for jobs and for eras where measured data may not be available or applicable (Armstrong et al., 1996; Lewis et al., 1997; Glass et al., 2000). These exposures were usually estimated with the aid of occupational hygienists from within the industry, and are discussed in more detail in Section 1.6.1. Smith et al. (1993) used such methodology to estimate total hydrocarbon exposure, from which Wong et al. (1999) estimated benzene exposure.

1.3.3 Exposure assessment for molecular epidemiology

Several factors should be considered in the design of epidemiological mechanistic studies. These include the congruency in the time period of effect or disease onset relative to exposure, the magnitude of effects observed, and inter- and intraindividual variability in the response.

For studies on cancer, long-term average exposure is relevant. The latency for leukaemia can be relatively short, for example less than 10 years (Finkelstein, 2000; Richardson, 2008), so exposure during this period should be characterized.

Shorter periods of more recent exposure should be considered for other end-points such as leukopenia (Lan et al., 2004), or chromosomal aberrations (Zhang et al., 1998; Marchetti et al., 2012) including genetic damage (Liu et al., 1996; Zhang et al., 2016). To identify changes in leukocyte numbers, for example, exposure to benzene in the 180 days before blood collection is relevant (Ward et al., 1996).

In a cross-sectional study, it is important to collect both exposure and outcome data for the same individuals to account for inter- and intraindividual variability associated with relevant parameters, for example, diet, smoking, shift work, and time-of-day effects. Data describing these factors should be collected systematically and incorporated within the analyses.

In assessing the exposure, a sufficient number of participants are needed to account for the variability in uptake and human metabolism, particularly where the biomarker of effect is labile (e.g. oxidative stress). In addition, repeated measurements to estimate average exposure are advisable to account for day-to-day variability in exposure.

Investigators should use recognized and validated methods of collection and analysis, ensuring quality by taking into account the most relevant parameters, including the limit of detection.

1.4 Occurrence and exposure

1.4.1 Occupational exposure

Benzene is a ubiquitous pollutant that is present in several industries and occupations, including the production and refining of oil and gas, the distribution, sale, and use of petroleum products, coke production, the manufacture and use of chemical products, automobile repair, shoe production, firefighting, and various operations related to engine exhaust. Due to the high volatility of benzene, occupational exposure to benzene mainly occurs via inhalation. Benzene also penetrates skin, but the degree of dermal absorption of benzene will depend upon the exposure scenario. Dermal absorption will vary according to the tasks being performed (e.g. dipping machinery parts, immersion of hands, or using petroleum-based products as degreasing agents), the benzene content of the product, the composition of the product containing benzene, contact time, and the area of the body on which the chemical resides (Kalnas & Teitelbaum, 2000; Williams et al., 2011; Jakasa et al., 2015). In these scenarios, the exposure will not usually be to pure benzene.

The major industries and occupations in which workers are potentially exposed to benzene are reviewed in the following sections. This

summary is not exhaustive, and the interested reader is referred to several reviews of occupational exposure to benzene across industries that have been published for Europe and North America (Runion & Scott, 1985; Nordlinder & Ramnäs, 1987; van Wijngaarden & Stewart, 2003; Capleton & Levy, 2005; Williams et al., 2008) and Asia (Kang et al., 2005; Liang et al., 2005; Navasumrit et al., 2005; Liu et al., 2009; Park et al., 2015). For some industries or applications, information in the literature is limited. For example, the use of pure benzene as a solvent and reagent in chemical laboratories is well known, but no report on exposure level of benzene was found for laboratory technicians apart from in the petroleum industry.

Although not exhaustive, Table 1.2 gives a summary of reported personal full-shift airborne benzene concentrations, while Table 1.3 summarizes biomonitoring data for the industries.

(a) *Production, refining, and distribution of petroleum and petroleum-derived products*

The petroleum industry can be divided into upstream and downstream segments. The upstream segment refers to conventional exploration, extraction, and production of crude oil and natural gas, described in the following section, as well as unconventional oil and gas development (UOGD). UOGD involves high-volume hydraulic fracturing, commonly referred to as "fracking", which is coupled to (vertical or horizontal) drilling to extract oil and gas from shale formations (i.e. extraction of materials other than crude oil and natural gas). UOGD includes the process of injecting large volumes of water, proppants (often sand), and potentially hazardous chemicals into wellbores at high pressure, fracturing the rock and enabling the outflow of trapped oil or gas from shale formations (EPA, 2013). The downstream segment consists of refinery operations (production and ancillary operations within the refinery and distribution depots, e.g. tank dipping, pump repairs, filter cleaning), distribution (loading of ships, railcars and road tankers, delivery to service stations), and retail of the petroleum fractions (attendant or self-service filling of customer vehicles).

(i) *Upstream petroleum industry (conventional oil and gas extraction)*

During drilling, the revolving steel bit must be lubricated and cooled, the well requires pressure support, and the rock cuttings must be transported to the surface. Drilling fluid, a complex oil- or water-based mixture, is used for these purposes. The characteristics of the hydrocarbon base oils in the drilling fluids have changed over time. Diesel as a base oil for drilling was gradually replaced in the early 1980s in the United Kingdom and Norway by petroleum-mineral oils with a reduced aromatic content; non-aromatic mineral oils (aromatic content < 0.01%) were used after 1998 (Gardner, 2003; Steinsvåg et al., 2006, 2007; Bråtveit et al., 2012). The mud-handling areas were originally designed for water-based mud that did not generate vapours, with open flow lines and mud pits. Other than measurements of oil mist and oil vapour, there have been very limited attempts to characterize the exposure regarding its composition. Theoretically, however, hydrocarbon and benzene exposure can occur through contamination of the drilling fluid from the geological formation in which it is drilled, or from hydrocarbons that are added to the drilling fluid to improve drilling properties, as in diesel and drilling fluids containing aromatics in the 1980s (Verma et al., 2000; Steinsvåg et al., 2007). With the exception of eight area measurements made during drilling in Canada showing a full-shift concentration of 0.006 mg/m^3 (with a highest measurement of 0.019 mg/m^3 and one personal measurement of < 0.010 mg/m^3), no information on this exposure scenario was available (Verma et al., 2000).

Table 1.2 Occupational exposure to benzene in air: personal measurements

Reference	Location, collection year	Occupational description, setting	*n*	Sampling time	Exposure concentration (mg/m³)[a]	Exposure range (mg/m³)[a]	Comments/additional data[a]
Upstream petroleum industry, unconventional							
Esswein et al. (2014)	USA, 2013	Flowback operations, workers gauging tanks	17	Full shift (typically 12 h)	AM (SD), 0.25 (0.16) ppm [0.8 (0.51)]	0.01–0.37 ppm [0.032–1.18]	Task-based short-term (2.5–30 min)
		Flowback operations, workers not gauging tanks	18	Full shift (typically 12 h)	AM (SD), 0.04 (0.03) ppm [0.13 (0.096)]	0.004–0.05 ppm [0.013–0.16]	
Upstream petroleum industry, conventional							
Bråtveit et al. (2007)	Norway, 2005	Process operators	35	657 min (range, 450–730 min)	AM (SD), 0.042 (0.132) ppm [0.13 (0.42)] GM, 0.005 ppm [0.016]	< 0.001–0.69 ppm [< 0.003–2.2]	Exposure varied according to tasks performed
		Flotation work	6		AM (SD), 0.221 (0.267) ppm [0.71 (0.85)] GM, 0.114 ppm [0.360]	0.030–0.688 ppm [0.095–2.2]	
		Sampling	11		AM (SD), 0.005 (0.005) ppm [0.16 (0.16)] GM, 0.003 ppm [0.096]	< 0.001–0.014 ppm [< 0.003–0.04]	
		Miscellaneous	18		AM (SD), 0.005 (0.01) ppm [0.16 (0.03)] GM, 0.003 ppm [0.096]	< 0.0010.023 ppm [< 0.003–0.07]	
Steinsvåg et al. (2007)	Norway, 1994–2003	Process and drilling operations (12 installations)	367	12 h	AM (SD), 0.037 (0.099) ppm [0.12 (0.32)] GM (GSD), 0.007 (5.7) ppm [0.22 (18.21)]	< LOD–2.6 ppm [< LOD–8.31]	165 measurements < LOD were set to LOD/√2

Table 1.2 (continued)

Reference	Location, collection year	Occupational description, setting	*n*	Sampling time	Exposure concentration (mg/m³)[a]	Exposure range (mg/m³)[a]	Comments/additional data[a]
Steinsvåg et al. (2007) (cont.)		Deck workers	29		AM (SD), 0.17 (0.51) ppm [0.54 (1.63)] GM (GSD), 0.010 (14) ppm [0.03 (44.7)]	< LOD–2.6 ppm [< LOD–8.31]	> LOD = 10
		Process operators	204		AM (SD), 0.036 (0.097) ppm [1.15 (0.31)] GM (GSD), 0.008 (5.3) ppm [0.026 (16.93)]	< LOD–0.97 ppm [< LOD–3.1]	> LOD = 101
		Laboratory	40		AM (SD), 0.012 (0.019) ppm [0.038 (0.06)] GM (GSD), 0.006 (3.7) ppm [0.019 (11.82)]	< LOD–0.11 ppm [< LOD–0.35]	> LOD = 13
		Mechanics	78		AM (SD), 0.006 (0.011) ppm [0.019 (0.035)] GM (GSD), 0.002 (4.5) ppm [0.006 (14.37)]	< LOD–0.08 ppm [< LOD–0.26]	> LOD = 37
		Electricians	16		AM (SD), 0.015 (0.017) ppm [0.048 (0.05)] GM (GSD), 0.007 (5.7) ppm [0.019 (18.85)]	< LOD–0.05 ppm [< LOD–0.16]	> LOD = 4
Kirkeleit et al. (2006b)	Norway, 2004	Crude oil production, vessel	139	592 min (range, 43–931 min)	AM, 0.43 ppm [1.37] GM (GSD), 0.02 (12.42) ppm [0.06 (39.7)]	< 0.001–16.75 ppm [< 0.003–53.5]	LOD, 0.001 ppm [0.003]

Table 1.2 (continued)

Reference	Location, collection year	Occupational description, setting	*n*	Sampling time	Exposure concentration (mg/m³)[a]	Exposure range (mg/m³)[a]	Comments/additional data[a]
Kirkeleit et al. (2006b) (cont.)		Process operators	30	669 min (range, 182–915 min)	AM, 0.39 ppm [1.25] GM (GSD), 0.01 (9.68) ppm [0.03 (30.92)]	< 0.001–7.3 ppm [< 0.003–23.32]	
		Deck workers	47	564 min (range, 43–866 min)	AM, 0.89 ppm [2.84] GM (GSD), 0.02 (19.11) ppm [0.06 (61.04)]	< 0.001–16.75 ppm [< 0.003–53.5]	The high exposure levels represent cleaning and maintenance of crude oil cargo tanks
		Mechanics	31	632 min (range, 257–705 min)	AM, 0.07 ppm [0.22] GM (GSD), 0.007 (12.04) ppm [0.02 (38.4)]	< 0.001–0.51 ppm [< 0.003–1.63]	
		Contractors	31	518 min (range, 190–931 min)	AM, 0.11 ppm [0.35] GM (GSD), 0.05 (4.90) ppm [0.16 (15.65)]	< 0.001–0.42 ppm [< 0.003–1.34]	
Verma et al. (2000)	Canada, 1985–1996	Conventional oil/gas	198	Long-term	AM, 0.206 GM, 0.036	0.003–7.78	For occupational groups see paper
		Conventional gas	838		GM, 0.010	0.006–57.6	
		Pipeline	8		AM, 0.392 GM, 0.350	0.16–1.54	
		Heavy oil processing	236		AM, 0.112 GM, 0.051	< 0.003–1.60	
Oil spill clean-up operations							
Gjesteland et al. (2017)	Norway, 2016	Sampling boats	21	10.8 h (range, 5.2–14.3 h)	AM, 0.43 ppm [1.37] GM (GSD), 0.20 (4.52) ppm [0.64 (14.44)]	0.01–1.52 ppm [0.03–4.86]	Field trial with spill of two fresh oils (22 workers, > 2 d)
		Workers on release ship and oil recovery ship	11	9.8 h (range, 5.2–12.5 h)	AM, 0.05 ppm [0.16] GM (GSD), 0.02 (0.02) ppm [0.064 (0.064)]	0.002–0.10 ppm [0.006–0.32]	
Downstream, petroleum refinery industry							
Almerud et al. (2017)	Sweden, 2009–2011	Process technicians, refinery I	132	Full shift (8 or 12 h)	AM, 0.153	95% CI, 0.01–0.022; maximum, 3.77	
		Maintenance workers, refinery I	67		AM, 0.0059	95% CI, 0.004–0.009; maximum, 1.32	

Table 1.2 (continued)

Reference	Location, collection year	Occupational description, setting	*n*	Sampling time	Exposure concentration (mg/m³)[a]	Exposure range (mg/m³)[a]	Comments/additional data[a]
Almerud et al. (2017) (cont.)		Process technicians outdoor, refinery II	66		AM, 0.0137	95% CI, 0.0083–0.023; maximum, 0.27	
		Laboratory workers, refinery I	25		AM, 0.0046	95% CI, 0.0034–0.0062; maximum, 0.0154	
		Laboratory workers, refinery II	11		AM, 0.0084	95% CI, 0.0034–0.021; maximum, 0.02	
Akerstrom et al. (2016)	Sweden, 2011–2013	Turnarounds, refinery I	43	Full shift (8 or 12 h)	AM, 0.61	95% CI, 0.23–1.60 µg/m³	
		Turnarounds, refinery II	26		AM (SD), 0.96 (1.3) GM (GSD), 0.23 (0.0075)	0.007–4.5	
		Oil harbour workers (jetty workers and dock workers)	34		AM, 0.31	95% CI, 0.08–1.2	
		Sewage tanker drivers	16		AM, 0.36	95% CI, 0.068–1.9	
Widner et al. (2011)	USA, 1977–2005	Refinery and dock workers	406	480–661 min	NR	0.006–15 ppm [0.19–47.9]	GM not calculated because > 50% of measurements < LOD
		Dock connecting crew	179	535–664 min	GM (GSD), 0.023 (11) ppm [0.073 (35.1)]	0.010–15 ppm [0.03–47.9]	
		Contractor–tankerman	38	326–463 min	GM (GSD), 0.25 (8.8) ppm [0.8 (28.1)]	0.010–9.8 ppm [0.03–31.3]	
Kreider et al. (2010)	USA, 1977–2006	Routine operation, all areas and job titles	624	> 180 min	AM, 0.091 ppm [0.29]	Minimum–maximum detected, 0.004–6.0 ppm [0.013–19.2] 75th, 95th percentile, 0.043, 0.31 ppm [0.14–0.99]	GM not calculated because > 50% of measurements < LOD

Table 1.2 (continued)

Reference	Location, collection year	Occupational description, setting	*n*	Sampling time	Exposure concentration (mg/m³)[a]	Exposure range (mg/m³)[a]	Comments/additional data[a]
Kreider et al. (2010) (cont.)		Start up	50		AM, 0.046 ppm [0.15]	Minimum–maximum detected, 0.015–0.29 ppm [0.048–0.93] 75th, 95th percentile, 0.05, 0.17 ppm [0.16, 0.54]	
		Turnaround	471		AM, 0.17 ppm [0.54] GM (GSD), 0.032 (6.7) ppm [0.1 (21.4)]	Minimum–maximum detected, 0.004–9.200 ppm [0.013–29.4] 75th, 95th percentile, 0.12, 0.68 ppm [0.38, 2.17]	
CONCAWE (2002)	Europe, 1999–2001	Offsite refinery operator	6	451–498 min	AM, 0.3 GM, 0.2	10–90th percentiles, 0.1–0.5	
		Laboratory technician blending test gasoline for research	7	215–487 min	AM, 3.7 GM, 1.6	10–90th percentiles, 0.2–8.3	
CONCAWE (2000)	Europe, 1993–1998	Onsite operators (including catalytic reformer, gasoline blending)	97	Full shift	AM, 0.22	0.008–7.88	91% corresponding to full shift (8 or 12 h)

Table 1.2 (continued)

Reference	Location, collection year	Occupational description, setting	*n*	Sampling time	Exposure concentration (mg/m³)[a]	Exposure range (mg/m³)[a]	Comments/additional data[a]
CONCAWE (2000) (cont.)		Refinery offsite operators (tank farm, including dipping, sampling, valve operation, dewatering, loading rail cars)	321		AM, 0.32	0.008–23.3	
		Refinery maintenance workers (pump maintenance, instrument calibration, enclosed equipment)	373		AM, 0.41	0.008–18.1	
		Refinery laboratory technicians (including product analysis, octane rating testing)	628		AM, 0.30	0.0015–5.0	
		Tank cleaners (including sludge cleaning)	49		AM, 2.10	0.008–38.7	
Downstream, distribution							
Lovreglio et al. (2016)	Italy, NR	Fuel tanker drivers	17	8 h	AM (SD), 0.28 (0.248) Median, 0.246	0.0074–1.017	
CONCAWE (2002)	Europe, 1999–2001	Rail car operators, top loading with vapour recovery)	21	64–363 min	AM, 0.5 GM, 0.4	10–90th percentiles, 0.2–0.7	
		Rail car operators, top loading without vapour recovery	16	165–450 min	AM, 4.0 GM, 1.4	10–90th percentiles, 0.3–10	
CONCAWE (2002)	Europe, 1999–2001	Road tanker distribution; drivers, bottom loading with vapour recovery	33	185–555 min	AM, 0.6 GM, 0.4	10–90th percentiles, 0.2–1.2	Pre-2000 specification gasoline

Table 1.2 (continued)

Reference	Location, collection year	Occupational description, setting	*n*	Sampling time	Exposure concentration (mg/m^3)[a]	Exposure range (mg/m^3)[a]	Comments/additional data[a]
CONCAWE (2000)	Europe, 1993–1998	Marine and rail loading; ship deck crew, open loading	41	Full shift	AM, 0.56	0.08–5.4	91% corresponding to full shift (8 or 12 h)
		Ship deck crew, closed loading	2		AM, 0.56	0.51–0.6	
		Ship deck crew, unloading	32		AM, 0.51	0.023–3.7	
		Jetty staff	46		AM, 0.37	0.023–1.7	
CONCAWE (2000)	Europe, 1993–1998	Road tanker distribution Road tanker drivers, top loading	69	Full shift	AM, 2.07	0.04–48.2	
		Road tanker drivers, bottom loading (without vapour recovery)	223		AM, 0.82	0.008–15	
		Road tanker drivers, bottom loading (with vapour recovery)	137		AM, 0.37	0.03–1.99	
		Drivers, other category or unspecified	56		AM, 1.26	0.07–19.2	
		Road tanker terminal rack operators	126		AM, 0.64	0.003–4.2	
		Road tanker terminal supervisors/ operators	151		AM, 0.36	0.001–3.1	
		Road tanker terminal maintenance	52		AM, 0.52	0.001–7.9	
Foo (1991)	Singapore	Petroleum delivery tanker drivers	14	Full shift	AM, 1.10 ppm [3.51] GM, 0.81 ppm [2.59]	0.08–2.37 ppm [0.26–7.57]	21 gasoline stations Short-term exposure ($n = 7$): AM, 3.1 ppm [9.9], range 0.07–11.6 ppm [0.22–37.05]

Table 1.2 (continued)

Reference	Location, collection year	Occupational description, setting	*n*	Sampling time	Exposure concentration (mg/m³)[a]	Exposure range (mg/m³)[a]	Comments/additional data[a]
Petrochemical manufacturing							
Sahmel et al. (2013)	USA, 1974–1999	Use of petroleum-based raw materials;	2359	8 h	AM (SD), 0.54 (5.0) ppm [1.72 (15.97)] Median, 0.042 ppm [0.13]	NR	
		Routine employee exposure (all) 1974–1986	1289		AM (SD), 0.885 (6.72) ppm [2.83 (21.47)] Median, 0.12 ppm [0.38]	NR	Median exposure during time periods corresponding to year when OEL changed
		1987–1999	1070		AM (SD), 0.125 (0.676) ppm [0.4 (2.16)] Median, 0.016 ppm [0.051]	NR	
		1974–1983	916		AM (SD), 1.103 (7.739) ppm [3.52 (24.72)] Median, 0.19 ppm [0.61]	NR	Median exposure during time periods stratified according to key process changes
		1984–1991	865		AM (SD), 0.206 (2.024) ppm [0.66 (6.47)] Median, 0.01 ppm [0.03]	NR	
		1992–1999	578		AM (SD), 0.148 (0.578) ppm [0.47 (1.85)] Median, 0.021 ppm [0.067]	NR	
Williams & Paustenbach (2005)	USA, 1976–1987	Petrochemical manufacturing facility (acetic acid); mainly process operators	749	4–10 h	AM (SD), 1.75 (3.8) ppm [5.59 (12.14)]	NR	See paper for mean exposure levels for various production processes/areas
Coke production							
He et al. (2015)	China, NR	Topside, plant A	27	8 h	AM (SD), 0.705 (0.259)	0.268–1.197	Plant A: top charging of coal; no air pollution control
		Topside, plant B	28		AM (SD), 0.290 (0.11)	0.085–0.489	Plant B: stamp charging of coal; bag house for air pollution control

Table 1.2 (continued)

Reference	Location, collection year	Occupational description, setting	*n*	Sampling time	Exposure concentration (mg/m³)[a]	Exposure range (mg/m³)[a]	Comments/additional data[a]
Bieniek & Łusiak (2012)	Poland, 2005–2010	Cokery workers	202	Full shift	0.15	0.01–1.79	
		Coke oven workers	122		Median, 0.16	5–95th percentile, 0.04–0.60	
		Coke by-products	37		Median, 0.37	5–95th percentile, 0.061–1.39	
		Other workers in the coke plant	43		Median, 0.09	5–95th percentile, 0.011–0.292	Electricians and supervising personnel
Kivistö et al. (1997)	Estonia, 1994	Cokery workers	18	Full shift	AM (SD), 1.3 (2.7) ppm [4.15 (8.62)] Median, 0.4 ppm [1.28]	0.09–11.7 ppm [0.29–37.37]	
		Benzene factory workers	20		AM (SD), 1.6 (3.3) ppm [5.11 (10.54)] Median, 0.6 ppm [1.92]	0.06–14.7 ppm [0.19–46.96]	
Drummond et al. (1988)	UK, 1986	Battery workers	NR	Full shift	AM, 0.31 ppm [0.99]	NR	Each worker measured for 3–5 consecutive shifts
		Refining process of benzene	NR		AM, 1.32 ppm [4.22]	Maximum, 4.3 ppm [13.74]	
Petrol stations							
Campo et al. (2016)	Italy, 2008–2009	Petrol station attendants	89	~5 h	Median, 0.059	5–95% CI, 0.005–0.284	
Lovreglio et al. (2016)	Italy, NR	Filling station attendants	13	8 h	AM (SD), 0.02 (0.015) Median, 0.0138	0.0045–0.0534	
Lovreglio et al. (2014)	Italy, NR	Filling station attendants	24	8 h	AM (SD), 0.023 (0.017) Median, 0.02	0.0045–0.0663	
Bahrami et al. (2007)	Islamic Republic of Iran, NR	Petrol station workers	25	2–4 h	AM (SD), 1.40 (0.80) ppm [4.47 (2.56)]	0.2–3.1 ppm [0.64–9.9]	
Navasumrit et al. (2005)	Thailand, NR	Petrol station attendants	50	8 h	AM (SD), 121.67 (14.37) ppb [0.39 (0.046)] GM, 86.4 ppb [0.28]	2.80–439.9 ppb [0.0089–1.42]	
CONCAWE (2002)	Europe, 1993–1998	Service station attendants, without vapour recovery	26	189–465 min	AM, 0.3 GM, 0.3	10–90th percentile, 0.2–0.5	Pre-2000 specification gasoline

Table 1.2 Occupational (continued)

Reference	Location, collection year	Occupational description, setting	*n*	Sampling time	Exposure concentration (mg/m³)[a]	Exposure range (mg/m³)[a]	Comments/additional data[a]
CONCAWE (2002) (cont.)		Service station attendants, with vapour recovery	7	288–437 min	AM, 0.1 GM, 0.1	10–90th percentile, 0.1–0.1	
		Service station cashiers	13	235–490 min	AM, 0.2 GM, 0.2	0.1–0.2	
		Service station workers, miscellaneous	6	237–280 min	AM, 0.2 GM, 0.1	0.1–0.2	
CONCAWE (2000)	Europe, 1993–1998	Service station attendants, without vapour recovery	417	Full shift	AM, 0.25	0.001–1.9	91% corresponding to full shift (8 or 12 h)
		Service station cashiers	268		AM, 0.05	0.001–1.92	
		Petrol pump maintenance workers	2		AM, 0.55	0.16–0.93	
		Service station workers, miscellaneous	5		AM, 0.03	0.01–0.10	
Lagorio et al. (1994)	Italy, 1991–1992	Petrol station attendants	27	8 h	AM (SD), 1.73 (5.53)	NR	Alkylated and lead-free gasoline: 2.86% and 2.65% benzene by volume, respectively
Lagorio et al. (1993)	Italy, 1992	Filling station attendants	111	8 h	AM (SD), 0.55 (2.46) GM (GSD), 0.12 (3.82)	0.001–28.02	111 filling stations
Foo (1991)	Singapore, NR	Gasoline kiosk attendants	54	Full shift	AM, 0.20 ppm [0.64] GM, 0.16 ppm [0.51]	0.028–0.71 ppm [0.89–2.27]	21 gasoline stations Short-term exposure (n = 49): AM, 6.6 ppm [21.08]; GM, 1.0 ppm [3.19]; range, 0.064–179 ppm [0.20–571.78]
Runion & Scott (1985)	USA, 1978–1783	Retail service stations	1478	Full shift	AM (SD), 0.06 (0.02) ppm [0.19 (0.06)] GM (GSD), 0.02 (5.4) ppm [0.06 (17.25)]	< 1.0 to > 10 ppm [< 3.19 to > 31.9] Range, NR	

Table 1.2 (continued)

Reference	Location, collection year	Occupational description, setting	*n*	Sampling time	Exposure concentration (mg/m³)[a]	Exposure range (mg/m³)[a]	Comments/additional data[a]
Automobile repair							
Egeghy et al. (2002)	USA, 1998–1999	Mechanics	197	4 h	AM (SD), 0.118 (0.166) Median, 0.0597	< 0.009–1.14	Self-administered sampling; benzene content of gasoline < 1%
Javelaud et al. (1998)	France, 1996	Mechanics	65	8 h	AM (SD), 0.48 (1.49) GM, 0.06 Median, 0.14	< 0.005–9.31	23 garages
Hotz et al. (1997)	Country NR, 1994–1995	Mechanics	156	8 h	Median, 0.01 ppm [0.032]	5–95th percentile, < LOD–0.14 ppm [< LOD–0.45]	
Foo (1991)	Singapore, NR	Motorcar service mechanics	54	Full shift	AM, 0.17 ppm [0.54] GM, 0.10 ppm [0.32]	0.014–1.7 ppm [0.045–5.43]	21 gasoline stations
Nordlinder & Ramnäs (1987)	Sweden, NR	Mechanics, small garage (summer)	> 100	Full shift	AM, 1.6	NR	
		Mechanics, small garage (winter)			AM, 6.8	NR	
		Mechanics, medium and large garages (summer)			AM, 0.4	NR	
		Electricians, medium and large garages (summer)			AM, 1.0	NR	
		Mechanics, medium and large garages (winter)			AM, 0.8	NR	
		Electricians, medium and large garages (winter)			AM, 1.4	NR	
Exposure from engine exhaust							
Arayasiri et al. (2010)	Thailand, 2006	Traffic police	24	8 h	AM (SD, 0.0382 (0.0027) Median, 0.039	(0.0155–0.069)	
		Office police	24		AM (SE), 0.007 (0.0005) Median, 0.0062	0.0036–0.014	
Manini et al. (2008)	Italy, 2005	Traffic police	19	6 h	AM, 0.0061	0.0003–0.012	

Table 1.2 (continued)

Reference	Location, collection year	Occupational description, setting	*n*	Sampling time	Exposure concentration (mg/m³)[a]	Exposure range (mg/m³)[a]	Comments/additional data[a]
Bahrami et al. (2007)	Islamic Republic of Iran, NR	Taxi drivers	60	2–4 h	AM (SD), 0.31 (0.22) ppm [0.99 (0.7)]	0.07–0.95 ppm [0.22–3.03]	
Manini et al. (2006)	Italy, 2004	Taxi drivers	37	24 h	AM (SD), 0.006 (0.0017)	NR	Non-smokers, ambient concentration in taxi during the 12 h shift, 0.0075 (0.0019)
Crebelli et al. (2001)	Italy, 1998–1999	Traffic police	139	7 h	AM (SD), 0.009 (0.011) GM (GSD), 0.0068 (0.002)	0.0013–0.0767	
		Office police	63		AM (SD), 0.0038 (0.0015) GM (GSD), 0.0035 (0.0015)	0.0011–0.0083	
Fustinoni et al. (1995)	Italy, 1994	Traffic wardens, urban and outdoors	20	5 h	AM (SD), 0.053 (0.03)	0.02–0.108	
		Traffic wardens, indoors (clerks)	19		AM (SD), 0.029 (0.008)	0.017–0.044	
Navasumrit et al. (2005)	Thailand, NR	Cloth vendors	22	8 h	AM (SD), 22.61 (1.32) ppb [0.073 (0.004)] Median, 21.1 ppb [0.067]	13.9–40.7 ppb [0.044–0.13]	
		Grilled-meat vendors	21		AM (SD), 28.19 (2.23) ppb [0.09 (0.007)] Median, 24.61 ppb [0.078]	16.8–52.0 ppb [0.054–0.17]	
Shoemaking							
Azari et al. (2012)	Islamic Republic of Iran, NR	Shoemakers, 12 workshops (October)	48	8 h	Mean (SE), 1.10 (0.11) ppm [3.51 (0.35)]	NR	Three consecutive months (October–December), examined effects of climate change and restriction of air flow due to closure of windows and shutdown of general ventilation systems
		Shoemakers, 12 workshops (November)			Mean (SE), 1.37 (0.14) ppm [4.38 (0.45)]	NR	
		Shoemakers, 12 workshops (December)			Mean (SE), 1.52 (0.18) ppm [4.86 (0.57)]	NR	

Table 1.2 (continued)

Reference	Location, collection year	Occupational description, setting	*n*	Sampling time	Exposure concentration (mg/m^3)[a]	Exposure range (mg/m^3)[a]	Comments/additional data[a]
Estevan et al. (2012)	Spain, 2002–2007	Shoemakers: 2002–2003	329	NR	AM (SD), 0.05 (0.15)	NR	18–26% of samples were ≥ LOD (0.01)
		Shoemakers: 2004–2005	218		AM (SD), 0.07 (0.14)	NR	
		Shoemakers: 2006–2007	302		AM (SD), 0.05 (0.14)	NR	
Zhang et al. (2011)	China, NR	Shoemakers	44	8 h	AM (SD), 44.81 (33.59) GM (GSD), 27.91 (3.29)	2.57–146.11	
Vermeulen et al. (2004)	China, 2000–2001	Large shoe factory (safety shoes): all workers	2667	8 h	AM, 3.46 ppm [11.05] GM (GSD), 1.28 (3.64) ppm [4.09 (11.63)]	10–90th percentiles, 0.20–7.00 ppm [0.64–22.4]	No glues reported to contain benzene
		Large shoe factory (safety shoes): cutting	427		AM, 0.45 ppm [1.44] GM (GSD), 0.34 (2.05) ppm [1.09 (6.55)]	0.17–0.15 ppm [0.54–3.67]	
		Large shoe factory (safety shoes): modelling	735		AM, 2.74 ppm [8.75] GM (GSD), 1.71 (2.81) ppm [5.46 (8.98)]	0.38–6.04 ppm [1.21–19.29]	
		Large shoe factory (safety shoes): fitting	1096		AM, 2.19 ppm [7] GM (GSD), 1.12 (2.98) ppm [3.58 (9.52)]	0.26–4.68 ppm [0.83–14.95]	
		Large shoe factory (safety shoes): finishing	241		AM, 8.35 ppm [26.67] GM (GSD), 2.91 (3.33) ppm [9.3 (10.64)]	0.65–11.69 ppm [2.08–37.34]	
		Large shoe factory (safety shoes): packing	168		AM, 15.55 ppm [49.67] GM (GSD), 7.60 (3.47) ppm [24.28 (11.08)]	1.43–43.06 ppm [4.57–137.55]	
		Small shoe factory (luxury shoes): all workers	116	8 h	AM, 21.86 ppm [69.83] GM (GSD), 14.4 (2.31) ppm [46 (7.38)]	10–90th percentiles, 5.23–50.63 ppm [16.71–161.73]	6 of 7 glues contained benzene (0.60–34% benzene)

Table 1.2 (continued)

Reference	Location, collection year	Occupational description, setting	*n*	Sampling time	Exposure concentration (mg/m³)[a]	Exposure range (mg/m³)[a]	Comments/additional data[a]
Vermeulen et al. (2004) (cont.)		Small shoe factory (luxury shoes): cutting	41		AM:10.96 ppm [35.01] GM (GSD), 10.24 (1.45) ppm [32.71 (4.63)]	6.53–16.26 ppm [20.86–51.94]	
		Small shoe factory (luxury shoes): modelling	18		AM, 9.04 ppm [28.88] GM (GSD), 7.75 (1.75) ppm [24.76 (5.59)]	4.45–18.69 ppm [14.21–59.7]	
		Small shoe factory (luxury shoes): fitting	47		AM, 29.31 ppm [93.62] GM (GSD), 21.34 (2.34) ppm [68.17 (7.47)]	7.06–65.17 ppm [22.55–208.17]	
		Small shoe factory (luxury shoes): finishing	10		AM, 54.64 ppm [174.54] GM (GSD), 28.03 (3.30) ppm [89.54 (10.54)]	7.62–179.60 ppm [24.34–573.69]	
Printing							
Portengen et al. (2016)	China, 1949 to after 2000	Printing	232	NR	AM, 94.1 GM (GSD), 8.2 (13.0)	NR	40% of measurements < LOD (3.19)
Kang et al. (2005)	Republic of Korea, 1992–2000	Offset printing	4	NR	AM (SD), 0.017 (0.012) ppm [0.0543 (0.038)] GM (GSD), 0.014 ppm [0.0447]	0.008–0.034 ppm [0.0255–0.11]	
Handling of jet fuel							
Smith et al. (2010)	USA, NR	US Air Force personnel					LOD, 0.9 µg/m³ Jet fuel JP-8 (0.004–0.007% benzene)
		All	69	Full shift	GM (GSD), 0.0016 (0.0035)	< LOD–0.0364	
		Group assumed exposed to concentration	25	Full shift	GM (GSD), 0.0029 (0.0034)	< LOD–0.0364	
Egeghy et al. (2003)	USA, NR	US Air Force personnel					Jet fuel JP-8 (0.0002–0.0123 weight% benzene)
		Group exposed to low concentration	140	4 h	Median, 0.0031	< 0.001–0.0613	
		Group exposed to moderate concentration	38	4 h	Median, 0.0074	0.0014–1.85	
		Group exposed to high concentration	114	4 h	Median, 0.252	0.0061–6.63	

Table 1.2 (continued)

Reference	Location, collection year	Occupational description, setting	*n*	Sampling time	Exposure concentration (mg/m³)[a]	Exposure range (mg/m³)[a]	Comments/additional data[a]
Holm et al. (1987)	Sweden, 1983–1984	Swedish National Defence	92	12 h	GM (GSD), 0.06 (4.0)	Maximum, 7.2	Jet fuel MC-77 (equivalent to JP4 (< 1% benzene)
		All samples	46	8 h (TWA)	GM (GSD), 0.06 (4.1)	Maximum, 4.1	
		Jet fuel handling	6		GM (GSD), 0.03 (2.6)	Maximum, 0.1	
		Flight service	28		GM (GSD), 0.08 (3.6)	Maximum, 1.2	
		Workshop service	12		GM (GSD), 0.05 (5.7)	Maximum, 4.1	
		Pure jet fuel exposure	38		GM (GSD), 0.06 (4.2)	Maximum, 4.1	
		Mixed solvents exposure	8		GM (GSD), 0.11 (3.1)	Maximum, 0.5	
Firefighting							
Reinhardt & Ottmar (2004)	USA, 1992–1995	Initial attack (full shift)	45	13.3 h (range, 12–18 h)	GM, 3 ppb [0.096]	Maximum, 24 ppb [0.077]	13 d of initial attack incidents
		Initial attack (at fires)		3.3 h (range, 2–10 h)	GM, 14 ppb [0.045]	Maximum, 43 ppb [0.14]	
		Project wildfires (full shift)	84	13.9 h (range, 4–24 h)	GM, 4 ppb [0.013]	Maximum, 249 ppb [0.8]	17 d at eight separate project wildfires
		Project wildfires (at fires)		10.4 h (range, 2–24 h)	GM, 6 ppb [0.019]	Maximum, 384 ppb [1.23]	
		Prescribed burns (full shift)	200	11.5 h (range, 6–18 h)	GM, 16 ppb [0.051]	Maximum, 58 ppb [0.19]	39 prescribed burns
		Prescribed burns (at fires)		7 h (range, 2–13 h)	GM, 28 ppb [0.089]	Maximum, 88 ppb [0.28]	
Austin et al. (2001)	Canada, NR	Structural fires	9	Short-term	AM (SD), 3.38 (3.45) ppm [10.8 (11.02)]	0.12–10.76 ppm [0.38–34.37]	Area samples (not personal)
Bolstad-Johnson et al. (2000)	USA, 1998	Structural fires	95	Short-term	AM (SD), 0.383 (0.425) ppm [1.22 (1.36)]	0.07–1.99 ppm [0.22– 6.36]	25 fires

AM, arithmetic mean; CI, confidence interval; d, day(s); GM, geometric mean; GSD, geometric standard deviation; h, hour(s); LOD, limit of detection; min, minute(s); *n*, number of measurements; NR, not reported; OEL, occupational exposure limit; ppb, parts per billion; ppm, parts per million; SD, standard deviation; SE, standard error; TWA, time-weighted average

[a] Exposure concentrations and range given in mg/m³, unless indicated otherwise; if published in another unit, the concentration in mg/m³ is given in square brackets

Table 1.3 Summary of selected studies on the biological monitoring of occupational exposure to benzene

Reference	Country, year	Occupational description	No. of participants	Benzene exposure in air (μg/m³)[a,b]	Urinary t,t-MA (μg/g creatinine)[a,c]	Urinary SPMA (μg/g creatinine)[a,c]	Urinary benzene (μg/L)[a,d]	Other biomarkers[a]
Campo et al. (2016)	Italy, NR	Filling station attendants	89	Median, 59 (5–284)[e]	Median, 127 (27–522)[e] μg/L	Median, 0.19 (< 0.1–1.28)[e] μg/L	Median, 0.339 (0.090–2.749)[e]	NR
		Unexposed workers	90	4 (1–18)[e]	Median, 117 (< 20–509)[e] μg/L	Median, < 0.1 (< 0.1–0.99)[e] μg/L	Median, 0.157 (0.054–2.554)[e]	NR
Lv et al. (2014)	China, NR	Shoe manufacturing workers	55	GM, 6980	NR	GM, 99[e]	NR	NR
Fustinoni et al. (2011)	Poland, NR	Petrochemical refinery workers	71	Median, 190 (50–2310)[e]	NR	Median, 0.65 (0.12–5.3)[e]	Median, 0.55 (0.117–7.487)[e]	NR
		Petrochemical office workers	97	NR	NR	Median, 0.40 (< 0.10–2.29)[e]	Median, 0.32 (0.083–2.316)[e]	NR
Carrieri et al. (2010)	Italy, 2006	Petrochemical workers	29	0.014 (< 0.001–0.280) ppm [45 (< 3–890)]	101 (< 6.86–746)	2.8 (< 0.06–38.59)	NR	NR
Lovreglio et al. (2010)	Italy, NR	Fuel tanker drivers	18	307 (7.4–1017)	134 (16–400)	2.94 (0.25–12.13)	2.96 (0.16–10.4)	Urinary phenol, 19 (5.0–33.0) mg/L
		Filling station attendants	23	23.5 (4.5–66.3)	86 (11–157)	0.79 (0.05–3.33)	0.62 (0.04–2.87)	Urinary phenol, 17.1 (8.0–29.0) mg/L
		Controls	31	4.6 (< 3.0–11.5)	93 (13–734)	0.65 (0.03–4.48)	1.23 (< 0.02–11.4)	Urinary phenol, 18.6 (3.0–36.0) mg/L
Hoet et al. (2009)	NR	Petrochemical workers	110	< 0.1 ppm [< 320]	50 (< 20–980)	0.97 (0.21–12.78)	0.270 (< 0.10–5.35)	Blood benzene, 0.405 (< 0.10–13.58) μg/L
Bråtveit et al. (2007)	Norway, 2004–2005	Petrochemical workers	12	0.042 (< 0.001–0.69) ppm [130 (< 3–2200)]	NR	NR	3.9 (0.5–34) nmol/L	Post shift; blood benzene, 1.8 (1.0–4.0) nmol/L
		Catering operator and office employees	9	NR	NR	NR	1.6 (0.5–4.0) nmol/L	Blood benzene, 1.8 (1.0–4.0) nmol/L

Table 1.3 (continued)

Reference	Country, year	Occupational description	No. of participants	Benzene exposure in air (µg/m³)[a,b]	Urinary t,t-MA (µg/g creatinine)[a,c]	Urinary SPMA (µg/g creatinine)[a,c]	Urinary benzene (µg/L)[a,d]	Other biomarkers[a]
Bahrami et al. (2007)	Islamic Republic of Iran, NR	Taxi drivers	60	0.31 (0.07–0.95) ppm [990 (220–3030)]	310 (90–1270)	NR	NR	NR
		Petrol station workers	9	1.40 (0.2–3.1) ppm [4470 (640–9900)]	2640 (1200–3280)	NR	NR	NR
		Controls	18	ND	170 (10–350)	NR	NR	NR
Manini et al. (2006)	Italy, 2004	Taxi drivers	21 NS	7.5	122	GM, 2.14	GM, 0.44	NR
			16 S	8.1	154	GM, 3.79	GM, 2.58	NR
Kim et al. (2006a)	China, 2000–2001	Shoemaking factory workers	164 women	Median, 1.28 (0.017–88.9) ppm [4090 (54–284 000)]	Median, 13.5 (0.644–426) µmol/L	Median, 262 (1.50–29 400) nmol/L	Median, 283 (6.21–53 900) nmol/L	NR
			86 men	Median, 1.05 (0.122–50.2) ppm [3350 (390–160 350)]	Median, 10.3 (1.50–370) µmol/L	Median, 137 (3.68–33 000) nmol/L	Median, 216 (19.4–42 600) nmol/L	NR
		Clothes manufacturing workers (controls)	87 women	Median, 3.40 (0.146–21.2) ppb [10.86 (0.47–67.72)]	Median, 1.06 (0.152–6.17) µmol/L	Median, 1.94 (0.591–86.4) nmol/L	Median, 1.48 (0.091–7.47) nmol/L	NR
			52 men	Median, 3.71 (0.146–533) ppb [11.85 (0.47–1702.55)]	Median, 1.09 (0.132–5.78) µmol/L	Median, 3.24 (0.591–68.1) nmol/L	Median, 1.59 (0.091–130) nmol/L	NR
Fustinoni et al. (2005)	Italy, 1999–2000	Filling station attendants	78	Median, 61 (11–478)	NS: Median, 49 (< 10–581) µg/L	Median, 5.8 (0.2–10.9) µg/L	Median, 0.342 (0.042–2.836)	NR
					S: Median,144 (15–321) µg/L	Median, 7.5 (0.2–24.8) µg/L	Median, 1.168 (0.055–5.111)	Section 1.01 NR
		Traffic police	77	Median, 22 (9–316)	NS: Median, 82 (< 10–416) µg/L	NS: Median, 5.3 (0.2–13.8) µg/L	NS: Median, 0.151 (0.025–0.943)	NR
					S: Median, 213 (52–909) µg/L	S: Median, 9.1 (2.4–13.8) µg/L	S: Median, 0.753 (0.054–4.246)	NR
		Office workers	58	Median, 6 (< 6–115)	NS: Median, 33 (< 10–1089) µg/L	NS: Median, 4.1 (0.2–12.5) µg/L	NS: Median, 0.133 (< 0.015–0.409)	NR
					S: Median, 71 (< 10–270) µg/L	S: Median, 8.0 (0.2–13.9) µg/L	S: Median, 0.331 (0.064–4.615)	NR

Table 1.3 (continued)

Reference	Country, year	Occupational description	No. of participants	Benzene exposure in air (μg/m³)[a, b]	Urinary t,t-MA (μg/g creatinine)[a,c]	Urinary SPMA (μg/g creatinine)[a,c]	Urinary benzene (μg/L)[a,d]	Other biomarkers[a]
Fustinoni et al. (2005) (cont.)		Bus drivers	152	Median, 21 (< 6–92)	NS: Median, 57 (< 10–536) μg/L	NS: Median, 5.6 (0.2–13.3) μg/L	NR	NR
					S: Median, 174 (< 10–695) μg/L	S: Median, 9.3 (0.2–65.9) μg/L	NR	NR
		Researchers	49	Median, 9 (< 6–46)	NS: Median, 51 (< 10–181) μg/L	NS: Median, 9.0 (0.2 – 182.2) μg/L	NR	NR
					S: Median, 195 (< 10–444) μg/L	S: Median, 13.7 (3.0–19.9) μg/L	NR	NR
Chakroun et al. (2002)	Tunisia, NR	Tanker fillers	20	0.16 (0.02–0.42) ppm [510 (63.89–1340)]	350 (80–1110)	NR	NR	NR
		Filling station attendants	10	0.20 (0.09–0.52) ppm [640 (290–1660)]	310 (150–590)	NR	NR	NR
		Controls	20	ND	110 (20–390)	NR	NR	NR
Waidyanatha et al. (2001, 2004)	China, ~1995	Rubber, adhesive, and paint manufacturers	42	14.5 (1.65–30.6) ppm [46 320 (5270–97 740)]	16 200 (1140–77 800) μg/L	712 (050–5890) μg/L	8.42 (0.837–27.9)	NR
				109 (31.5–329) ppm [348 180 (100 620–1 050 920)]	51 300 (7250–133 000) μg/L	9420 (123–27 500) μg/L	50.2 (1.30–284)	NR
		Sewing machine manufacturing workers (controls)	41	0.015 (0.0–0.11) ppm [48 (0.0–350)]	108 (020–338) μg/L	21 (2–79) μg/L	0.145 (0.027–2.06)	NR
Kivistö et al. (1997)	Estonia, 1994	Benzene production (in winter)	25	1.6 (0.06–14.7) ppm [5110 (190–46 960)]	38 (< 0.2–210) μmol/L	99 (< 0.3–1030)	965 (10–6250) nmol/L	Benzene in blood, 174 (8–1160) nmol/L
		Cokery workers (in winter)	27	1.3 (0.09–11.7) ppm [4150 (290–37 370)]	11 (< 0.2–35) μmol/L	73 (< 0.3–1020)	372 (22–1750) nmol/L	Benzene in blood, 160 (18–1690) nmol/L
		Rural controls (in winter)	10	0.009 ppm [28.75]	0.8 (< 0.2–8.1) μmol/L	2.1 (< 0.3–18)	12 (2–45) nmol/L	Benzene in blood, 7 (< 3–22) nmol/L

Table 1.3 (continued)

Reference	Country, year	Occupational description	No. of participants	Benzene exposure in air (µg/m³)[a, b]	Urinary t,t-MA (µg/g creatinine)[a,c]	Urinary SPMA (µg/g creatinine)[a,c]	Urinary benzene (µg/L)[a,d]	Other biomarkers[a]
Boogaard & van Sittert (1995, 1996)	Several countries including Belgium, Germany, and the Netherlands, 1992–1994	Natural gas production platforms	24	< 100–19 200	< 10–9920 µmol/mol creatinine	< 0.5–378 µmol/mol creatinine	NR	NR
		Chemical manufacturing	130	< 10–100 000	< 10–31 300 µmol/mol creatinine	< 1–1096 µmol/mol creatinine	NR	NR
		Oil refineries with aromatic plants	16	110–3300	8–1200 µmol/mol creatinine	0.9–46.4 µmol/mol creatinine	NR	NR
		Fuel tanker drivers and gasoline attendants	14	NR	9–830 µmol/mol creatinine	0.5–8. µmol/mol creatinine	NR	NR
		Employees without potential benzene exposure	38 NS	NR	29 µmol/mol creatinine	0. 94 µmol/mol creatinine	NR	NR
			14 S	NR	46 µmol/mol creatinine	1.71 µmol/mol creatinine	NR	NR

GM, geometric mean; ND, not detected; NR, not reported; NS, non-smokers; ppb, parts per billion; ppm, parts per million; S, smokers; SPMA, *S*-phenylmercapturic acid; t,t-MA, *trans*,*trans*-muconic acid

[a] Benzene exposure and biomarker concentrations are reported as mean (minimum–maximum), if not indicated otherwise

[b] Exposure concentration and range given in µg/m³ unless indicated otherwise; if published in another unit, the conversion to µg/m³ is given in square brackets

[c] Exposure concentration and range given in µg/g creatinine unless indicated otherwise; if published in another unit, the conversion to µg/g creatinine is given in square brackets

[d] Exposure concentration and range given in µg/L unless indicated otherwise; if published in another unit, the conversion to µg/L is given in square brackets

[e] 5–95th percentile

The separation and processing of crude oil and natural gas into crude oil, condensate, gas, and produced water before transport to shore via pipelines or tank ships takes place in a closed processing equipment and pipeline system. All four petroleum streams contain benzene, however, and the likelihood of exposure to benzene increases whenever the system is opened. The composition of crude oil and gas condensate varies between oil and gas fields and depends upon several factors, such as geological conditions in the reservoirs and the production age of the oil field, but typically lies within the range of < 0.01 and 3.0% by weight (Verma & des Tombe, 1999; Verma et al., 2000; Kirkeleit et al., 2006a), with benzene content in condensate being higher. The full-shift mean exposure in the production of oil and natural gas is usually well below 1 ppm [3.19 mg/m^3] benzene, the 8-hour permissible exposure limit set by the Occupational Safety and Health Administration (OSHA, 2017), during ordinary activity (Glass et al., 2000; Verma et al., 2000; Kirkeleit et al., 2006a; Bråtveit et al., 2007; Steinsvåg et al., 2007) (Table 1.2). However, some specific tasks, such as cleaning and maintenance of tanks and separators, pipeline pigging operations, and storage tank gauging, may cause short-term exposures in excess of this (Runion, 1988; CONCAWE, 2000; Glass et al., 2000; Verma et al., 2000; Kirkeleit et al., 2006a; Esswein et al., 2014).

With technological advances and more efficient reservoir completion techniques, UOGD has grown in the past decades. The only study available for this segment indicates that the potential for exposure is higher than for conventional oil and gas extraction (Esswein et al., 2014; Table 1.2).

(ii) Downstream petroleum industry: refining

The full-shift exposure to benzene during ordinary activity in the refining petroleum industry tends to be higher than for upstream activities, but still with average values well below 1 ppm [3.19 mg/m^3] (Nordlinder & Ramnäs, 1987; Verma et al., 1992, 2001; CONCAWE, 2000, 2002; Glass et al., 2000; Akerstrom et al., 2016; Almerud et al., 2017) (see Table 1.2). However, the range of exposure indicates potential for exceeding 1 ppm [3.19 mg/m^3]; this is particularly true for refinery maintenance, laboratory technicians, and dock workers. Specific tasks such as sampling, opening of vessels for maintenance and cleaning, and loading of petrol may cause high short-term exposure (Runion, 1988; Hakkola & Saarinen, 1996; Vainiotalo & Ruonakangas, 1999; Davenport et al., 2000; Verma et al., 2001; Kreider et al., 2010; Widner et al., 2011). However, while workers before 2000 were likely to have been exposed to higher concentrations of benzene because of a higher content of benzene in reformate stream (Burns et al., 2017), the range of benzene exposures reported in recent studies is considerably reduced (Campagna et al., 2012; Akerstrom et al., 2016; Almerud et al., 2017; Burns et al., 2017). Some of the reported exposure levels are given in Table 1.2.

(iii) Downstream petroleum industry: distribution

In the petroleum transport chain there is a potential for exposure at each point where the products are stored and transferred, and the reported exposures tend to be higher than for production and refinery workers (Halder et al., 1986; Javelaud et al., 1998; CONCAWE, 2000, 2002; Glass et al., 2000). However, because of a lowered content of benzene in petrol (Verma & des Tombe, 2002; Williams & Mani, 2015), as well as the recent introduction of vapour recovery systems in the petroleum distribution chain in at least developed countries, the exposure to benzene for these groups of workers has declined over the years. Some of the reported exposure levels are given in Table 1.2.

Williams et al. (2005) reviewed the available industrial hygiene data describing exposure during the marine transport of products