UNITED NATIONS

UNEP

MC

UNEP/MC/COP.4/INF/12*

MINAMATA CONVENTION ON MERCURY

Distr.: General
23 September 2021

English only

Conference of the Parties to the Minamata Convention on Mercury
Fourth meeting
Online, 1–5 November 2021 and Bali, Indonesia, 21–25 March 2022
Agenda item 4 (i)

Matters for consideration or action by the Conference of the Parties: effectiveness evaluation

Giving effect to article 22: effectiveness evaluation

Guidance on monitoring of mercury and mercury compounds to support evaluation of the effectiveness of the Minamata Convention

Note by the secretariat

1. In paragraph 2 (a) of decision MC-3/10, on the arrangements for the first effectiveness evaluation of the Convention, the Conference of the Parties to the Minamata Convention on Mercury requested the secretariat to advance the work on the effectiveness evaluation by securing services for drafting guidance on monitoring to maintain harmonized, comparable information on mercury levels in the environment.

2. In response to decision MC-3/10, the secretariat, in consultation with the Bureau of the fourth meeting of the Conference of the Parties, prepared a road map[1] outlining an iterative and participatory process for the development of guidance on monitoring in the context of the effectiveness evaluation. In line with the road map, the secretariat developed a draft annotated outline of the monitoring guidance and held open online information sessions in June 2020 to discuss the development of the guidance. Subsequently, parties and organizations were invited to identify experts to contribute to the drafting of the guidance,[2] and three consultants were engaged by the secretariat to draft chapters on mercury monitoring in air, biota and humans.

3. The first online meeting of the experts and consultants was held on 15 September 2020, and the final annotated outline of the guidance was developed taking into account the comments received. Further thematic online meetings were held from September 2020 to March 2021 to develop the guidance. Subsequently, the secretariat, working with the consultants and supported by the experts identified by parties and organizations, developed a first draft of the guidance, which was made available for comments by parties and relevant stakeholders on 15 April 2021. A total of 14 submissions were received from 8 countries and 6 organizations. After further consultation with the experts, a second draft of the guidance, along with supplementary material, was developed and made

* Reissued for technical reasons on 11 March 2022.

[1] The documents and submissions mentioned in the present note are available online at https://www.mercuryconvention.org/en/meetings/cop4#sec971.

[2] At the time of drafting of the present note, 37 experts had been identified by 16 parties and 42 by organizations to contribute to the development of the guidance.

available for review by parties and organizations on 15 July 2021. A total of 15 submissions were received during the commenting period, of which 10 were from parties and 5 from organizations.

4. Throughout the development of the guidance, an attempt was made to address all comments and suggested amendments in an inclusive manner consistent with the annotated outline. Several bilateral discussions took place, between countries or organizations and the secretariat or consultants, in an attempt to fully address the comments and suggested amendments. Parties and organizations were also invited to submit additional information on existing monitoring programmes and available standard operating procedures. Despite the efforts made, some comments and suggestions could not be reflected in the guidance, in particular those requesting the removal of text that had been part of the annotated outline and had received support from other reviewers. To support transparency and maintain open communication, parties and organizations were invited to contact the secretariat to discuss questions and comments related to the development of the guidance, including with regard to instances in which their comments had not been fully reflected in the revisions.

5. The resulting text, entitled “Guidance on monitoring of mercury and mercury compounds to support evaluation of the effectiveness of the Minamata Convention”, is contained in the annex to the present note and is presented without formal editing. The guidance consists of six chapters: (1) Introduction and objectives; (2) Comparable monitoring data and the effectiveness evaluation; (3) Atmospheric mercury monitoring; (4) Biota mercury monitoring; (5) Human biomonitoring; and, (6) Cross-media data management and analysis. It also has an executive summary, a list of references to the publications cited, and an annex containing an overview of a tiered approach to monitoring mercury in the environment and in humans.

6. A supplement to the main guidance document, entitled “Supplementary material – guidance on monitoring of mercury and mercury compounds to support evaluation of the effectiveness of the Minamata Convention” (UNEP/MC/COP.4/INF/25), has two parts: part A, containing an overview of existing monitoring programmes organized by matrix (air, biota and human biomonitoring), an overview of existing gaps, and a non-exhaustive list of standard operating procedures; and part B, which contains an overview of quality assurance and quality control procedures in laboratory analysis and data management and a draft template for the submission of monitoring data.

Annex

GUIDANCE ON MONITORING OF MERCURY AND MERCURY COMPOUNDS TO SUPPORT EVALUATION OF THE EFFECTIVENESS OF THE MINAMATA CONVENTION*

Draft of 23 September 2021

Table of Contents

* The annex has not been formally edited.

Acknowledgements

The Secretariat of the Minamata Convention wishes to acknowledge and thank the countries, institutions and experts who made the development of this guidance possible. First among those are over 80 experts – scientists from chemical, biological and human health sciences, policymakers, Indigenous representatives, and health practitioners, amongst others – who have given freely their time and expertise to shape and develop this document, acting as peer-reviewers of the successive drafts of this document.

A special thanks go to the four Lead Authors of the different chapters – Lynwill Martin (chapter 3), David Evers (chapter 4), Niladri Basu (chapter 5) and Colin Thackray (chapter 6) – who have each gone well beyond the call of duty and donated a substantial contribution of time, effort and intellectual capital to bring each chapter to fruition, over the last two years and many drafts and short deadlines.

We express our gratitude for the support provided by other UN organizations, including the World Health Organization (WHO), the International Labour Organization (ILO), the International Atomic Energy Agency (IAEA), UN Environment Programme (UNEP) and the Secretariat of the Basel, Rotterdam and Stockholm Conventions, as well as to various national Governments that have participated in the different rounds of review and helped improve the guidance with a multidisciplinary focus.

Last, but not least, appreciation goes to all members of the Minamata Convention family who were actively engaged throughout the process to conceptualize and develop this guidance.

List of Acronyms

AAS	Atomic absorption spectrometry
AFS	Atomic fluorescence spectrometry
AIC	Akaike information criterion
AMAP	Arctic Monitoring and Assessment Programme
ASGM	Artisanal and small-scale gold mining
CARE	Collective Benefits, Authority to Control, Responsibility, and Ethics
CART	Classification and regression tree
CHMS	Canadian Health Measures Survey
CI	Confidence interval
CIOMS	Council for International Organizations of Medical Sciences
COP	Conference of the Parties
CV	Cold vapour
DHS	Demographic and Health Surveys
DMA	Direct mercury analyzer
DOC	Dissolved Organic Carbon
dw	Dry weight
EHMS	Environmental Health Monitoring System
FAIR	Findability, Accessibility, Interoperability, and Reuse
FAO	Food and Agriculture Organization
fw	Fresh weight
fww	Fresh wet weight
GAM	Generalized additive model
GEF	Global Environment Facility
GEM	Gaseous elemental mercury
GerES	German Environmental Survey
GLM	Generalized linear model
GLMM	Generalized linear mixed model
GOM	Gaseous oxidised mercury
HBM	Human biomonitoring
Hg	Mercury
IAEA	International Atomic Energy Agency
ILO	International Labour Organization
INSPQ	Institut national de santé publique du Québec
IUCN	International Union for Conservation of Nature
KoNEHS	Korean National Environmental Health Survey
LSM	Large-scale mining
MeHg	Methylmercury
MK	Mann-Kendall
MRPP	Multiple-response permutation procedure
NCP	Northern Contaminants Program
NHANES	National Health and Nutrition Examination Survey
NIES	National Institute for Environmental Studies, Japan
NIST	National Institute of Standards and Technology
NTP	National Toxicology Program
OCAP	Ownership, Control, Access, and Possession

PAS	Passive air sampler
PM	Particulate matter
PBM	Particle-bound mercury
PCA	Principal component analysis
PMF	Probability mass function
PSCF	Potential source contribution function
QA/QC	Quality assurance and quality control
QMS	Quality management systems
RGM	Reactive gaseous mercury
SD	Standard deviation
SOC	Soil organic carbon
SOP	Standard operating procedure
STROBE	Strengthening the reporting of observational studies
TGM	Total gaseous mercury
THg	Total mercury
TSS	Total suspended solids
UNDP	United Nations Development Programme
UNEP	United Nations Environment Programme
UNIDO	United Nations Industrial Development Organization
US CDC	United States Centers for Disease Control and Prevention
USEPA	United States Environmental Protection Agency
USAID	United States Agency for International Development
WHO	World Health Organization
WMO	World Meteorological Organization
ww	Wet weight

Executive Summary

In paragraph 2 of article 22, on effectiveness evaluation, the Minamata Convention on Mercury requires the Conference of the Parties to make "arrangements for providing itself with comparable monitoring data on the presence and movement of mercury and mercury compounds in the environment as well as trends in levels of mercury and mercury compounds observed in biotic media and vulnerable populations".

The "Guidance on monitoring of mercury and mercury compounds to support evaluation of the effectiveness of the Minamata Convention" (hereinafter the "monitoring guidance") provides scientific and technical guidance to support the Conference of the Parties in obtaining comparable monitoring data for the effectiveness evaluation. The overall aim of the monitoring guidance is to (i) explain the role of monitoring in the effectiveness evaluation and set realistic expectations about what can be learned over time; (ii) provide guidance to parties and organizations that are currently conducting monitoring programmes on what data and accompanying information would inform the effectiveness evaluation; and (iii) provide guidance to parties and organizations who wish to develop new monitoring programmes or improve existing ones, with a view to contributing to the effectiveness evaluation.

The following four overarching policy questions have been proposed to help frame the Effectiveness Evaluation:[3]

(a) Have the parties taken actions to implement the Minamata Convention?

(b) Have the actions taken resulted in changes in mercury supply, use, emissions and releases into the environment?

(c) Have those changes resulted in changes in levels of mercury in the environment, biotic media and vulnerable populations that can be attributed to the Minamata Convention?

(d) To what extent are existing measures under the Minamata Convention meeting the objective of protecting human health and the environment from mercury?

Monitoring levels of mercury in air, biota, and humans can contribute to addressing the third and fourth policy questions above.

The monitoring guidance describes the scientific and technical processes and guiding principles for compiling and/or generating comparable monitoring data. It also suggests methods that can be used for understanding the presence, movements and trends of mercury in the environment and humans based on monitoring data, in order to inform the effectiveness evaluation. Throughout the guidance, monitoring activities have been grouped to achieve six objectives:

Objective 1: Estimation of mercury concentrations for areas without (i.e., background sites) or with (i.e., affected sites) local anthropogenic sources.

Objective 2: Identification of temporal trends.

Objective 3: Characterization of spatial patterns.

Objective 4: Estimation of source attribution.

Objective 5: Estimation of exposure and adverse impacts.

[3] Document UNEP/MC/COP.3/14.

Objective 6: Quantification of key environmental processes to improve understanding of cause-effect relationships.

For each of these six monitoring objectives, questions have been formulated to guide the collection and analysis of the relevant monitoring data and to inform the effectiveness evaluation in complementary ways. These guiding questions are set out in chapter 2 of the monitoring guidance. Answers to the guiding questions provide several lines of evidence with different strengths and challenges. Together, they form a range of scientific weight of evidence that can give evidence-based support to the effectiveness evaluation.

To strengthen the scientific evidence for the effectiveness evaluation, comparable and high-quality monitoring data should be used. The quality assurance and quality control (QA/QC) protocols employed by existing monitoring programmes will provide a basis to inform the development of comparable data of high quality. Data generated from different monitoring programmes may be supplemented, as appropriate, with comparable and high-quality data from academia and research. This may be accomplished through a well-documented and transparent set of "data flags" that will enable the use of data from different sources with different levels of QA/QC.

Air, biota and humans were identified as key matrices for monitoring trends in the movement of mercury from its sources to the environment and into human populations. A tiered approach to monitor trends in these different matrices is presented in the monitoring guidance, with a view to supporting parties and organizations who wish to develop new monitoring programmes or improve existing ones.

The tiered approach for the three matrices can differ in terms of which monitoring objectives are primarily being targeted; however, for the most part, tier 1 aims to provide evidence to support the achievement of objectives 1, 2 and partially 5; tier 2 aims to provide information that supports objectives 3, 4, and 5; and tier 3 aims to support objective 6, which in turn will improve the scientific strength of the data for the achievement of the other five objectives. Each tier builds upon the former tier to provide a better overall weight of evidence. Overall, the tiered approach is as follows:

Tier 1 is intended to provide guidance on baseline mercury monitoring under a limited set of parameters for circumstances where available resources are limited. The methods in tier 1 are cost-effective, practical, feasible and sustainable.[4] The tier 1 methods are intended to provide information that is useful in identifying and characterizing gaps and needs of national, regional or local interest and to provide information that is useful to the collective effort for the effectiveness evaluation. While the implementation of tier 1 actions may not fully address the monitoring objectives, it will contribute valuable information and create a foundation for tier 2 monitoring.

Tier 2 is intended to build upon tier 1 methods and create a basis for assessing source attribution at the local, national and global scales. The methods and approaches in this tier may be more expensive or complex than those under tier 1. Although implementation of tier 2 is not required by all parties, the more tier 2 approaches that are implemented, the better the weight of evidence for the effectiveness evaluation will be.

[4] In decision MC-1/9, the Conference of the Parties noted that the monitoring arrangements should take into consideration cost-effectiveness, practicality, feasibility and sustainability.

Tier 3 identifies research methods and approaches that may play a vital role in supporting the tier 1 and tier 2 programmes and the effectiveness evaluation, primarily by improving understanding of key processes that link sources to environmental concentrations and exposures (objective 6). Tier 3 focuses on processes; thus, the results would likely yield insights that are broadly applicable and strengthen the weight of scientific evidence used to support the other monitoring objectives. Tier 3 information should therefore be taken into consideration in the effectiveness evaluation where available.

The tiered recommendations are further elaborated for each of the key matrices in chapters 3 (air), 4 (biota) and 5 (humans). While the overall tier principles are similar in each of the matrices, there are some differences in the recommended approaches. For example, in the approaches to monitor mercury in air, the primary differences between the tiers are the methods employed to collect data. In the biota chapter, the main differences between the tiers reflect how sites are selected and sampled, as well as what ancillary measurements are collected. In the human biomonitoring chapter, the three tiers are primarily differentiated by the target human population and how they are sampled. The annex to the monitoring guidance presents a tabular summary of the recommended data to be collected under each tier, for each of the three matrices. Chapter 6 discusses how single- and cross-matrix analyses of the observations can be performed using various mechanistic and statistical models to support the monitoring objectives and inform the effectiveness evaluation.

Atmospheric mercury monitoring

Mercury levels in the atmosphere are linked to mercury emissions from natural, geogenic and anthropogenic sources. Key anthropogenic sources of atmospheric mercury influenced by the Convention include the point sources listed in annex D to the Convention and the intentional use of mercury in artisanal and small-scale gold mining (ASGM) and other industrial products and processes. In the context of the effectiveness evaluation, it will be relevant to estimate how significant the contribution of sources influenced by the Convention are compared to total anthropogenic emissions, as well as legacy and natural emissions, and how these emissions travel and impact the receiving environment. Many of the Convention measures to control mercury supply, use, emissions, releases, storage and disposal are expected to reduce levels of mercury in the atmosphere.

Chapter 3 identifies different methods parties and organizations can use to monitor atmospheric mercury and generate comparable data to support the effectiveness evaluation. Atmospheric mercury has been successfully monitored for decades but not all regions have been covered equally, with the biggest data gaps occurring in the southern hemisphere. The suggested tiers for air monitoring gives parties and organizations an opportunity to start, expand or improve their monitoring programmes in a manner such that comparable data can be generated to support the effectiveness evaluation.

Air monitoring is well established in many areas. The guidance offers the opportunity of joining or employing one of the several existing monitoring programmes or networks to draw from the experience and information that these established networks can provide. Automated atmospheric mercury data collection is the predominant method within existing monitoring networks; however, passive and manual sampling of atmospheric mercury are two other options also presented for consideration. The advantages and disadvantages of employing each method are presented in chapter 3.

Depending on the specific needs of the monitoring initiative, the monitoring guidance puts forward different methods at tier 1 as the minimum step to start generating comparable atmospheric mercury data of high quality. The objective of tier 1 air monitoring is to provide comparable data to identify temporal trends and characterize spatial patterns to gain an understanding of the changes in the distribution of mercury over time around the world. Wet deposition of mercury from the atmosphere, which is one of the methods included at the tier 1 level, is a well understood method that provides comparable results helpful for understanding part of the atmospheric deposition of mercury to a receiving environment. Therefore, the tier 1 recommendations offer scientifically sound and cost-effective means of acquiring comparable and high-quality data on mercury concentrations in air.

It is important for each monitoring initiative to identify sites that can provide insights into the guiding questions. Thus, recommendations are provided on where to monitor mercury in the air in order to best observe changes from emissions, inform atmospheric model capabilities and fill data gaps. A variety of site locations should be considered, including background/remote, rural, urban and contaminated/industrial sites. Each site type addresses a different monitoring requirement and should be carefully chosen to focus on the appropriate question. To the extent possible, the air monitoring sites should be coordinated with sites (or vulnerable populations) in which mercury is monitored in biota or humans.

A wealth of experience on key elements and processes related to good QA/QC of the data is available from existing atmospheric mercury monitoring programmes and networks. Details on how best to implement good QA/QC programmes are identified both in the main guidance document and the supplementary material.

Overall, the elements put forward in chapter 3 will help answer the different monitoring guiding questions for the effectiveness evaluation with regard to atmospheric mercury monitoring. Furthermore, chapter 3 provides parties and relevant organizations with the means of starting, improving or expanding on their initiatives for monitoring atmospheric mercury to enable them to deliver comparable data for the effectiveness evaluation.

Biota mercury monitoring

The approach to monitoring mercury in biota in support of the effectiveness evaluation takes into account: (a) the monitoring objectives described above and the guiding questions identified in chapter 2; (b) the current scientific understanding of mercury's biogeochemical cycle, including its transport, transformation and bioaccumulation, as well as atmospheric deposition, local pressures and large -scale drivers that affect these processes; and (c) the tiered approach presented to expand and develop monitoring programmes with available resources.

Mercury transport, transformation and bioaccumulation in the marine and continental environment is known to be influenced by a number of competing processes that ultimately determine how much mercury is found in a given biotic sample. The biomagnification and bioaccumulation of mercury in the food chain will depend on both the bioavailability of methylmercury and the food-web dynamics. While many of these processes are known, their relative strength and complexity is site- and location-dependent. This complexity makes site classification according to land use, habitat and ecosystem characteristics critical in data collection. When assessing biotic results, external pressures such as atmospheric deposition, industrial/agricultural or ASGM activity and large-scale drivers (for example, climate change) that can influence the system should be taken into account. Choice of bioindicators and related types of tissue are also critical decision points, as biotic methylmercury concentrations can vary

significantly by trophic level and are often impacted by life history and ecological factors. Thus, the recommended tiers in the biota chapter reflect these and other considerations in its design. The necessary elements of monitoring mercury in biota have been arranged into tiers to include the selection of monitoring sites, bioindicators, tissue type and ancillary measurements.

For tier 1, it is recommended that the chosen sites compromise a mixture of (a) remote sites, with little local anthropogenic input that will be representative of background conditions and (b) sites with well-known anthropogenic impacts. As several routinely used methods for analysing mercury concentrations in biota exist, it is important to use the same method consistently over time in the chosen sites and to sample the sites annually to inform robust trend analysis. Sites that are governed by well-known biogeochemical processes and co-located with monitoring efforts in air or human biomonitoring should be prioritized. All these sites should be classified according to their land-use, habitat and ecosystem characteristics. Total Hg in muscle, blood, egg and keratin tissue of the monitored fish or birds on trophic level 3 and 4 are recommended because this trophic level is the most commonly measured and used as food by humans; choosing bioindicators at this trophic level is particularly suitable for ensuring consistency with ongoing monitoring efforts and for estimating exposure and adverse effects in humans. Ancillary measurements should be taken based on known (or suspected) co-variates of interest to normalize mercury concentrations for trend analysis. Where little or no prior information exists, experience with the use of geographic information system (GIS) maps gained during the Minamata Initial Assessments might also be helpful in choosing sites.

The tier-2-level recommendations include the addition of more locations that represent different site characteristics than those chosen at the tier 1 level and/or that are particularly suited to understanding the impact of a specific input, pressure or driver. Where beneficial, measurements can be collected at the additional locations on a rotational basis, resulting in every site being monitored every few years. It is recommended that, during the rotation, the same species be sampled in all sites. If that is not possible, sampling all the species used in the programme at least at some sites is recommended, to establish statistical relationships between the expected mercury levels. The tier 2 recommendations are aimed at the collection of additional ancillary measurements known to impact the inputs, pressures and drivers of interest at all sites. For example, carbon (^{13}C) and nitrogen (^{15}N) stable isotope measurements help assess changes in organic matter sources and the food-web. Further, water chemistry parameters such as dissolved organic matter and carbon, suspended solids, pH, dissolved oxygen and salinity can, in turn, give an indication of the impact from local pressures and large-scale drivers. Mercury measurements in the underlying sediments can be useful for tracking changes in a local inputs or pressures. These ancillary measurements, together with the site classification system introduced in tier 1, will also help to establish how widely the biogeochemical processes governing a particular site can be generalized with models. More details on the recommended ancillary measurements can be found in annexed tiered approach table.

Tier 3 recommendations build on tiers 1 and 2. Site selection and bioindicator sampling are the same, but other biota are suggested to be added to the data collection. Tier 3 recommendations also include the introduction of “supersites”, where a specific catchment or area of specific interest is monitored intensively, and “satellite sites” (sites with supporting data) in the vicinity of the supersite, by which the representativeness of the observed biogeochemical relationships can be established. Additional ancillary measurements, particularly of stable mercury isotopes, are also recommended to establish cause-effect relationships between

mercury levels in biota and the inputs, pressures and drivers that influence them. All the elements in this tier will therefore help to quantify the key environmental processes that govern mercury levels in biota and strengthen the weight of evidence that biota monitoring adds to the effectiveness evaluation.

Human biomonitoring

Human health may be negatively impacted by mercury exposure. Human populations may be exposed to elemental and inorganic mercury in occupational settings (for example, in ASGM and dentistry), from contact with certain products (for example, dental amalgams, some skin-lightening creams, broken fluorescent bulbs and other waste products) and from environmental contamination and dietary sources, including but not limited to shellfish, fish and marine mammals contaminated with methylmercury. Measuring mercury levels in the blood, hair and/or urine of individuals from target populations provides direct information on human exposures to mercury, from which risks to human health can be assessed.

Article 22 of the Convention requires the Conference of the Parties to establish arrangements to provide monitoring data on the trend in mercury levels in vulnerable human populations. This human biomonitoring data will help address the six monitoring objectives and support the effectiveness evaluation. Chapter 5 provides essential guidance and links to key resources, for parties and relevant organizations to consider in terms of using existing, and generating new, human biomonitoring data for the effectiveness evaluation.

There are several databases of human biomonitoring information and resources that can be used to help understand human exposures to mercury before the Minamata Convention's entry into force. This information helps to establish a baseline for the effectiveness evaluation. In terms of data to be collected in the future, there are two sources to consider. First, there is the biomonitoring data generated by existing government-led national biomonitoring programmes, regional initiatives and/or academic-led studies. Second, parties and relevant organizations can further support the effectiveness evaluation by implementing new biomonitoring studies in a harmonized way so that they are purposefully designed to fill data gaps and build capacity.

Human biomonitoring data can be designed as part of a tiered approach to inform new monitoring programmes or improve existing ones. The recommended activities in tier 1 are geared towards initiatives seeking to create a human biomonitoring programme or expand a minimal programme with modest resources. The goal of tier 1 is to focus on a vulnerable subpopulation and take total mercury measurements in blood, urine or hair. This activity should ideally be repeated in the same population every 2 to 5 years. A good starting point for tier 1 is the recent guidance from the World Health Organization for characterizing prenatal mercury exposure.[5] The tier 2 recommendations are aimed at the collection of data to inform all monitoring objectives and call for more in-depth analysis of the tier 1 subpopulation groups or incorporation of mercury biomonitoring into other, in-depth health surveys or cohort studies. Tier 3 aims to increase understanding of key processes that link mercury sources to human exposures, and thus resource-intensive research methods and approaches are required. These include national human biomonitoring programmes and surveys for comparison to vulnerable subpopulations, and coordination of human biomonitoring activities with air and biota monitoring where relevant.

[5] https://apps.who.int/iris/handle/10665/334181.

Key elements of all human biomonitoring studies that need to be considered include: (a) defining the target and sample population (which usually focus on groups vulnerable to mercury, i.e., those in early life stages or those with relatively high exposures); (b) selecting and measuring the appropriate biomarkers to help define exposure to different sources and forms of mercury (with total mercury measurements in hair, urine, blood and cord blood being most commonly used and accepted); (c) administering surveys to gather supportive information (e.g., on socio-demographics, occupational practices, dietary habits) to deepen understanding and assist in interpretation; and (d) managing and analysing data as per the guiding policy question. All these elements must be performed in a responsible and ethical manner.

Cross-media data management and analysis

From primary release to human exposure, mercury can undergo many physical and (bio-)chemical changes that interact with each other over a large range of timescales and can be influenced by human behaviour. Attribution of observed trends to specific drivers such as direct anthropogenic mercury releases, legacy mercury, process-driven releases of natural or anthropogenic influence and non-mercury environmental or behavioural drivers requires the use of models that resolve the intervening processes, supplemented or calibrated by empirical statistical approaches. Separating the relative magnitude of the inputs, pressures and drivers influenced by the Convention from those that are not will be key to assessing the effectiveness of already implemented policies. This makes cross-media analysis involving both mechanistic and statistical modelling in all relevant media a vital part of the scientific weight of evidence used to evaluate the effectiveness of the Convention.

By analysing monitoring data, temporal and spatial trends in the levels of mercury in specific environmental media or human matrices can be derived. These trends provide a first-level indication of whether the Convention may be contributing to protecting human health and the environment from the adverse effects of mercury by assessing whether levels in the environment and humans are changing. Analyses of the monitoring data collected in each matrix separately will be informative, but this monitoring data can also be used in an integrated manner, where multiple complementary analysis approaches are combined to answer the same question. This will improve robustness and increase the scientific weight of evidence. As more comparable and high-quality monitoring data becomes available and our understanding of the intervening processes improves, more detailed questions can be answered with a higher level of confidence.

To estimate levels of mercury in locations with or without known anthropogenic mercury sources, simple analyses can be conducted on monitoring data at sites chosen for this purpose. These observations, together with suitable models, can be used to conduct trend analysis that gives a transparent presentation of the confidence with which a trend has been detected, as well as its magnitude.

To characterize spatial patterns, several atmospheric chemical transport models can be used, supplemented with statistical models where beneficial to quantify the representativeness of the observed levels and trends in air and to extrapolate ambient air concentrations and wet deposition to areas with sparse monitoring data. Spatially resolved models in air and other media can be used to interpolate levels and trends of mercury while accounting for the drivers of spatial and temporal differences.

Two types of analyses can be employed when using models to estimate source attribution and exposure for the effectiveness evaluation: a “bottom-up” or process-based analysis that estimates effects of drivers on observable quantities, and a “top-

down" or observation-based analysis that identifies drivers. Bottom-up analyses can be used whenever suitable input parameters and a sufficient process-level understanding of the relevant system exists. Top-down analyses can be used whenever sufficient ancillary data and/or measurements are available (or suitable surveys, in case of human biomonitoring). These two approaches can be used separately, but the strongest weight of evidence is obtained when they are used together in a complementary manner. At intensively monitored sites, combined top-down and bottom-up analyses can be performed for air, biota and human biomarkers.

Finally, the quantification of key environmental processes can improve our understanding of cause-effect relationships, which in turn will improve the confidence with which models can be used to answer the guiding questions. An increased understanding of mercury processes can be obtained through the comparable and high-quality monitoring data compiled for the effectiveness evaluation, as well as through other experimental, monitoring, computational and modelling studies made available for the evaluation. The strength of the scientific weight of evidence available for the effectiveness evaluation will therefore improve in an iterative manner from one evaluation cycle to the next.

To improve transparency, understanding and legitimacy of the models used for the effectiveness evaluation, models can be evaluated and inter-compared to give a clear understanding of the confidence of their outputs with respect to the question(s) being asked. Key assumptions, parameters and functions, and the consequences of these choices, can be presented to all stakeholders. Participatory processes can also be used for model selection and/or construction to improve ownership of the results among policymakers.

In addition to the main document, the monitoring guidance offers supplementary material organized in two parts: part A, which contains an overview of existing monitoring programmes organized by matrix (air, biota and human biomonitoring), an overview of existing gaps and a non-exhaustive list of standard operating procedures; and part B, which contains an overview of quality assurance and quality control procedures in laboratory analysis and data management and a draft template for the submission of monitoring data.

Chapter 1. Introduction and Objectives

1.1. Introduction

The objective of the Minamata Convention on Mercury (herein referred to as the Convention) is to protect the human health and the environment from anthropogenic emissions and releases of mercury and mercury compounds (Article 1). The Convention contains, in support of this objective, provisions that relate to the entire life cycle of mercury, including controls on the supply and trade of mercury, products and processes where mercury is used, emissions and releases of mercury, and management of waste and contaminated sites (Articles 3-12). The Convention also includes provisions that support the Parties to fulfil their obligations (Articles 13 and 14), health aspects (Article 16), and measures to enhance knowledge and information (Articles 17-19). Article 22 of the Convention requires the Conference of the Parties (COP) to periodically evaluate the effectiveness of the Convention, and to perform this evaluation on the basis of available scientific, environmental, technical, financial and economic information. Comparable monitoring data on the presence and movement of mercury and mercury compounds in the environment, as well as trends in levels of mercury and mercury compounds observed in biotic media and vulnerable human populations, are of particular interest to the COP in the context of the Effectiveness Evaluation.

1.2. Objectives

This document, as requested by the COP in its decision MC-3/10 in November 2019, provides scientific and technical guidance to support the COP to obtain comparable monitoring data for the Effectiveness Evaluation. The primary objectives of this document are to:

(a) Explain the role of monitoring in the Effectiveness Evaluation and set realistic expectations about what can be learned over time.

(b) Provide guidance to Parties and organizations, which are currently conducting monitoring programmes, on what data and accompanying information would inform the Effectiveness Evaluation.

(c) Provide guidance to Parties and organizations who wish to develop new monitoring programs, or improve existing ones, with a view to contributing to the Effectiveness Evaluation.

This document describes the scientific and technical processes and guiding principles for compiling and/or generating comparable monitoring data, as well as methods to use such monitoring data for understanding the presence, movements and trends of mercury in the environment and humans, in the context of evaluating effectiveness of the Convention.

Chapter 2 builds on the four overarching policy questions proposed for the Effectiveness Evaluation and establishes five categories of monitoring activities that can produce comparable data to address these questions. It explains the rationale for selecting air, biota and human as core matrices for monitoring activities, and presents general guidance that is relevant to all matrices to support efforts towards obtaining comparable monitoring data.

Following chapters address monitoring of mercury in specific matrices: air (chapter 3), biota (chapter 4) and humans (chapter 5). These chapters describe the significance of monitoring the matrices, and provide guidance on the selection of monitoring sites, sampling and measurement methods, quality control and assurance, and data collection, management, analysis and evaluation.

Chapter 6 discusses how these matrix-specific data can be compiled, analysed and synthesized, how those data can be used in mechanistic and statistical models, and how observed changes in mercury levels in environmental media and humans observed can be interpreted.

The Annex contains a proposed tiered approach for programmes to monitor mercury and mercury compounds to support the Effectiveness Evaluation.

The Supplementary material to the guidance presents an overview of existing monitoring activities undertaken by Parties and other stakeholders, as well as a review of gaps in the monitoring of key matrices. The Supplementary Material will be a "living document" that may be updated to support the COP in identifying available monitoring information for the Effectiveness Evaluation, as well as to support Parties and relevant organizations to consider whether their monitoring activities could contribute to filling the gaps. Other supplemental information will be developed to support the use of this document, including the comparison of existing standard operating procedures, international QA/QC programmes, and available reference materials.

Chapter 2. Comparable Monitoring Data and the Effectiveness Evaluation

2.1. Introduction

Paragraph 2 of Article 22 on Effectiveness Evaluation of the Minamata Convention requires the Conference of the Parties (COP) to make "arrangements for providing itself with comparable monitoring data on the presence and movement of mercury and mercury compounds in the environment as well as trends in levels of mercury and mercury compounds observed in biotic media and vulnerable populations". It has been proposed that the Effectiveness Evaluation of the Convention should address the following four overarching policy questions:[6]

(a) Have the Parties taken actions to implement the Minamata Convention?

(b) Have the actions taken resulted in changes in mercury supply, use, emissions and releases into the environment?

(c) Have those changes resulted in changes in levels of mercury in the environment, biotic media and vulnerable populations that can be attributed to the Minamata Convention?

(d) To what extent are existing measures under the Minamata Convention meeting the objective of protecting human health and the environment from mercury?

Monitoring levels of mercury in air, biota, and humans can contribute to addressing the third and fourth policy questions above. Detecting changes in mercury levels, estimating the human or ecosystem health impacts of those changes, and attributing them to actions influenced by the Minamata Convention require the use of mechanistic and/or statistical models. Therefore, observations are needed not only to detect and quantify changes, but also to improve and evaluate models of mercury transport, fate, exposure, and impacts. Monitoring activities have been grouped to achieve six objectives:

Objectve 1: Estimation of contemporary mercury concentrations for areas without (i.e., background sites) or with (i.e., affected sites) local anthropogenic sources.

Objectve 2: Identification of temporal trends.

Objectve 3: Characterization of spatial patterns.

Objectve 4: Estimation of source attribution.

Objectve 5: Estimation of exposure and adverse impacts.

Objectve 6: Quantification of key environmental processes to improve our understanding of cause-effect relationships.

2.2. Weight of evidence and guiding questions

From each of the monitoring objectives above, guiding questions can be drawn, as outlined in Table 2.1, to guide the collection and analysis of the relevant monitoring data and inform the Effectiveness Evaluation in complementary ways. Different types of observations and sources of data may be most appropriate for addressing different questions during the Effectiveness Evaluation. National and multi-country monitoring programmes, including those identified in the Supplementary Material, as well as programmes and projects overseen by international organizations, such as WHO and the GEF, may be prioritized as the preferred sources of comparable monitoring data. In the absence of those, additional sources of comparable data may also provide valuable information to

[6] Document UNEP/MC/COP.3/14 and further information therein.

support the Effectiveness Evaluation. Quality control measures, including those listed in chapters 3-5 and Supplementary Material, will be needed to assess the usefulness and validity of different data sets and maximize scientific weight of evidence.

The more basic scientific questions on levels and trends of mercury in humans and the environment are often more easily answered with a high level of confidence than the more complex questions related to source attribution and exposure assessment. However, answers to the more complex questions are more accurate indicators of the effect of specific measures under the Convention. How confidently and accurately a particular question can be answered will, in turn, depend on the quality and representativeness of the available data, and the robustness of the scientific analysis. Together, the monitoring objectives and guiding questions in Table 2.1, and the confidence and accuracy by which they can be answered form a continuum of scientific weight-of-evidence that will help us understand the effectiveness of the Convention.

Starting with the first guiding question, i.e., how levels of mercury in background and impacted locations compare to established benchmark values, is relatively easily addressed with a high level of confidence. When discussed together with other available information from the EE process, these mercury levels alone give valuable information for the Effectiveness Evaluation, even if the scientific links to specific measures influenced by the Convention are weak. Identifying possible temporal trends and spatial patterns adds further to the weight-of-evidence, because impacted sites can often be expected to react faster than background sites to measures influenced by the Convention, even if no formal cause-effect relationships is established. The levels of mercury and the identification of temporal trends can be determined with the sampling strategies and analytical methods (see Tier 1 below).

Estimating exposure and adverse impacts from mercury in habitats, ecosystems or populations will further help to understand the effectiveness of the Convention. This information is independently valuable for the Effectiveness Evaluation, but when combined with the methods used to estimate source attribution, the full pathway from source to impact can be described. By conducting a formal analysis that can estimate what sources are causing changes in levels of mercury in humans or the environment, the weight of evidence can be further increased, compared with just describing temporal trends or spatial patterns.

Statistical methods can be used to infer relationships between observed mercury levels and potential drivers. Mechanistic models that represent physical processes can be used to examine the consistency of these inferred relationships with what is known about other processes and to estimate mercury levels in the environment. These two methodological approaches (statistical and mechanistic modelling) can be used separately, but the strongest weight of evidence is obtained when they are used together in a complimentary manner. How accurately levels of mercury in humans and the environment can be attributed to changes in specific sources will depend on the available data, and the robustness and confidence of the used model.

The continuum of scientific weight of evidence that the guiding questions in Table 2.1 provide is reflected in the three tiers presented in this guidance. Together, answers to the guiding questions form a successive continuum of scientific weight of evidence that can form an evidentiary basis for addressing policy questions 3 and 4 as part of the Effectiveness Evaluation.

Table 2.1. Monitoring objectives and associated guiding questions.

1. *Estimation of mercury concentrations for areas without (i.e., background sites) or with (i.e., affected sites) local anthropogenic sources*
   - (a) What are the levels and form of mercury found in sites that are considered to be remote from anthropogenic sources?
   - (b) What are the levels and form of mercury found in sites that are expected to be affected by local anthropogenic sources?
2. *Identification of temporal trends*
   - (a) Do the levels and form of mercury in the observed matrix (air, biota, human) at a given location change over time – for example, in the short term (< 5 years), medium term (5 to 20 years) and long term (> 20 years)? Is there a long-term trend or trajectory (a signal) that can be separated from the temporal variability (noise)?
   - (b) How do observed temporal variations and trends differ spatially, and how do they differ among matrices?
   - (c) How do observed temporal variations and trends in mercury compare to, or co-vary with, variations and trends of:
     - (i) Mercury in different forms (chemical species) or within other matrices?
     - (ii) Mercury emissions and releases?
     - (iii) Related pollutants/emissions or environmental variables?
3. *Characterization of spatial patterns*
   - (a) What are the levels and form of mercury in the observed matrix (air, biota, human) at a given location and time?
   - (b) Taken together, what does the available data suggest about:
     - (i) Spatial variability in environmental mercury concentrations?
     - (ii) Variability in mercury concentrations within and among human populations, wildlife populations and their habitats, and ecosystems?
   - (c) Do the observed spatial variations and patterns differ among:
     - (i) Forms (chemical species) of mercury?
     - (ii) Air, biota and human matrices?
   - (d) How do the observed spatial variations and patterns or gradients compare to those of:
     - (i) Mercury emissions and releases?
     - (ii) Related pollutants/emissions or environmental variables?
4. *Estimation of source attribution*
   - (a) Using models and statistical analyses consistent with observational data, how can the observed levels, temporal trends, spatial patterns and adverse impacts on species, ecosystem services, biodiversity and human populations be attributed to changes:
     - (iii) In mercury of anthropogenic, legacy or natural origin?
     - (iv) In anthropogenic sources (local, regional, global) of mercury?
     - (v) Influenced by the Convention?

(vi) Not influenced by the Convention?

5. Estimation of exposure and adverse impacts

(a) How do the observed levels of mercury in air, biota and humans compare to established national and international benchmark levels associated with adverse effects on human health, wildlife and environmental sustainability?

(b) How significant are the observed changes in exposures for different types of impacts on humans and wildlife in regions that are remote from sources, as well as those that are locally impacted by anthropogenic sources?

(c) Are observed changes in exposure attributable to mitigation measures or changes influenced by the Convention?

6. Quantification of key environmental processes to improve understanding of cause-effect relationships

(a) How do ancillary measurements contribute to establishing the level, spatial pattern or temporal trends of mercury and improve understanding about the relative importance of environmental processes and parameters driving transport and fate?

(b) How consistent are the observed levels, temporal trends and spatial patterns with the modelled estimates and what lessons can be learned from them to improve the existing models?

2.3. Monitoring matrices

Mercury is a chemical of global concern owing to its long-range atmospheric transport, persistence in the environment, and ability to biomagnify and bioaccumulate in ecosystems leading to adverse effects on human health and the environment. For the purpose of evaluating the effectiveness of the Convention, in the light of its objective (i.e., "to protect the human health and the environment from anthropogenic emissions and releases of mercury and mercury compounds"), it is important to monitor temporal and spatial changes in the movement of mercury from its sources to the environment and into human populations. As such, air, biota and humans have been identified as the key matrices for tracking mercury (Figure 2.1) and designing a monitoring strategy to link observed changes in these matrices will be important to evaluate the effectiveness of the Convention.

Air: Mercury levels in the atmosphere are linked to mercury emissions from natural and anthropogenic sources. Key anthropogenic sources of atmospheric mercury include point sources listed in Annex D of the Convention, the intentional use of mercury in artisanal and small-scale gold mining (ASGM), and in certain industrial products and processes. In the context of the effective evaluation, it will be relevant to estimate how large the contribution of sources in Annex D is as compared to total anthropogenic emissions. Many of the Convention measures to control mercury supply, use, emissions, storage, and disposal are expected to reduce levels of mercury in the atmosphere.

Biota: Mercury released into the environment may be converted to other forms, such as methylmercury, and accumulated in fish and wildlife, and it can negatively impact fish, wildlife and human health through the consumption of contaminated prey and food. Mercury may also cause significant negative impacts to the environment, for example by adversely impacting ecosystem services and leading to the loss of rare species and potentially biodiversity. Article 22 of the

Convention requires the COP to establish arrangements to provide monitoring data on the trends in levels of mercury observed in biotic media.

Human biomonitoring: Human health may be negatively impacted by mercury exposure. Human populations may be exposed to (a) elemental and inorganic mercury in occupational settings (e.g., in ASGM and dentistry), from contact with certain products (e.g., dental amalgams, some skin-lightening creams, broken fluorescent bulbs and other waste products), and from environmental contamination; and (b) organic mercury largely from dietary sources (including but not limited to shellfish, fish, and marine mammals contaminated with methylmercury). Human biomonitoring (i.e., measuring mercury levels in the blood, hair and/or urine of individuals from a target population, depending on the form of mercury exposure) provides direct information on human exposures to mercury, from which risks to human health can be assessed. Article 22 of the Convention requires the COP to establish arrangements to provide monitoring data on the trend in mercury levels in vulnerable populations.

Other matrices: Available monitoring data for environmental matrices such as freshwater, sediments, vegetation, snowpacks, soils, and oceans may also be useful, in certain contexts, to support the Effectiveness Evaluation. Levels of mercury in freshwater may be helpful to assess environmental contamination in a given area. However, due to the complexity of tracking mercury contamination in biota resulting from water contamination, direct measurements of mercury concentrations accumulated in fish, marine mammals, sea turtles and birds offer a more practical indicator for assessing environmental contamination. Soil and sediment monitoring can provide data for an assessment of the local environment, especially in heavily polluted areas, and information from such monitoring may be used to inform the effectiveness of specific measures in certain areas, such as those addressing ASGM. As the commercial seafood market is the largest source of human exposure to MeHg on a global basis, levels of mercury in the oceans' surface may also contribute to the assessment of global environmental transport of mercury.

Based on these considerations, this document provides guidance on air monitoring, biota monitoring and human biomonitoring. Examples of ancillary measurements from water and sediment samples associated with biota monitoring are included in chapter 4. Review of environmental monitoring methods that may be used to assess ASGM sites is also being developed[7] to support the implementation of Article 7 of the Convention. An overview of existing monitoring programs is provided in the Supplementary Material to this guidance.

[7] The Secretariat is developing a separate document which will be made available for consideration by the fourth meeting of the COP.

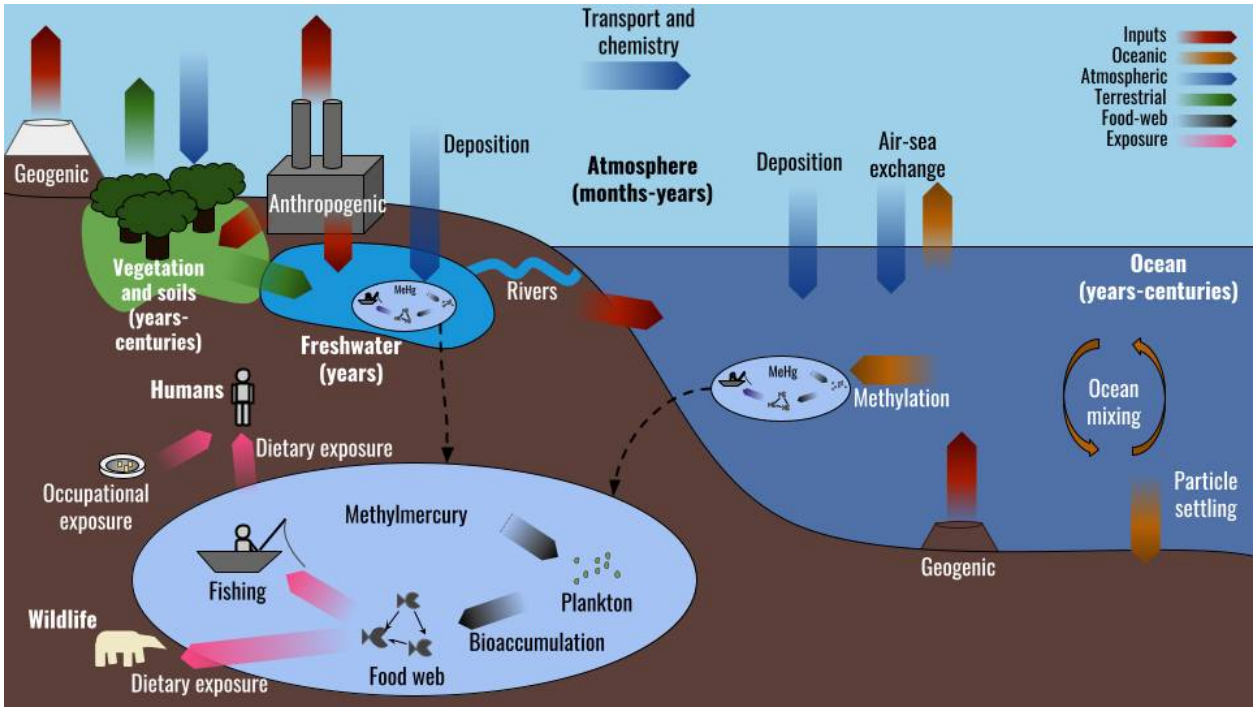

Figure 2.1: Flows of mercury among matrices. Arrow colours indicate which models simulate those flows (red - model inputs, blue - atmospheric models, green - terrestrial models, orange - ocean models, black - food web models, magenta - exposure models), while values in parentheses indicate relevant time scales in each matrix (C. Thackray, unpublished).

2.4. Tiered approach for developing and improving monitoring programmes

To support Parties and organizations who may wish to develop new monitoring programs, or improve existing ones, with a view to contributing to the Effectiveness Evaluation, this document identifies a tiered approach for monitoring each of the three media (air, biota, humans):[8]

- **Tier 1** is intended to provide guidance on mercury monitoring under a limited set of parameters for circumstances where available resources are not sufficient to implement the actions in Tier 2. Following guidance by the COP,[9] the methods in Tier 1 are cost effective, practical, feasible, and sustainable. The Tier 1 methods are intended to provide information that are useful in identifying and characterizing gaps and needs of national, regional, or local interest and to provide information that is useful to the collective effort for the Effectiveness Evaluation. While the implementation of Tier 1 actions may not fully address the questions in Table 2.1, it will contribute essential information and create a foundation for Tier 2 monitoring.
- **Tier 2** is intended to build upon Tier 1 methods to provide information that will address the questions identified in Table 2.1, and to create a basis for assessing source attribution at the local, national, and global scales. The methods and approaches in this tier may be more expensive or complex than those under Tier 1. The more comparable data from Tier 2 becomes available, the more robust the Effectiveness Evaluation will be.
- **Tier 3** identifies research methods and approaches that may play a vital role in supporting the Tier 1 and Tier 2 programs and the Effectiveness Evaluation, primarily by improving our

[8] It is noted that the Convention does not impose any obligation upon Parties to conduct monitoring. As such, the tiered approach and any other activities or recommendations contained in this guidance are voluntary and presented with the sole purpose of supporting Parties who may wish to develop new monitoring programs, or improve existing ones, with a view to contributing to the Effectiveness Evaluation.

[9] Decision MC-2/10 pursuant to the terms of reference to Ad-hoc Technical Expert Group on Effectiveness Evaluation.

understanding of key processes that link sources to environmental concentrations and exposures. Because Tier 3 focuses on processes, the results would likely yield insights that are broadly applicable and that should be taken into consideration in the Effectiveness Evaluation when available.

2.5. Quality of monitoring data

The Effectiveness Evaluation of the Convention will require monitoring data that is comparable and credible. The Quality Management Systems (QMS) employed by existing monitoring programs typically include Quality Assurance and Quality Control (QA/QC) measures that will provide a basis to inform the development of comparable data for use in the Effectiveness Evaluation. Data generated from different monitoring programs may be supplemented, as appropriate, with comparable data from academia and research. This may be accomplished through a well-documented and transparent set of "data flags" that will enable the use of data from different sources with different levels of QA/QC.

Understanding the presence and movement of mercury in the environment and in humans will require implementation of different monitoring programmes that yield good quality data via QA/QC systems and protocols. Such programmes need to be well documented data to enable their data to be compared with other data for the purposes of the Effectiveness Evaluation.

Examples of criteria to assess the quality of mercury monitoring data include:

- Selection bias – describe the location/population/setting (Is sampling performed in a consistent manner, representative of the target location/population/setting, and bias-free?);
- Exposure detection – describe how mercury was measured in a given sample (Is the measurement value accurate and precise, and does it follow a standardized and scientifically credible method? Are additional metadata[10] provided to do exposure/attribution assessment?);
- Statistical parameters – describe the sample size and its adequacy, and whether basic and essential data is present, complete and well summarized (Is the data useful to address the questions in Table 2.1?).

In order to determine adequate sampling site location, sample sizes, site densities, sampling frequencies, or observational period needed to detect expected changes in the context of the Effectiveness Evaluation, an analysis based on the spatial and temporal variability of existing data and model projections of potential changes may be necessary. Chapters 3, 4 and 5 describe specific QA/QC considerations for each medium or matrix.[11]

2.6. Data management

To the extent possible and in accordance with requirements of individual data providers, data used in the Effectiveness Evaluation should follow the FAIR principles (findable, accessible, interoperable and reusable)[12] for data management and stewardship, including the following elements:

Findable:

- A searchable and interoperable database acting as a repository of available data;
- Unique identification systems (e.g., "Digital Object Identifiers" or "DOIs") and controlled vocabulary to facilitate searching and retrieval of information;

[10] The term "metadata" refers to data that provides information about other data (source: Merriam-Webster Dictionary).
[11] The terms "medium/media" and "matrix/matrices" are being used interchangeably throughout this document.
[12] Wilkinson et al. (2016).

- Detailed metadata associated with each data record to facilitate the submission, searching, location and retrieval of information;

Accessible:

- Free and open access to the data to Governments, Indigenous Peoples, and relevant stakeholders, taking into account the relevant ethical considerations;

Interoperable:

- An interoperability mechanism to facilitate the exchange of information across different programmes and databases;

Reusable:

- Data usage license/agreement identifying the terms and conditions for further use of the data;
- Metadata including enough information describing how the data were collected/produced to enable an assessment of the quality and comparability of the data, reproducibility and further analyses.

Further to the FAIR principles to facilitate increased data sharing, principles to support an ethical use of data should also be followed (see chapter 5). Ethical considerations associated with Indigenous Peoples, including with regards to self-determination of research, research ethics, data considerations, utilization of Indigenous Knowledge, and communication of results, should be guided by principles such as the "CARE Principles for Indigenous Data Governance".[13]

Major international and national monitoring programmes and networks have their own data management systems, including data repositories, portals, or catalogues. Some of these may be used as primary repositories, along with other sources of data, including Indigenous Knowledge in mercury monitoring efforts, to gather and exchange information so that monitoring data, relevant metadata, ancillary data and QA/QC information from different programs can be used for the Effectiveness Evaluation. There are also global initiatives to develop monitoring data management platforms that enable access to data from multiple monitoring programmes, networks and existing primary data repositories. An overview of the existing data management systems and networks will be provided in a Supplementary Material to this guidance.

[13] CARE Principles for Indigenous Data Governance: https://www.gida-global.org/care.

Chapter 3. Atmospheric Mercury Monitoring

3.1. Introduction

Monitoring of atmospheric mercury has been identified as one of the primary and most appropriate types of monitoring to help evaluate the Convention's effectiveness (AMAP 2011; Evers et al. 2016; Gustin et al. 2016; Sprovieri et al. 2016; UNEP 2019). Monitoring of atmospheric mercury generally has the following objectives: 1) to gather information on temporal and spatial trends of changes in mercury concentrations in the atmosphere; 2) assessing atmospheric mercury inputs to aquatic and terrestrial ecosystems and 3) to provide data for the development and improvement of transport and chemistry models (Obrist et al. 2018; Saiz-Lopez et al. 2020; Skov et al. 2020). Monitoring may also provide insight on how climate change affects, and is affected, by atmospheric mercury (Li et al. 2020).

This chapter aims to provide guidance to Parties and organizations on the different methods to monitor mercury in air, and what procedures need to be in place to continue, expand or start monitoring mercury in air to generate comparable monitoring data to support the effectiveness evaluation. Given the differences in capacity and objectives among monitoring programmes, a tiered approach offers a way forward to support Parties and organizations in starting or expanding their monitoring programmes (section 3.3 and annex to the present document).

3.2. Significance of air as a matrix for mercury monitoring

Mercury is a naturally occurring element and is emitted to the atmosphere from a variety of natural, geogenic and anthropogenic sources. Mercury that has been deposited on land and ocean surfaces can be re-emitted to the atmosphere, through so-called "legacy emissions" (Driscoll et al. 2013). Atmospheric deposition represents the major pathway of mercury input to terrestrial and aquatic ecosystems outside areas with substantial sources that directly release mercury to land or water.

The temporal and spatial scales of mercury transport in the atmosphere and its transfer to aquatic and terrestrial ecosystems depend primarily on its chemical and physical forms. There are three forms or species of mercury commonly found and measured in the atmosphere, namely:

- Gaseous elemental mercury (GEM),
- Gaseous oxidized mercury (GOM)
- Particle-bound mercury (PBM)

with the last two species being operationally defined.[14]

In addition to the above, atmospheric mercury can also be measured as "total gaseous mercury" (TGM) which equals to GEM and GOM combined and as reactive mercury which equals to GOM and PBM combined (Martin et al. 2017).

Mercury is transformed in the atmosphere, cycling between GEM, GOM, and PBM, through gas and aqueous photochemical reactions. The apparent or effective atmospheric lifetime of GEM is relatively long, whereas GOM and PBM have short residence times in the atmosphere. The overall atmospheric lifetime of TGM against deposition is approximately 6 months (Horowitz et al. 2017; Skov et al. 2020). This, together with the influence of re-emissions of previously deposited mercury (and the presence of geogenic sources), means that the change in atmospheric concentrations or deposition is likely to be less than proportional to the change in primary anthropogenic emissions.

[14] Operationally defined species: Their chemical and physical structure cannot be exactly identified by experimental methods but are instead characterised by their properties and capability to be collected by different sampling equipment (Schroender et al. 1998).

Furthermore, some fraction of the change may be observed immediately, but the full benefits of the abatement will occur over time (Sprovieri et al. 2016). The lag time (in the atmosphere) is expected to be much shorter than the time response for mercury in other reservoirs (soils, surface waters, biota and ocean) where residence times are much longer (approximately decades) and mercury levels are complicated by several other factors (Lyman et al. 2020). Therefore, at large regional scales, the atmosphere is expected to be one media where changes in environmental levels due to changes in emissions influenced by the Convention will be reflected earlier, than in other matrices. Atmospheric monitoring (ambient air concentrations and atmospheric deposition) can thus be seen as one scientifically sound approach to help evaluate the Convention's effectiveness (Gustin et al. 2016).

An overview of existing programmes and networks for monitoring atmospheric mercury, including standard operating procedures, and data gaps is available in the Supplementary Material to this guidance.

3.3. Tiered approach for atmospheric mercury monitoring

The tiered monitoring approach presents a framework to identify and prioritize monitoring needed to 1) determine whether mercury concentrations in air are changing over time, and 2) whether observed changes in concentrations may be attributable to controls on emissions, releases, supply and use of mercury effected by the Convention. This approach seeks to build-on/expand existing mercury air and wet deposition monitoring efforts to promote consistency in data collection and advance collaboration across sampling activities. The guidance offered in this chapter describes criteria to consider when deciding which measurement methods to use, the frequency of measurements, and where to potentially locate sites depending on the monitoring programme needs and objectives. For mercury in air monitoring the Tiered approach can be seen as follows:

Tier 1 – The objectives of this tier are to estimate background and impacted levels and identify temporal trends. It documents trends and spatial distribution in air (TGM/GEM) and in wet deposition over broad geographic areas and provides information to inform atmospheric modelling (statistical and mechanistic). The measurement methods included in this tier are cost effective, practical, feasible, and sustainable. Tier 1 offers an entry point for Parties and relevant organizations who wish to pursue one or more of these sampling options (e.g., automated, manual, passive), where feasible, to contribute to the Effectiveness Evaluation.

Tier 2 – The objectives of Tier 2 include source attribution, characterization of spatial patterns, and estimation of exposures of humans and wildlife. This tier explains temporal trends and attributes mercury sources to mercury concentrations in other matrices. Sampling is more intensive than Tier 1 sites and includes more ancillary measurements. Tier 2 will also enable "top-down" attributive analysis of TGM/GEM levels, using speciated mercury data and air-quality tracers. The more Tier 2 information are available, the more robust the Effectiveness Evaluation will become.

Tier 3 – In Tier 3 the observational strategy is designed in such a way as to understand key processes affecting Hg fate and transport. This tier improves representativeness of the measurements and understanding of key processes (e.g., related to transformation and deposition) using new, advanced measurement techniques and intensive research. Where Tier 3 efforts and results are available, the information should be taken into consideration in the Effectiveness Evaluation.

3.4. Atmospheric mercury deposition

After mercury is emitted into the atmosphere, it eventually returns to the Earth's surface via wet and dry deposition processes. The pathway for wet deposition occurs when mercury is deposited with precipitation.

The amount of precipitation is the main driver of wet mercury deposition to aquatic and terrestrial habitats but, in areas with a lot of rainfall, the amount of deposition may be limited by the availability of atmospheric mercury (Prestbo et al. 2009; Weiss-Penzias et al. 2016; Sprovieri et al. 2017). Forest and land fires also have an impact on the mercury release to the atmosphere and on wet deposition. GOM and PBM are water soluble and the primary atmospheric forms responsible for wet deposition of mercury in precipitation.

Wet deposition can be measured directly by collecting precipitation such as rain or snow and measuring the amount of mercury relative to the quantity of precipitation. Monitoring mercury in precipitation is an important way of determining inputs of mercury into aquatic and terrestrial ecosystems (Aas et al. 2019; Sprovieri et al. 2017). When compared to automated atmospheric mercury monitoring, wet deposition monitoring is relatively easy to start off with and is considered a very reliable method to achieve comparable data across different monitoring initiatives or programmes (Brown et al. 2010; Sheu et al. 2019). It should be noted that total mercury deposition is determined by both wet and dry deposition (Brown et al. 2010). Standardized sampling equipment deployed and consistently operated across many locations coupled with consistent sample analyses will help facilitate the production of comparable mercury wet deposition measurement data. Several national mercury wet deposition networks described in the supplemental material offer examples of systematic mercury wet deposition networks for parties and other organizations to consider.

In the absence of precipitation, dry deposition is the transfer of atmospheric mercury (either a gas or a particle) to vegetation, soil, water, and snow, controlled by the characteristics of the atmosphere, the surface, and the mercury species. While no techniques are available to routinely measure mercury dry deposition in a network configuration, it is possible to estimate dry deposition on the basis of GEM and GOM measurements using a number of methods, such as surrogate surface, litterfall, and throughfall measurements (Wright et al. 2016 and references therein). More research is needed to resolve uncertainties in measurements based on surrogate surfaces, micro-meteorological methods, and dynamic flux chambers, which may be included in Tier 3 activities. Inferential approaches to estimate dry deposition and bi-directional air-surface exchange models are recommended for use in Tier 2 activities, acknowledging that these methods are subject to uncertainties in measurements and underlying assumptions (Lyman et al. 2020; Zhang et al. 2016, 2019). Since vegetation can uptake Hg^0 from air and litterfall is an important Hg dry deposition pathway, it is suggested that at forest sites the litterfall should be monitored to calculate Hg dry deposition fluxes (Wang et al. 2016, 2019). It should be noted that dry deposition is significant during mercury depletion episodes in the Arctic, and it is dominating in dry arid climates (Steffen et al. 2014, 2015).

3.5. Forms of mercury

Quantification of atmospheric deposition of the different forms of mercury to various surfaces is useful in assessing the impacts on the environment and human health. Following emission, GEM can be transported long distances before oxidation and/or removal by particle and gas-phase dry deposition or scavenging by precipitation. Due to its relatively long residence time in the atmosphere, GEM can be transported and deposited to remote locations such as the Arctic (Skov et

al. 2004; Sprovieri et al. 2005; Sprovieri et al. 2010) and Antarctic (Dommergue et al. 2010; Angot et al. 2016).

GOM and PBM have shorter atmospheric lifetimes than GEM and as a result are generally deposited closer to emission sources (Hedgecock et al. 2006; Jung et al. 2009). Oxidized mercury compounds often have a more local impact than elemental mercury because they are water-soluble, are more reactive and thus deposit more quickly (Hedgecock and Pirrone 2004; Fu et al. 2015; Lyman et al. 2020). Measurements of mercury species such as oxidized and particle-bound mercury compounds are important as they help to improve the understanding of short-term oxidation processes regarding the removal of mercury from the atmosphere (Pirrone et al. 2009; Fu et al. 2015; Weiss-Penzias et al. 2015; De Simone et al. 2016).

Mercury speciation measurements are made in various networks using standard operating procedures which are widely available. Comparison studies between different monitoring networks have delivered satisfactory results (Steffen et al. 2012; Gustin et al. 2015; Sprovieri et al. 2016). Although they are very important in helping to understand the global mercury cycle as well as improving model output, speciation measurements are quite complex, costly, and require very skilled operators to avoid analytical interferences in GOM and PBM measurements (Gustin et al. 2015). Furthermore, recent information also demonstrates that GOM might be underestimated by this method (Gustin et al. 2020). Therefore, for the purpose of evaluating the effectiveness of the Minamata Convention, GOM and PBM are not recommended for monitoring in Tier 1. Several monitoring networks and research groups perform mercury speciation measurements in a comparable manner and are encouraged to share these results, as their data will be helpful in answering questions for the Effectiveness Evaluation. Further scientific work on these methods will improve understanding of biases in existing methods and comparability across measurement techniques.

3.6. Monitoring sites

Site characteristics can affect the concentration levels of mercury in air. Therefore, site selection is a critical part of any monitoring network's design (Schmeltz et al. 2011). The selection of monitoring sites to support the Effectiveness Evaluation of the Minamata Convention and help address policy questions 3 and 4 listed in the previous chapter, should be based on the sites' potential to provide insights into changes in atmospheric mercury levels, to assess levels in sensitive or vulnerable ecosystems (e.g., the Arctic), and to help evaluate atmospheric models.

A diversity of site locations should be considered, as different types of sites may provide different types of information, including (a) background or remote, (b) rural, (c) urban and (d) contaminated or industrial sites. Background and remote sites, including forests not impacted by activities such as ASGM, provide information on determining long-term global trends and provide data for evaluating and refining transport models; rural sites provide information on mercury concentrations that are regionally representative as long as the influence of significant local pollution sources is limited; urban sites can provide information on non-point sources (pollution that comes from many places, all at once) such as Hg transported in cities, and will be useful to improve emissions inventories; and contaminated or industrial sites will assist with determining the effect on human exposure of communities living close to point sources (e.g., ASGM activities).

Mercury air and deposition monitoring site locations should prioritize co-location with other existing monitoring activities, including air quality sites and weather stations, where relevant air pollutants are monitored, such as ozone, sulfur dioxide, carbon monoxide, fine particulate matter ($PM_{2.5}$) and halogens, to make use of available infrastructure and co-measurements with relevant reactants.

Mercury air and deposition monitoring should also prioritize collocation with monitoring of Hg in other media to allow linkages to be quantified. Furthermore, in addition to spatial coverage, temporal coverage may also be a relevant element in the selection of monitoring sites so that areas with historic monitoring data, i.e., from a period before the implementation of the Convention, can be used as baseline.

One concept that has been widely adopted in most proposed monitoring program structures is to have so-called "intensive" sites where detailed measurements are made and have these sites then coupled with a number of other sites ("cluster" sites) that can expand the data collection regionally and spatially (Harris et al. 2007). For atmospheric Hg, the intensive sites could include atmospheric mercury speciation measurements (gaseous Hg speciation, aerosols, and wet deposition), and therefore would need the infrastructure to support such measurements, while at the cluster sites only TGM and weekly wet or bulk deposition could be measured. Given the development of passive sampling techniques, these could be used at the cluster sites allowing for the collection of data in remote areas where electricity is limited. The development of "cluster sites" can be viewed as following a Tier 1 approach and will be useful for Parties and Organizations with limited infrastructure or no monitoring program. Cluster sites will also provide a good baseline where no data is available. The number and location of the monitoring sites in each region is expected to be determined based on the distribution of major urban areas, and take into account differences in the vegetation and should include coastal as well as inland sites (Evers et al. 2008a). For example, Schmeltz et al. (2011) suggested site locations in the USA should be determined by the major ecoregions, as well as determined by other factors such as the likelihood for the site to experience change in concentrations over time. Additional criteria include their usefulness for model evaluation, and their importance in terms of evaluation of health risk to humans and wildlife (e.g., the Arctic). As noted above, intensive sites should also collect ancillary information (wind direction and speed, air temperature, relative humidity, solar radiation, precipitation) besides mercury measurements. Some ancillary data may be especially useful for estimating dry deposition and, where possible, they also should be measured at existing Hg monitoring sites under ongoing national, regional or global mercury monitoring programs.

During the past two decades, a number of mercury monitoring sites have been established in Europe, North America and Asia as part of regional or global monitoring networks (Pirrone et al. 2003; Steffen et al. 2008; AMAP 2011, 2021; Fu et al. 2012; Cole et al. 2014; Sprovieri et al. 2016) using the information presented above as guidance when new sites were established. The need to establish a global network to assess likely southern-northern hemispheric gradients and long-term trends has long been considered a high priority for policy and scientific purposes in order to evaluate the impact of mercury pollution. Consistent globally distributed Hg observations will help reveal trends in mercury concentration and deposition in different regions of the world. Building on and expanding existing infrastructure will help improve the process of evaluating the effectiveness of the Convention. Examples of global, regional and national monitoring programmes are provided in the Supplementary Material to this guidance.

3.7. Sampling and measurement methods

Several different methods are available for monitoring of atmospheric mercury. Selection of methods should be based on the purpose of monitoring. All methods employed in a monitoring program need to have been tested, intercompared and validated to ensure quality of data used for the Effectiveness Evaluation. This section aims to facilitate the selection of monitoring techniques for Tier 1 and Tier 2 that best meet monitoring objectives, taking into account the availability of resources and logistical constraints.

3.7.1. Active sampling

Active air sampling methods involve ambient air pulled through a pump at a constant flow rate through an active material, otherwise known as a trap, with laser and Zeeman-AAS techniques being the exception. The active material contained in these traps is often gold, but other materials such as sand mixed with gold or carbon are also used. Once the sample has been collected on the active material for a set amount of time, the mercury adsorbed is removed using thermal desorption and spectroscopic detection.

Active sampling of this kind can be undertaken in an instrument that both collects the sample and performs the analysis in situ (automatically) or can be undertaken by collecting the sample actively and then performing the analysis at a separate location or laboratory. The differences in the methods are explained below as (a) automated Mercury Air Measurements; and (b) manual mercury air measurements. Further information on instruments used for atmospheric mercury monitoring, including their advantages and their disadvantages, is available elsewhere (e.g., Gustin et al. 2015).

(a) Automated mercury air measurements

Currently, automated mercury measurements, which can be used for GEM/TGM and speciation measurements (see section 3.2.1), are recorded with different commercial instruments available from various manufacturers that are capable of detecting mercury at very low concentrations in air, i.e., nanograms or micrograms of gaseous pollutant per cubic meter of ambient air. These instruments have high temporal resolution, low limits of detection, established and proven quality assurance and quality control protocols (Angot et al. 2014; Gustin et al. 2015; Sprovieri et al. 2016; Slemr et al. 2020).

The following spectroscopic detection techniques are most commonly used for automated monitoring of mercury in air, either as total gaseous mercury (GEM + GOM) or as gaseous elemental mercury (GEM) are (USEPA 2006):

- Cold Vapour Atomic Absorption Spectroscopy (CV-AAS)
- Cold Vapour Atomic Fluorescence Spectroscopy (CV-AFS)

Mercury is present in very low concentrations in air (as nanograms of mercury per cubic meter of air (ng/m^3) at standard pressure and temperature), and it has a particularly strong absorption/emission line at 253.7 nanometer (nm). In most cases and with the exception of Zeeman AAS and laser techniques, the sensitivity for AAS and AFS is not sufficient for direct measurements of ambient concentrations and, therefore, mercury is pre-concentrated on gold coated surfaces (Gustin et al. 2010; Amos et al. 2012; Gustin et al. 2015). Zeeman AAS method can be applied in urban, regional, and rural sites for automated total gaseous mercury air measurements. Both AAS and AFS are subject to interference by molecular species (e.g., ozone, sulphur dioxide, organics) that absorb in the ultraviolet (UV) range close to 253.7 nm. AFS is less influenced by these molecular interferences than AAS is and does not require any sort of correction scheme. AFS is more sensitive in comparison to AAS, but it requires pure Argon (Ar) or Helium (He) gas during the desorption and detection step, whereas AAS uses mercury free air or nitrogen instead (Gustin et al. 2015).

The pure gold trap method with AFS has been chosen by national and international networks for use at background and regional sites and is cited extensively in the scientific literature (Sprovieri et al. 2016; Martin et al. 2017; Xu et al. 2017). Other methods may be good choices for applications where low detection limits are not required, where there is less concern about interference, and a lower QA threshold is set. Such applications may include monitoring near artisanal gold mining or industrial locations where levels are expected to be high. Automated instruments are used widely within different monitoring networks and programs and can generate data that will be comparable

(Steffen et al. 2012; Sprovieri et al. 2016). The cost (investment and running costs) associated to the use of automated analysers are substantially higher compared to other methods (passive samplers, wet deposition, and manual method). However, these instruments are able to deliver high frequency data in a short time span from as little as 5 seconds to 5 minutes.

(b) Manual mercury air measurements

With this technique, mercury in the atmosphere is collected manually on an adsorbent material over 24-hour periods, or one week, at a constant flow rate using a pump. The sampling instruments normally use a gold tube scavenger/impregnated active carbon trap and small air pump, and are portable. Unlike automated sampling, manual active sampling method is analysed by using an analytical instrument which, in most laboratories, is installed separately from the sampling equipment. For this reason, this technique is less site restricted. This method is relatively easy to setup and operate but requires a covering or shield to protect the trap from contamination or interference by weather elements.

Analysis of TGM in ambient air is possible when the sample is analysed after exposure using thermal desorption and spectroscopic detection in a laboratory. Additionally, since the scavenger material that is generally used in manual active sampling (e.g., gold cartridge) catches almost all amount of mercury in air, the manual active sampling has very little or almost no isotope effect.[15] Thus, collected air sample by manual active method is useful for isotope ratio analysis.

The measurement principle applied for manual active method is similar as that for automated active method and interferences are caused by same environmental factors (e.g., by ozone and high humidity). The data from this technique can be calculated from a calibration curve based on measurements of mercury gas that are accurately collected from a saturated mercury gas generator (bell jar calibration). Minimum time resolution is dependent on the measurement instrument used (around 3 hours for CVAAS and 1 hour for CVAFS depending on analytical condition and concentration of the air) and, in most cases, between the resolution of automated active and passive sampling methods. Analytical instruments are costly, but this method is cost-effective because scavenger materials are low cost and analytical instruments can be shared while being able to place sampling equipment in multiple sites.

The temporal resolution of this method is often lower than that for automated mercury monitoring, but 24h-average is suitable for long-term trend analyses. It should be noted that achieving quality results with this method, would require consistent low blanks and an operator with trace-clean technique experience. Interference from ozone and high humidity can also influence the performance. QA/QC procedures, such as field blank, co-located sampling instruments and staff proficiency, are important because multiple equipment and personnel are involved. General approaches for these QA/QC procedures are well documented (Munthe et al. 2001; Brown et al. 2010; MOEJ 2011), and methodological comparisons for automated active and passive sampling are in progress (e.g., in Japan).

3.7.2. Passive sampling

The development of methods for passive air samplers (PAS) for gaseous mercury has increased recently. While it is not possible to produce data at the same temporal resolution as automated or manual instruments (McLagan et al. 2016), PAS have been shown that they can produce air concentrations of mercury accurately and comparably with active air monitoring methods. PAS can increase spatial resolution of air concentration data and contribute to Hg source characterization.

[15] "Isotope effect" refers to the variation of certain characteristics, such as density and spectrum, of an element in accordance with the mass of the isotopes.

PAS work by uptake through a diffusive surface and accumulation of gaseous Hg onto an adsorbent scaffolding. The peculiarity of the passive samplers relies on the unassisted molecular diffusion of gaseous agents following Fick's First Law (i.e., volatile vapours of elemental mercury). Unlike actively pumped sampling, passive samplers require no electricity, have no moving parts, no pump operation or calibration, and are simple to use and low cost (Macagnano et al. 2018). PAS can be deployed at background, remote, urban, hotspots sites and without worry about media failure as a new sampler with active material is used each time.

After exposure, the PAS can be analysed with well-documented and credible methods in most analytical laboratories (Wängberg et al. 2016; McLagan et al. 2016, 2017, 2018; Macagnano et al. 2018). It should be noted that depending on the exposure period chosen, PAS have lower temporal resolution than other methods and its sampling rate can be affected by wind speed and temperature, which must be factored into the calculations (McLagan et al. 2017). Quality control of the samplers is necessary even when a new product is used (i.e., it cannot be left to providers), and appropriate QA/QC process is important for the entire survey and analysis, including at laboratories. Laboratory measurements of passive air samples are not as easy as other types of samples and, while interlaboratory comparisons may be challenging, a recent study shows good insight into how to undertake intercomparisons of data generated with different PAS (Naccarato et al. 2021). Moreover, PAS have shown to be useful for low level monitoring as well as at high concentration sites or hotspots, and they can assist with understanding contaminated sites emissions to air, specifically describing concentration gradients.

For initiatives that have no mercury air monitoring program or previous mercury air experience, passive sampling is considered a suitable method to start with (Tier 1; see tiered approach in chapter 2 and annex) or to complement to a limited number of active analysers. In the context of the Effectiveness Evaluation, if deployed at the relevant sites, PAS will contribute to answering questions related to spatial variability, trends and emissions and impacts on local ASGM communities.

3.7.3. Wet deposition sampling

The amount of mercury depositing from the atmosphere to the ecosystems provide a useful proxy for mercury emissions (Selin 2018; Travnikov et al. 2017). Mercury wet deposition sampling, which collects all water-soluble atmospheric mercury species (Hg^{II} + MeHg), can be carried out using a variety of commercially available precipitation collectors for either wet only or bulk collection with wet only sampling being the preferred method (Prestbo et al. 2009; Brown et al. 2010; Gichuki et al. 2013; Brunke et al. 2016; Weiss-Penzias et al. 2016; Risch et al. 2017). Wet only samples are collected using polytetrafluoroethylene tubing, borosilicate funnels, and polyethylene terephthalate glycol (PETG) bottles for sampling and shipping. When an event stops, the sensor dries out, and the lid returns to cover the funnels. For bulk collection, the sampler is open to the atmosphere constantly, and this method has delivered satisfactory results (Brunke et al. 2016; Sheu et al. 2019). The wet-only samplers have the advantage of avoiding particle dry deposition although the contribution to the measured wet deposition fluxes from gaseous or particulate mercury species is probably not large in non-industrialised or non-urban areas. For extended sampling periods it is also necessary to prevent significant gas phase diffusion of Hg^0 to the surface of the collected sample where it could contribute to the mercury content of the sample via oxidation to water-soluble forms. This can be easily done using a capillary tube between the funnel and the bottle. Shielding of the sample bottles from light is also necessary to avoid photo-induced reduction of the mercury in the precipitation sample. Within most Networks, wet deposition samples have been collected weekly (Prestbo et al. 2009; Brunke et al. 2016) but semi-weekly collection is also used (Sheu et al. 2019). Sampling frequencies provide data that are useful for quantifying total deposition or understanding

longer-term trends addressing key objectives as outlined in Chapter 2 of the Guidance Document. It is beneficial to do wet deposition sampling in close proximity to where automated or manual mercury air monitoring is taking place. The benefits of doing this, is that studying the correlation (if any) between total mercury concentrations in rainwater and mercury concentrations in surface-level ambient air will be possible (Fu et al. 2015; Brunke et al. 2016; Sprovieri et al. 2017). If only wet deposition sampling is possible, it is advisable to have a variety of sites located throughout a country or network (see section 3.4). Various studies have indicated that urban/industrial locations tend to have higher mercury wet deposition than rural/remote locations, but this association can be weak as atmospheric processes, not just local emissions, are important drivers of mercury uptake by precipitation (Sprovieri et al. 2010). The influence of emission sources on spatial trends in mercury wet deposition have been observed, even at the global scale (Fu et al. 2012; Weiss-Penzias et al. 2016) and this showed that wet deposition follows temporal trends in global and regional anthropogenic mercury emissions (Obrist et al. 2018). Therefore, when selecting a location for wet deposition sampling, the above factors should be taken into consideration.

When starting with wet deposition monitoring, the following factors should be considered when choosing a sampling location: (a) availability of stable electricity when using a wet-only collector (wet-only collectors can run on solar and battery power in some locations, without access to power lines; bulk collectors requires no electricity), (b) access to laboratory facilities to prepare samples, including treating glassware with acids and other chemicals, and to analyze the collected samples, (c) access to proper shipping, including portable sample trays and coolers for collecting and transporting field samples, (d) skilled operator to conduct analysis, (e) availability of meteorological data at sampling site, (f) refrigerator or other storage facility for samples, (g) site selection so it is following the general recommendations for precipitation measurements from WMO.

3.8 Advance techniques for atmospheric mercury measurements

The following techniques under Tier 2 and 3 air monitoring will help to understand key processes affecting mercury fate and transport. These advanced and resource intensive measurement techniques will provide valuable information to improve understanding of Tier 1 and Tier 2 measurements and data.

3.8.1. Dry deposition sampling

Direct measurement of mercury dry deposition is technically challenging but can be done by micro-meteorological methods (Brooks et al. 2006; Skov et al. 2006). It is also possible to model dry deposition on the basis of GEM and GOM measurements (see section 3.2.2). Although no methods currently exist to measure mercury dry deposition in a network configuration, dry deposition measurements may be useful as part of monitoring activities in Tier 3 (see annex). Further expert consultation is necessary to consider how dry deposition estimates or measurements may be compared.

3.8.2. Mercury isotope measurements

Mercury stable isotopes exist naturally in the environment and have different approximate abundances (e.g., ^{200}Hg = 23.14, ^{201}Hg = 13.17%, ^{202}Hg = 29.73%, Blum 2007; Kwon et al. 2020). Precise measurement of mercury isotopes ratios is possible using multi-collector inductively coupled plasma mass spectroscopy (MC-ICP-MS). The development of protocols to measure and report mercury isotope ratios (Blum et al. 2014) have also enabled consistent data comparison among international research groups.

Measurement of mercury isotope ratios in total gaseous mercury from different environmental samples is an effective means of distinguishing between anthropogenic and natural sources of mercury in the atmosphere (Kwon et al. 2020). The ability to make this distinction may help determining whether the observed changes in atmospheric mercury concentration are influenced by implementation of the Convention's provisions on anthropogenic mercury emissions or from other drivers such as climate change and human activities not influenced by the Convention.

3.8.3. Measurements of species important in the atmospheric oxidation of mercury

Oxidation of elemental mercury is a main species driving atmospheric cycling given the solubility and relative lifetimes of mercury in its different forms (Si et al. 2018; Lyman et al. 2020). The conversion from Hg^0 to Hg^{II} is an important step in the atmospheric and biogeochemical Hg cycling. Despite this fact, questions remain as to the dominant oxidation pathway(s) in the atmosphere, largely due to uncertainties around the kinetics associated with proposed mechanisms and inability to accurately determine the chemical form of GOM in ambient air (Obrist et al. 2018; Si et al. 2018). Bromine-induced oxidation has been proposed as the globally oxidant dominant pathway and some models are still based on this assumption (Amos et al. 2012; Horowitz et al. 2017).

Direct oxidation of Hg^0 by O_3 or OH is a pathway commonly used in many global and regional chemical models even though the gas-phase reaction of Hg^0 with these oxidants may be too slow to act as the dominant oxidation mechanism in the atmosphere (Driscoll et al. 2013; De Simone et al. 2014; Wang et al. 2014; Travnikov et al. 2017; Obrist et al. 2018). Recent studies have therefore suggested that both OH and Br radicals could contribute to the oxidation of Hg^0 to Hg^I-complexes and that O_3 is responsible for oxidizing these further into Hg^{II} species. (Shah et. al., 2021). Other oxidants such as Cl, H_2O_2 and NO_3 have also been proposed as potential oxidants for Hg^0 (Si et al. 2018; Lyman et al. 2020).

The measurement of these oxidant species may help to improve the chemical reactions and rates currently employed to interpret Hg chemistry in the atmosphere, and to predict observed atmospheric mercury concentrations and deposition with a higher degree confidence, as well as to identify other pathways for mercury oxidation in the troposphere. Further investigation with these techniques will help to understand key processes affecting Hg fate and transport in the atmosphere.

3.9. Frequency and duration of sampling

The following sampling periods and conditions are guidelines for the various sampling methods and site locations depending on how the data will be used. Once the sampling period and other conditions have been selected, it is important that site operators maintain that schedule.

(a) Automated measurement with CV-AFS/AAS

- A sampling time ranging from 5 min (at most locations) to 15 min (where the average concentration is lower than 1.0 ng/m^3 and mostly for sites in the southern hemisphere).
- Hourly averages can be used for trend analysis (Tiers 1 and 2) and half-hourly averages can be used for process studies (Tier 3).

(b) Manual measurements

- A sampling time of 24 hours and a constant flow rate is sufficient as several networks are currently using this parameter and available data can be used as comparison.
- A flow rate of between 0.1 and 1.0 L/min is acceptable but once selected, should be kept constant throughout and for each sampling episode.

- Studies have indicated that if the flow rate is below 0.1 L/min, it is difficult to hold a constant flow rate due to suction pump performance and a flow rate of over 1.0L/min, increases the possibility that the collection efficiency of the gold trap may begin to decline (MOEJ 2011).
- A flow rate of 0.5L/min is commonly applied as this have shown very good correlation with automated analyzers (Marumoto et al. 2019).
- Sampling frequency (weekly, monthly, etc) should be determined appropriately based on local conditions or site description, in order that reliable long-term average concentrations in the atmosphere is obtained that is comparable to data from automated sampling.

(c) Passive sampling

- To generate data with PAS, an exposure time of 1 to 3-months is suitable for a monitoring network (McLagan et al. 2016, 2018).
- If sufficient resources and manpower are available, the sampling time can be increased (e.g., to monthly exposure) which will increase data points and be very helpful.
- Monthly sampling can be considered for more impacted sites and 3-months for remote locations.

(d) Wet deposition

- Weekly sampling is recommended to provide an integrated 24-hour 7-day sample as this is the preferred procedure currently in use within most wet deposition networks.
- Sampling should occur during the rainy season depending on the geographical location of the sampling site selected.
- It is beneficial to start with wet deposition monitoring at the beginning of each rainy season and stop sampling once the rainy season is completed (The duration will be different for each sampling site each season due to the influence of local conditions).

3.10. Quality assurance and quality control (QA/QC) for field air monitoring operations

Below are some examples of best practices, but more detailed procedures are found in the field SOPs available for major mercury air and wet deposition networks for which references are available in the supplementary material.

(a) Instrumentation

It is important to ensure that instructions for the correct installation and operation of instruments, samplers, and collectors are followed. Instruments must meet minimum requirements for sampling, such as sensor sensitivity to chemical inertness.

(b) Sample collection and handling

Specific quality control procedures that prevent contamination from occurring during sample collection and handling for the various monitoring techniques include:

- Wearing disposable plastic gloves whenever handling precipitation collectors, passive samplers and transferring samples from field sites.
- Avoid handling samples in areas where there may be high levels of Hg present.
- Properly transporting samples by “double bagging” samples after collection.
- Checking for, and documenting, sample leaks in the field, during shipping, and upon receipt at the laboratory.

(c) Field notes

- Field notes should be written down and kept safe and dry with the samples, with copies kept at separate location.
- Any deviations from the standard sampling method should be indicated and all supporting information documented.
- Meteorological and other environmental conditions that may affect the measurements should always be indicated by site operators.
- For passive, wet deposition and manual active sampling all the relevant information must be collected (sample times, including on and off times, dates etc).

(d) Sample storage and shipping

Proper storage and shipping methods must be used to preserve the chemical and physical integrity of samples. Quality control procedures for this purpose include:

- For wet deposition samples, maintaining samples (precipitation collected) in cooled containers while in transit and when stored in laboratory.
- Weighing wet deposition samples to determine sample volume at the station and at the laboratory in order to detect leaks in transit.
- Precipitation samplers should not be stored longer than 6 months before being analysed, even if they are properly stored and preserved.
- Passive samples should be stored in a cool dry place and in double bags and sealed tightly after being collected from the field.

(e) Blanks

Field blanks are to be collected on a regular basis to ensure that sampling methods and materials do not interfere with sample chemistry. It is recommended that blanks be collected randomly at every site. For manual active sampling monitoring methods, field blank test is performed regularly (for example, once every 10 times). For wet deposition monitoring the blanks are to be collected by pouring an aliquot of deionised water into a dry sample container (e.g., bucket, bag, funnel-and-bottle) for a sampling period during which no precipitation occurred. The aliquot should be submitted to the laboratory in the same manner as precipitation samples. For passive samplers, the field blanks need to be exposed together with the samples, but they are hermetically sealed with electrical tape to prevent exposure to atmospheric air while in the field. As an additional negative control, passive samplers can also be sent to the field and returned unopened so that they experience the same transport and handling that exposed samplers do.

(f) Uncertainties associated with atmospheric measurements

When conducting mercury speciation measurements, it is important to consider instrument setup to remove any bias between GOM and PBM measurements. Field and laboratory studies have shown that GOM concentrations performed on commercial instruments underestimate the GOM concentration. The collection efficiency of KCl-coated denuders have shown to vary with environmental conditions (O_3, relative humidity) and Hg^{II} compounds in air. Temperature and atmospheric composition influence PBM measurements, so networks performing PBM measurements maintain the particulate module at 50 °C to avoid the effects that a temperature drop could have on the GOM and PBM concentrations (GOM will deposit to the walls if this

happens). Therefore, it is better to interpret total reactive mercury observations rather than PBM and GOM data separately (Gustin et al. 2013, 2015, 2021).

3.11. Ancillary data

Ancillary data are collected to allow the (mercury) data to be understood in a valid manner; they are not indispensable for using the data but serve as additional information valuable for interpreting it. The most relevant ancillary data for mercury in air monitoring are: (on Tier 1) meteorological variables such as temperature, pressure, precipitation, relative humidity, wind direction and wind speed and (on Tier 2) chemical variables such as carbon monoxide (biomass burning), sulphur dioxide (volcanic activity),ozone (Arctic Mercury Depletion Events) or $PM_{2.5}$ (biomass burning) and (on Tier 3) halogens and other oxidants, which can be used to identify sources and atmospheric processes. For the collection of ancillary data, the following WMO GAW Guidelines will be useful (GAW Report 183 (WMO 2009), GAW Report 192 (WMO 2010), GAW Report 201 (WMO 2014), GAW Report 204 (WMO 2012)). The ancillary data should be collected at the same sites and stored with the same metadata and data format as the mercury data.

3.12. Management, analysis and evaluation of atmospheric mercury data

The following tools, provided as a non-exhaustive list of examples, aim to provide Parties and organisations with a more holistic picture of the state of mercury in air by adding value to the monitoring data that is collected. Combining atmospheric mercury data with a trajectory model will enable researchers to investigate mercury sources and sinks relevant to their regions. Additional examples are provided in the annex.

Local, regional and hemispheric trend analysis:

- Atmospheric concentrations and wet deposition data.
- Mann-Kendall non-seasonal trend analysis, preferentially pre-whitened, or machine learning methods (e.g., Empirical Wavelet Methods).
- For trend analysis, groups of stations for a region, latitudes or even a complete hemisphere.
- Depending on how the data is collected, the minimum number of years and minimum data coverage to calculate trends is recommended to be 5 years or more, with a minimum of 60% of days in each month with the values giving an idea of what is deemed sufficient when trends are reported in literature).

Observation-based source apportionment:

- e.g., PMF (Probability Mass Function) analysis of sources using other measured species indicative of major Hg sources (e.g., SO_2, CO).
- Based on statistical methods this approach uses the secondary measurements to identify different sources. e.g., high Hg and high SO_2 could be a volcanic or coal combustion source. High Hg and high CO could be a burning source.
- The PMF method could be used once a single year with data coverage of (>70%) is achieved as a means of source/sink appointment.

Source receptor relationship based on footprints and trajectory analysis:

- Analysis of source regions using backward modeling datasets generated by Lagrangian models like Hybrid Single-Particle Lagrangian Integrated Trajectory (HYSPLIT) or FLEXPART.

Backward trajectories or 3-dimensional footprints of air parcels released at the measurement location and followed up to 5-20 days backwards in time.

- By combination of measurements and information on air mass origin it is possible to determine source/sink regions based on a PSCF (Potential Source Contribution Function). One can also combine several stations for a comprehensive map. The next step would be a feasibility analysis to explore the potential for gaining more quantitative emission estimates by inverse algorithms.

3.13. Conclusions

This chapter has identified different methods for monitoring atmospheric mercury. The elements put forward in this chapter will provide Parties and organizations with the means to start, improve or expand their initiatives for monitoring atmospheric mercury and support evaluation of the effectiveness of the Minamata Convention.

Monitoring of atmospheric mercury has been ongoing for decades but not all regions are equally covered with the biggest data gaps in the southern hemisphere. A tiered approach is proposed and gives Parties and organizations an opportunity to start, expand or improve their monitoring programmes in such a manner that comparable data can be generated to support the Effectiveness Evaluation. Moreover, the tiered approach also breaks down the monitoring requirements in such a manner that new atmospheric mercury monitoring initiatives have an opportunity of joining one of the several existing monitoring programmes and networks, thus drawing from the experience and information at hand that these established networks can provide.

Automated atmospheric mercury measurement is the preferred method within existing monitoring networks. While the instruments used in automated measurements are capable of detecting very low concentrations of mercury, these instruments are expensive and alternative options are available that can also deliver comparable data. Manual and passive sampling of atmospheric mercury are two such options, even if at a lower temporal resolution as compared to automated systems.

Depending on the specific needs of the monitoring initiative, this guidance puts forward different methods at Tier 1, as the minimum step to start generating comparable atmospheric mercury data. Wet deposition of Hg from the atmosphere is one of the methods included at Tier 1 level. The method is reasonably well understood and sufficient results have been achieved in networks, as well as on a global scale, through various studies and intercomparison exercises. Therefore, scientifically sound, and cost-effective methods and techniques to determine mercury concentrations in air are available and can generate comparable data.

Another important factor to take into account when performing mercury air monitoring is the location(s) where monitoring will take place. Monitoring at a variety of sites will provide a more comprehensive picture of the levels of Hg in the atmosphere. It is therefore important for each monitoring initiative to identify sites that can provide insights into changes in atmospheric mercury levels over time, including relevant and sensitive ecosystems. Carefully selected sites can also help develop more robust atmospheric models and fill data gaps.

Beyond sampling and analysis, for any monitoring programme to be successful, a strong quality assurance and quality control (QA/QC) program is needed. A wealth of experience on key elements and processes related to QA/QC is available from existing atmospheric mercury monitoring programmes and networks as seen in this chapter and Supplementary Material.

Chapter 4. Biota Mercury Monitoring

4.1. Introduction

Mercury emitted to the air and released to water and land can be retained in the environment for years to millennia and may be transported across great distances, where its fate is complex as it moves through and across terrestrial and aquatic ecosystems (Driscoll et al. 2013; Kocman et al. 2017). Inorganic mercury from natural or anthropogenic sources becomes more toxic in the environment when it is converted to methylmercury (MeHg) by microbes.

Methylmercury readily biomagnifies through both aquatic and terrestrial ecosystems, resulting in increasing concentrations as it moves from the base of the food web to higher trophic levels (Eagles-Smith et al. 2016b; Figure 4.1). Generally, each trophic change in the food web accounts for roughly an order of magnitude (10x) of increase in MeHg concentrations, with the largest enrichment step occurring between water and plankton in aquatic systems (Lee et al. 2016). As a result, predatory animals, including invertebrates, fish, reptiles, birds, and mammals, may have MeHg concentrations in their tissues that are many orders of magnitude higher (often > 10^6 to 10^7-fold) than the concentrations found in abiotic matrices of the surrounding environment.

Studies show that the biomagnification (trophic level enrichment) and bioaccumulation (body burden accumulation over time) of MeHg adversely affect the behavior, physiology, survival, and reproductive success of many wildlife species, including fish, reptiles, birds and mammals, across many habitats and geographic areas of the world (Ackerman et al. 2016; Evers 2018; Dietz et al. 2019). Moreover, dietary uptake of methylmercury by humans, primarily through the consumption of fish, but also of marine mammals and birds, is a primary health concern (Trasande et al. 2016; Dietz et al. 2018; Fielding et al. 2021).

Monitoring of mercury in biota can inform policy making and implementation at various levels and across sectors. It can also help support the objectives of the Effectiveness Evaluation of the Minamata Convention guided by questions identified in chapter 2 for the achievement of monitoring objectives (see Table 2.1).

Existing biomonitoring programs provide valuable information for the Effectiveness Evaluation and the breadth of existing data will provide a basis for establishing comparable bioindicators, geographic areas of interest and a baseline for estimating change, as well as for identifying data gaps. However, there are some challenges in using existing data that were not necessarily designed for describing spatial patterns or standardized tracking of temporal trends or linking with anthropogenic mercury sources, even if they can be alleviated by introducing standardized tracking over time and by adding suitable ancillary measurements (e.g., environmental conditions, chemical properties, species attributes). The extent to which the existing information may be used in the Effectiveness Evaluation will depend on the ability to quantify the resulting uncertainty and a willingness to accept it in the use of existing biotic mercury concentrations as baselines – that may be an important determinant for establishing time series and developing spatial patterns. Continuous monitoring programs (either existing or new ones) that are designed to produce long time series with comparable methods and ancillary data will be particularly valuable in this effort. Furthermore, linkages between biotic Hg concentrations and anthropogenic sources and uses of mercury as identified in the Minamata Convention can be conducted with statistical analysis and modelling, given the availability of suitable ancillary data that can be linked to inputs, pressures, and drivers (Harris et al. 2007; Knightes et al. 2009; Dietz et al. 2019; Schartup et al. 2019).

Fostering international collaboration and coordination among national and regional projects will be crucial to create harmonized regional approaches and to strive, where possible, to integrate biomonitoring activities in an interdisciplinary manner (i.e., including air and human as well as biota monitoring) to assess ecological and human health risk that can be merged to illustrate regional and eventually global temporal trends and spatial patterns. For example, samples collected and analyzed using standard operating procedures across defined regions (i.e., regional hubs) may be collectively used for global analyses.

This chapter provides a brief overview of our state of knowledge with regards to existing data in fish and wildlife and proposes a strategy for using both existing and new biomonitoring data to support the Effectiveness Evaluation. Accounting for our understanding of the drivers and variability of mercury methylation across and within ecosystems, it offers scientific and technical considerations for the selection of monitoring sites, bioindicators, tissue type and ancillary measurements in a tiered approach to answer all the guiding questions in Table 2.1, including adverse effects to the environment and human health.

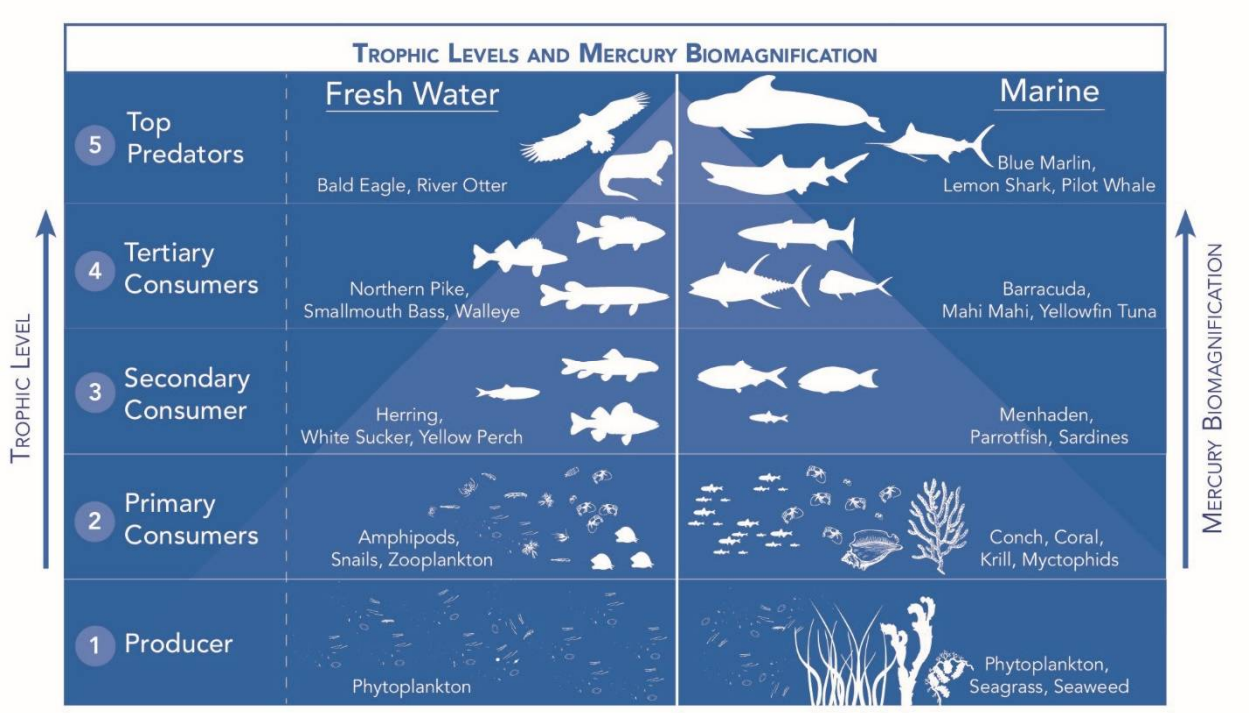

Figure 4.1. Examples of trophic level MeHg enrichment or "biomagnification" in freshwater and marine ecosystems.

4.2. State of knowledge

Mercury exposure has been well documented in fish and wildlife around the world. Published mercury concentration data for the target biota of the Minamata Convention exceed 530,000 data points and represent the world's oceans and continents (Evers and Sunderland 2019). Biotic mercury concentrations are most robust in fish, for both marine and freshwater ecosystems, and the number of analyzed samples are known to be much greater when including unpublished governmental and other datasets.

Numerous recent studies have documented adverse impacts across many fish species. In fish, adverse impacts of MeHg exposure include reproductive behavioural, and immunological impairment (Depew et al. 2012a; Scheuhammer et al. 2015; Carvan et al. 2017). Elevated

methylmercury concentrations may impact artisanal and commercial fisheries by reducing the viability and sustainability of fish populations, especially those in ecosystems with high sensitivity to Hg methylation (Evers et al. 2007) and at higher trophic levels – due to biomagnification. Therefore, such impacts can significantly reduce various ecosystem services[16] provided by fish, ranging from population sustainability and contributions to biological diversity, to food security and local availability of healthy food, to local livelihoods and commercial viability.

In birds, numerous studies document reduced reproductive success, behavioural change (e.g., reduced time incubating), and neurological problems (e.g., ataxia) (Depew et al. 2012a, b; Ackerman et al. 2016; Whitney and Cristol 2017; Evers 2018; Cristol and Evers 2020). Both avian piscivores and invertivores are at risk to MeHg availability in the environment because of trophic level enrichment. Mercury exposure varies greatly across habitats (Ackerman et al. 2016), continents (Evers and Sunderland 2019), and ocean basins (Albert et al. 2021).

In mammals, elevated MeHg concentrations can result in biochemical changes in the brain, ataxia, and reduced reproductive output (Dietz et al. 2013; Evers 2018; AMAP 2021). The effect thresholds for marine mammals are poorly understood, but based on mercury effect thresholds for terrestrial mammals, there could be significant adverse impacts on the reproductive success of marine mammals (Dietz et al. 2019), which likely generate further pressure on many species that are already threatened by other drivers of biodiversity loss (e.g., climate change, habitat loss, pollution, and overharvesting).

Existing mercury biomonitoring networks for biota that have ongoing and standardized measurements that can be used for objectives such as tracking temporal trends are relatively rare, with the exception of a few well-established regional initiatives, as documented by a review by UNEP (UNEP 2016). An overview of existing biota monitoring programmes, networks and databases is provided in the Supplementary Material.

4.3. Proposed elements for monitoring mercury in biota

4.3.1. Tiered approach to monitoring

To increase the comparability of biota monitoring data and improve our ability to generalize process-level knowledge a tiered approach based on a common site classification system is proposed, which takes into account: (i) the monitoring objectives and guiding questions in Table 2.1, (ii) the current scientific understanding of mercury's biogeochemical cycle, including its transport, transformation and bioaccumulation, as well as, atmospheric deposition, local pressures and large scale drivers that affect these processes, and (iii) the tiered approach presented in Chapter 2.4 to improve, expand and develop monitoring programs with available resources.

Given resource limitations, choices have to be made about what to monitor and where. A three-tier approach is recommended for Hg monitoring in fish and wildlife to answer the guiding questions in Table 2.1, starting with simple and low-cost arrangements, progressing towards more complex and resource intensive approaches. This tiered approach will ensure multiple lines of evidence that describe the effectiveness of the Convention in different ways. All the necessary elements of monitoring mercury in biota have been arranged into three tiers, including the selection of monitoring sites, bioindicators, tissue type, and ancillary measurements.

[16] Ecosystem services are the direct and indirect contributions of ecosystems to human well-being. These include provisioning services such as food and water; regulating services such as flood and disease control; cultural services such as spiritual, recreational, and cultural benefits; and supporting services such as nutrient cycling that maintain the conditions for life on Earth (Millennium Ecosystem Assessment 2005).

The tiered approach will enable: (i) estimation of background and impacted levels of mercury, particularly in regions where data gaps have been identified (see Supplementary Material Part A 2.5 for further information), and to establish new programs that meet statistical power of confidence to determine temporal trends and spatial patterns, and (ii) expand existing monitoring programs in support of trend analysis and source attribution to Hg source types. The tiered approach will also improve our understanding of (i) key environmental processes for mercury transport, methylation, and bioaccumulation, (ii) estimation of exposure pathways and adverse effects of mercury on target bioindicators, and (iii) increase our understanding of contributing factors that influence mercury transformation, biomagnification, and bioaccumulation in order to help normalize observed Hg concentrations and improve comparative analyses and models.

For mercury monitoring in biota, a low resolution-level set of parameters can be identified (Tier 1), including precise spatial coordinates of sampling site, tissue Hg concentrations, species information such as weight, length and age, and supplementary information on the lake and catchment (e.g., size, elevation, and land cover and use) or river (e.g., water level changes, current speed), or coastal area (e.g., coverage of area by mangroves, association of coral reefs). Medium resolution parameters (Tier 2) would include time series that would preferably be established for areas with well-known pollution loading (local catchment sources). Inclusion of monitoring sites with only external (long-range transported) pollution loads is crucial for effect evaluation in remote areas (e.g., boreal, subarctic, arctic). In addition to the main measurement parameters, ancillary data should also be included. The tiered approach would build upon the use of existing monitoring networks and stations. To best represent global patterns related to both local and long-range transport of mercury, additional monitoring sites could be needed to represent the degree of ecosystem sensitivity ("sensitivity" relates to variation in the environmental drivers of MeHg production which can be mapped with a certain level of value for decision-making) and have a mixture of background/reference sites together with sites with well-known local Hg sources. Bioindicators can be identified that are cost-effective and replicable over time. A proposed tiered approach for biota monitoring, which would be supported by existing and new programs, is shown in the Annex to this guidance document.

4.3.2. Monitoring sites

Ecosystems are variable in their relative sensitivity to mercury methylation. This is largely due to the heterogeneity of abiotic and biotic processes that influence the ability of any particular ecosystem to convert available inorganic mercury into its more bioavailable organic form (via the methylation process). Mercury transport, transformation, and biomagnification in the marine and continental environment is known or suspected to be influenced by several competing processes that ultimately determine mercury concentrations in a given individual. For example, sulphur and iron reducing bacteria are known to methylate mercury to MeHg, while light catalyzes the opposite reaction, demethylation. Concentrations of dissolved organic matter (DOM) and dissolved organic carbon (DOC) have, in turn, been shown to affect methylation and demethylation rates in both continental and marine ecosystems. Conversely, the biomagnification of mercury will depend on both bioavailability of MeHg and food web dynamics. While many of these processes are known, their relative strength and complexity is dependent on general location and specific monitoring site. This makes good site classification according to land use, habitat, and ecosystem characteristics important. It will improve comparability of observed mercury levels in biota and our understanding of how broadly the observed biogeochemical processes governing a particular location can be generalized with models. In addition, atmospheric deposition and potential vicinity to local pressures, like industrial, agricultural, or artisanal and small-scale gold mining (ASGM) activity can influence the system through direct and indirect mercury input by altering the biogeochemical cycle,

for example through increased solid and organic matter content caused by soil erosion. Finally, large scale drivers of Hg methylation, such as sea level and temperature rise, changes in ice cover, thawing permafrost, and deforestation can have a significant impact on mercury levels in fish and wildlife. Some of these site classification characteristics, pressures and large-scale drivers are described in more detail below in Table 4.1.

Table 4.1. Input, pressures, and drivers affecting mercury's biogeochemical cycle and its levels in biota, and mappable site characteristics that may affect the sensitivity of an ecosystem*.

Site characteristics	Description
Physical and (bio)chemical characteristics	
Land use/Land cover	Standardized land use and land cover classifications can provide a useful starting point to characterize potential sources, transport pathways, methylation mechanisms, exposure pathways, and relevant taxa. Areas with higher soil organic carbon (SOC) can have higher methylation potential and total mercury accumulation than areas with lower SOC levels.
Water body type and watershed morphology	Water body types (lakes, rivers, estuaries, bays, ocean) and watershed characteristics (catchment size, complexity of inputs and outputs) may be related to how mercury in biota respond to changes in mercury input. Ocean monitoring sites need to account for distance to shore and depth as important features.
Wetland type	Wetlands often have some of the highest rates of methylation of any land cover type. Multiple wetland types may be nested within land cover types, including shorelines of lakes, ponds, rivers; and swamps, bogs, and peatlands.
Ecosystem classification	Standardized classifications of ecosystem type may be useful to for comparison purposes across regions and the world.
Habitat and food web characteristics	Biogeochemical characteristics and food webs may vary over very fine spatial scales.
Atmospheric deposition	Wet and dry atmospheric mercury deposition are often the primary inputs to ecosystems that are remote from anthropogenic sources.
Vicinity to local pressures	
Artisanal and Small-Scale Gold Mining (ASGM) activity	On a global basis, ASGM is the single biggest activity that releases mercury to air, water, and land. Hg is emitted to air as it is evaporated from the amalgam. Hg is directly released to water from tailings (as water soluble complexes). Increased deforestation and soil erosion lead to increased mobilisation of soil bound Hg and particulate matter.
Industrial activity	Coal-fired electric power plants, non-ferrous metal production, and cement production are the largest industrial sources of mercury emissions and releases. Quantifying the amount of industrial activity within watersheds is an important indicator of local mercury emission and release potential.
Waste disposal sites	Waste disposal sites may have emissions or releases associated with the disposal of mercury containing products. Documentation of active

	and former waste disposal sites within watersheds is important for guiding monitoring of potential contamination.
Dams and water reservoirs	The creation of reservoirs creates a pulse of mercury through the release from soils, sediments and drowned vegetation and can last 1-2 decades. The management of water levels thereafter can further exaggerate the shoreline methylation process through frequent wet-dry cycles that can lead to increases in methylation exposure to aquatic and terrestrial biota.
Agricultural activity	Agricultural activities, such as flood irrigation, can increase Hg methylation. For example, rice paddy fields, which are a dominant agricultural land use throughout Asia, have been identified as important sites for MeHg production and a primary pathway of MeHg exposure to humans in mercury mining areas.
Soil erosion and soil leaching	Soil Erosion and soil leaching is the primary process that carries mercury from the land into freshwater ecosystems. There are many factors influencing soil erosion and soil leaching, and it is responsible for releasing mercury into the air and water, especially ground water. It is particularly pronounced where ASGM activity and deforestation occur but is not limited to these areas. Soil erosion and soil leaching are good proxy for habitat degradation, and an important indicator of the mercury transport process in terrestrial and freshwater environments
Fires	Fires are a natural disturbance process in many ecosystems. However, the frequency and intensity of fires has been influenced by climate change in many ecosystems, including forests and wetlands in the tropics. Fires result in the natural release of mercury into the air, and the more fires there are, the more mercury is likely emitted.
Fisheries and Aquaculture	The proximity of fisheries (or fishing grounds/areas) and aquaculture is an indicator of potential human exposure.
Proximity to large scale drivers of mercury release and/or methylation	
Thawing permafrost	Areas with permanently frozen soil are one of the largest immobilized mercury reservoirs globally. Melting of permafrost therefore promotes direct emissions of mercury into air and mobilization into river systems, that transport it further to the marine environments. Thawing permafrost also leads to the creation of wetlands and anoxic methylation hotspots.
Deforestation	Deforestation is one of the most important process driving mercury releases to the water in tropical forest regions. The amount of deforestation in a watershed is an important indicator of disturbance and potential mercury releases.
Sea level rise and warming	Sea-level rise, ocean warming, and ocean acidification may impact methylation rates and MeHg availability.

*This table is not all-inclusive but is meant to provide for identifying variables of greatest interest for understanding the sensitivity of ecosystems to Hg methylation.

Even in areas where Hg deposition is low, concentration in fish and wildlife may be disproportionately high if conditions are conducive to MeHg production and biomagnification (i.e., the amount of total mercury in any given location does not necessarily correlate to adverse impact). For example, ecosystems that are highly sensitive to mercury methylation may require only limited amounts of inorganic mercury to pose risks to organisms. Similarly, ecosystems with little to no sensitivity to mercury methylation may experience high levels of mercury inputs with limited impacts to the environment and human health. To credibly assess the potential threat of mercury to biota, biodiversity, ecosystem services and people, it is important to collectively assess Hg sources (from potential multiple local and remote inputs) and ecosystem sensitivity (ability to convert available mercury into its more toxic, bioavailable form). This makes it important to choose monitoring sites and ancillary measurements according to the monitoring objective(s) of interest. The selection of monitoring sites must account for the broad geographic range of methylation abilities in oceanic and continental areas. The response from one site is not necessarily relatable to the response of a neighboring site that has different habitat characteristics. Once monitoring sites are chosen, tracking temporal trends will be possible by performing consistent sampling over multiple years.

Understanding this variability (with a particular interest for highly sensitive areas) is important during the process of identifying monitoring sites especially when addressing the guiding questions (from Table 2.1) to characterize spatial patterns and estimate exposure and adverse impacts. Conversely, selecting monitoring sites that are less sensitive may be important for tracking temporal trends to reduce confounding variables. Therefore, a mix of monitoring sites that represent both sensitive and less-sensitive ecosystems can address multiple questions is viewed as most useful and identifying the key variables that may provide some direction for selection are important.

The selection of monitoring sites is dependent on what monitoring and analysis questions are of most interest. The recommendations presented here have been organized into three tiers with the goal of informing the Effectiveness Evaluation by addressing each of the guiding questions.

The recommended approach for Tier 1, for Parties and organizations with limited resources, is to focus on a mixture of locations that are (a) remote from anthropogenic sources and expected to represent background conditions and (b) impacted by known anthropogenic sources. These sites should be visited annually. The potential to link biota data from sites where air, deposition, and human biomonitoring activities are taking place should also be considered in selecting Tier 1 sites. Selecting sites that may be expected to be simpler to understand may be better as a starting point than selecting very complex sites. Where little or no prior information exists, experience from the development of Mercury Initial Assessments suggests that mapping ecosystem and landscape characteristics overlaid with sources of mercury emissions and releases can be a helpful tool to inform the selection of monitoring locations. Such mapping may consider landscape or ecosystem characteristics, local pressures, and large-scale drivers, examples of which are outlined in Table 4.1. Based on recent evidence of elevated mercury loading into coastal oceans from rivers and the importance of river deltas for artisanal fisheries and biodiversity, such ecosystems may be emphasized for mercury biomonitoring efforts to measure regional trends and potential health impacts (Amos et al. 2014; Liu et al. 2021).

Progressing to Tier 2, more sites should be added to cover a wider range of landscapes and geochemical characteristics and local and large-scale pressures. Tier 2 sites may be visited less frequently (every 2-5 years) on a rotating basis, allowing a greater diversity of sites to be sampled. As observations are collected, information on mercury levels and ancillary parameters can be incorporated into the mapping exercise described above, helping to guide future site selection.

Tier 3 sites, which are focused on understanding the underlying processes controlling the presence and movement of mercury and mercury compounds in the environment, should be selected to characterize a specific watershed or coastal area of interest. This may be accomplished by having a primary location for a suite of detailed measurements (supersite) and an array of secondary sites (satellites) for limited measurements to capture the variability across the watershed or coastal area. Tier 3 might be designed to collect quantitative information that would allow for weights to be assigned in future iterations of the mapping exercise discussed above with the help of models that have been validated for the specific site characteristics.

To evaluate the effectiveness of the Minamata Convention and answer all the monitoring objectives and guiding questions in Table 2.1 will require combining a mix of discrete categorical and continuous data. Methods are being developed that ensure both consistency and transparency in this approach, as well as the ability to down-scale this approach for application at regional and local-levels to make use of critical information not available at a global scale (e.g., point-source data). As water is a major pathway for mercury through ecosystems, evaluating the threat of mercury via watersheds up to coastal areas has emerged as an important part of monitoring as a justifiable, hierarchical approach to assessment across many spatial scales (Evers and Sunderland 2019).

Creating new models of risk, sensitivity, and threat of mercury impacts to the environment and people and emphasizing land-sea connectivity of watershed to coastal areas, will significantly improve the selection of priority sites for global mercury biomonitoring that will most effectively use limited resources. Information from these biomonitoring priority areas can, in turn, be used to adaptively manage and improve the usefulness of mercury threat-related assessments over time. This supports the application of “systems-thinking” considered necessary to chemicals and waste problem-solving in which a set of synergistical analytical skills is used to improve the capability of identifying and understanding systems, predicting their behaviors, and devising modifications to them to produce desired effects (Arnold and Wade 2015).

4.3.3. Selection of bioindicators

Mercury monitoring using biotic media requires the careful selection of aquatic and terrestrial bioindicators and associated tissues that can realistically respond to key objectives of identifying temporal trends, spatial patterns, and linking with mercury source types. The selection of bioindicators will vary according to monitoring activities and associated guiding questions (see Table 2.1), geography, habitat, and ultimately to national interests.

For Tier 1 monitoring, high trophic level biota can be effective bioindicators because they have a strong nexus to Hg concentrations that may be of concern to ecological and human health. A previous bias toward sampling higher trophic level species has generated extensive mercury exposure data in the published literature (Evers and Sunderland 2019). Based on the knowledge of existing biotic Hg data and only using comparable data (e.g., trophic level 3 or 4 species that can be regularly sampled for comparable purposes for understanding spatiotemporal patterns) for relevant terrestrial biomes and associated marine areas, a matrix of available data can respond to questions related to spatial patterns, temporal trends, and linkages with mercury source types. While monitoring mercury in trophic level 3 and 4 biota can be useful for assessing potential mercury exposure for humans and top predators, which play important roles in maintaining ecosystem health and high levels of biodiversity (Sergio et al. 2008), attribution of causative relationships for observed temporal trends is more complex. Rich supporting information on the processes that affect bioaccumulation and MeHg availability is therefore needed to conduct robust trend analysis and separate anthropogenic influences from each other. Such information is readily available from

several long-term mercury monitoring programs in the Northern Hemisphere, and they have already been used for trend assessment on a pan-regional scale (AMAP 2021).

An important step towards developing comparable biotic monitoring data to inform the Effectiveness Evaluation is to define regional bioindicators for monitoring in order to minimize the effects of species-specific physiological differences. For example, there are several game fish species that are found in northern Europe and North America that accumulate significant amounts of Hg due to their high trophic level and are frequently used by Hg biomonitoring programs (Depew et al. 2013; Eagles-Smith et al. 2016a; Olk et al. 2016). To be able to potentially explain the main drivers behind the spatial patterns and temporal trends of fish Hg concentrations, and how these patterns and trends change under influence of different and emerging drivers (including environmental / climate change and deposition change in addition to changes in emissions and releases), a set of minimum target information could be developed. While adult predatory fish, and piscivores birds and mammals may be useful for characterizing spatial gradients and estimating exposure and adverse impacts on human health and the environment, younger fish are better for tracking temporal trends in availability of MeHg at the base of the food web, to avoid confounding effects related to varying bioaccumulation rates over times and shifts in diet (Vander Zanden et al. 1999; Wiener et al. 2012a).

When establishing new programs primarily intended for analyzing trends, detecting short-term (<10 years) trends in changes of mercury in biota are best viewed through young individuals where age and, therefore, bioaccumulation (important for high risk, long- lived species that can increase Hg body burdens over their lifespan) is not such a significant confounding factor. For fish, selecting individuals <2 years of age is suitable (Wiener et al. 2012b). When using long-lived birds as bioindicators both short-term (i.e., blood) and longer term (e.g., feathers) temporal objectives can be met simultaneously with individual sampling, which have been useful for mercury monitoring programs in Canada and the United States over the past three decades (e.g., loon species; Evers et al. 2008b, 2014; Scheuhammer et al. 2016; Yang et al. 2020).

A key initial step in bioindicator selection is to decide on the temporal and spatial integration that is to be represented and whether the aim of biota monitoring is linked to environmental impact and/or human exposure assessments. It is often possible to select organisms that provide monitoring data for both purposes. Careful selection of bioindicators could further provide information about the potential impacts of MeHg contamination on biodiversity, including threatened species (e.g., IUCN Red List of Threatened Species™),[11] keystone species, as well as other species of national and global interests for their conservation and protection in so far as non-lethal and no-impact sampling methods can be used. Moreover, monitoring of such species in tier 3 will also provide insights on the relationship between mercury and additional stressors such as habitat degradation, climate change, and overharvesting. For example, ASGM activities in tropical systems significantly contribute to environmental MeHg loads as well as severely altering habitat quality in areas with high endemism (Gearson et al. 2021) – such locations are therefore of interest for biomonitoring. Ultimately, careful selection of bioindicators can include taxa supporting objectives of other multilateral environmental agreements (e.g., Convention on Biological Diversity).

Minamata Convention Article 19 (b) states that the "modelling and geographically representative monitoring of levels of mercury and mercury compounds in vulnerable populations and in environmental media, including biotic media such as fish, marine mammals, sea turtles and birds, as well as collaboration in the collection and exchange of relevant and appropriate samples". The extensive data on Hg in biota found in the published literature, can inform the selection of bioindicators for monitoring. Informed selection can ensure cost-effective comparability at regional and global scales. Table 4.2 lists a number of species and species' groups that are well described and may serve as useful bioindicators for ecosystem health and human exposure assessment,

categorized within their respective biomes and associated aquatic ecosystems (Evers et al. 2016). Appropriate tissue types for varying objectives are shown in a tiered approach in Annex 1. Biomonitoring for tracking temporal trends should be consistent with species, tissue, and location sampled, sampling methodologies, and analytical approaches. and migration, if applicable, is critical for interpreting Hg concentrations with high certainty.

Table 4.2. Examples of trophic levels 3 and 4 biota that could serve as bioindicators grouped by major biomes and associated nearshore areas (based on Evers et al. 2016).*

Terrestrial biomes and associated marine areas	Bioindicators for assessment of potential environmental impact			Bioindicators for assessment of potential human exposure[17] (which can also be used for assessing environmental impacts)		
	Freshwater Birds	Marine Birds	Marine Mammals and other	Freshwater Fish	Marine Fish	Marine Mammals
Arctic Tundra and Arctic Ocean	Loons, Songbirds	Fulmars, Murres	Polar Bears, Seals	Arctic Char, Arctic Grayling	Arctic Char, Cod, Halibut	Beluga, Narwhal, Seals
Boreal Forest-Taiga and North Pacific and Atlantic Ocean	Loons, Eagles, Osprey, Songbirds (invertivores only)**	Osprey, Petrels	Otter, Seals	Perch, Pike, Walleye	Bluefish, Cod, Tuna	Pilot Whale
Temperate Mixed Forest and Pacific and Atlantic Ocean	Loons, Egrets, Herons, Eagles, Osprey, Terns, Songbirds (invertivores only)	Osprey, Terns	Otter, Seals	Perch, Bass, Walleye	Barracuda, Mackerel, Mahi mahi, Sharks, Tuna	Pilot Whale
Tropical Rainforest and South Pacific and Atlantic and Indian Ocean	Egrets, Herons, Kingfishers, Songbirds (invertivores only)	Albatrosses, Frigatebirds, Shearwaters, Terns, Tropicbirds	Dolphins, Otter, Seals and Sea Turtles	Catfish, Cichlids, Snook	Barracuda, Grouper, Mahi mahi, Sharks, Swordfish, Tuna	Pilot Whale

***** Trophic level 3 or 4 young individual fish (<2 years) can be used for tracking temporal trends (see below). ** Songbirds foraging within invertivore foodwebs are at trophic level 3 or higher (Cristol and Evers 2020).

Fish size normalization is an additional tool for comparative purposes. Fish Hg concentrations can be standardized for size by converting to standard units to account for variation related to length and, as a proxy, age using a general linear mixed model[18]. Trophic level 3 and 4 fish may still be used for

[17] Rice grown near mercury contaminated sites can also be a significant source of human exposure through dietary intake. However, this chapter is focused on animals as bioindicators of environmental and human exposure to mercury.

[18] Using the natural log transformed THg concentration as the response while the total length is the fixed effect with a random effect of length by genus allows for the calculation of the genus-specific effect of length on muscle THg concentrations (conditional $R^2 = 0.58$). The residuals of this model can then be added to the predicted value to obtain a length-standardized Hg concentration for each individual sample.

this objective, but younger rather than older individuals could be sampled. The use of trophic level 3 and 4 fish, both young (for temporal trends) and adults (for spatial patterns) can simplify sampling efforts and preferably multiple age groups, size classes, and trophic levels should be sampled to best understand Hg concentrations in the fish community.

While Table 4.2 focuses on vertebrates, there are some predatory aquatic invertebrate taxa that occupy trophic level 3 and are also effective bioindicators of food web Hg exposure. Dragonfly larvae (Order: Odonata, suborder: Anisoptera) are an example of one such aquatic invertebrate group that makes an effective bioindicator. As such, some countries (e.g., United States) have implemented long-term national scale programs using dragonfly larvae as bioindicators of the risk posed by mercury in public lands (Eagles-Smith et al. 2020).

Final selections of target biota for monitoring Hg and its impacts on the environment and human health should be evaluated for their life history characteristics, as well as their plasticity in foraging ecology and habitat, spatial use and movement/migration patterns, variability in growth rates, temperature, and general water quality tolerances, geographic distribution, and socioeconomic interests for humans.

4.3.4. Tissue types

The selection of tissue types will vary according to monitoring objectives and associated guiding questions (see Table 2.1), geography, taxa being monitored, and ultimately to national interests. Examples of the proper selection of tissue type are well-established with associated information about the percent MeHg content in the tissue and the preferred type of tissue preparation (Table 4.2). Most muscle, blood, egg and keratin-based (e.g., scutes, feathers, and fur) tissues primarily contain MeHg and can be sampled through non-lethal methods (with some exceptions such as whole-body analyses of small fish). This is important for the simpler and more cost-effective laboratory analyses of total mercury concentrations that can assume 95% or more MeHg content. It is also important when sampling species that are threatened, sacred and/or legally protected. Field protocols are available for all tissue types (see Supplementary Material).

Table 4.3. Major biota groupings and tissues recommended for MeHg monitoring.*

Biota Group	Tissue Type	% MeHg	Sample preparation type**	Analysis type	Source reference for % MeHg	Comments
Fish	Muscle fillet	75-95% (but varies on average as low as 65%)	ww or dw	THg	Bloom 1992; Lescord et al. 2018	Recent evidence indicates that %MeHg may be lower for some fish species (Manceau et al. 2021) and for some cooking approaches (Wang et al. 2013) so to confirm the expected amounts. 10% of fish should be analysed for MeHg content.
	Muscle biopsy	75-95% (but varies)	dw	THg	Peterson et al. 2004	dw is best because of moisture loss concerns. Muscle biopsy to muscle fillet has a r2 = 0.96. Biopsy plug depth may impact Hg measured – 5mm plugs are best below dorsal fin (Cizdziel et al. 2002) and are without skin and adipose tissue.
	Fin clips, muscle fillet and whole body	varies	dw	THg or MeHg	Cerveny et al. 2016	There is a significant correlation between fin clips and muscle fillet/whole body ($p<0.01$).
	Blood	>95%	ww or dw	THg		Assumed to be >95% MeHg based on other vertebrates.
Sea Turtles	Scutes / Carapace fragments	~10%	fw (or dw if scutes need washing)	THg	Rodriguez et al. 2019	While scutes are keratinized material the %MeHg may be relatively low and needs more data.
	Blood	>95%?	ww or dw	THg		Assumed to be >95% MeHg based on other vertebrates.
	Muscle	>95%?	ww or dw	THg		Assumed to be >95% MeHg based on other vertebrates.
Birds	Blood	>95%	ww or dw	THg	Rimmer et al. 2005; Edmonds et al. 2010	Elimination of MeHg in blood comprises an initial fast phase, with half-time of 1 day, and a slow terminal phase with half- time between 44-65 days. Molt is a crucial factor in determining the rate of MeHg elimination (Monteiro and Furness 2001).

	Feather	~100%	dw or fwwhole feathers	THg	Burger 1993	Use feathers with caution; see Peterson et al. (2019) for a tool and guidelines for feather processing, analysis, and Hg interpretation.
	Eggs	>96%	dw or ww	THg	Ackerman et al. 2013 (96% for 22 species)	ww and dw can be problematic if eggs are not collected immediately after laying (Dolgova et al. 2018).
	Eggshells and membranes	>95%	dw	THg	Peterson et al. 2017	Membranes are assumed to be primarily MeHg, but shells are entirely inorganic Hg.
	Muscle	>95%	ww or dw	THg		MeHg comprised over 99% of total Hg in breast muscle of waterfowl(Sullivan and Kopec 2018)
Mammals	Skin	>90%	dw	THg	Wageman et al. 1998	Muktuk (in marine mammals) includes layers of skin and blubber
	Fur or hair	>90%	dw (or fw if fur is not washed)	THg	Evans et al. 2000	Use fur with caution; fur/hair may not relate to blood and muscle depending on growth patterns (Peterson et al. 2016)
	Muscle	>90%	ww or dw	THg	Wageman et al. 1998	

* Except for the whole-body samples, all tissues can be non-lethally sampled.**Reported as wet weight (ww), dry weight (dw) or fresh weight (fw). Fw denotes keratin-based samples that are not cleaned or dried prior to total Hg analyses.

4.3.5. Ancillary measurements

As described above, the conditions and processes that drive levels of mercury in biota vary geographically. Certain ecosystem conditions (e.g., acidic wetlands with fluctuating water levels) can encourage the production and bioavailability of MeHg in the environment. Bacteria often produce more MeHg under low oxygen (hypoxic or anoxic) conditions (Hsu-Kim et al. 2013, 2018). Light and microbes are known to promote the opposite reaction, de-methylation (Poste et al. 2015; Klapstein et al. 2016; Kronberg et al. 2018; Eckley et al. 2021). Environmental factors such as pH, dissolved organic carbon, total suspended solids, and sulphur concentrations are important in influencing inorganic Hg input, transport, and net methylation rates (Wyn et al. 2009; Gabriel et al. 2014; Gorski et al. 2008; Chételat et al. 2018; Rudd et al. 2018; Broadley et al. 2019; Taylor et al. 2019; Braaten et al. 2018). Ancillary measurements must therefore be taken to understand the relative strength of these processes, to improve comparability between sites, and to normalize trends and to perform source attribution.

The complex chemical conversions and cycling of Hg make it challenging, but not impossible with suitable ancillary measurements that can be used to parametrize models, to predict the concentration of MeHg in fish and wildlife from concentrations of inorganic mercury in air, water, and sediments (Chen et al. 2014; Gustin et al. 2016; Sunderland et al. 2016; Eagles-Smith et al. 2018). At more intensively monitored Tier 2 sites relationships between abiotic inputs and MeHg in fish and wildlife may be established. Inclusion of the abiotic ancillary measurements will also

become increasingly important as the historical connection between biotic Hg and Hg concentrations in air become more complex and the relative contribution of re-mobilized legacy Hg increases (Wang et al. 2019). Abiotic ancillary measurements will also often react faster to changes in inputs, pressures, and drives that could have been influenced by the Convention (Valdes et al. 2017; Bierregaard et al. 2020).

Ancillary measurements also help to normalize observed Hg concentrations with respect to known co-variates that subsequently facilitate interpretation of temporal trends, spatial gradients, health and environmental impacts, and source attribution. Changes in trends of ancillary measurements may represent different pressures and drivers (see Table 4.1.). Moreover, by establishing quantitative relationships between drivers/predictors and associated responses of Hg concentrations in biota that will subsequently improve models and risk assessments. While some ancillary measurements need to be measured on site, others can be observed with available datasets.

Table 4.4. Examples of matrices by tier level (low to medium to high resolution for characterizing a monitoring site) for sampling and analyzing biota (and for tier 2 and 3, abiotic matrices) in conjunction with ancillary measurements.

Tier	Matrix	Ancillary measurement examples
1 (low resolution)	Aquatic Biota (e.g., fish using muscle samples; birds using feathers)	Species information, body length, body mass, spatial coordinates, sex
2 (medium resolution)	Aquatic Biota (e.g., fish using muscle samples; birds using blood and feathers) Sea and freshwater Surface sediment Air	Species information, body length, body mass, spatial coordinates, sex, carbon ($\delta^{13}C$), nitrogen ($\delta^{15}N$), MeHg From surrounding abiotic media: Water: DOM/DOC/TOC, TSS, salinity, DO, (pH), N and P, phytopigments (e.g., chlorophyl *a*); Surface sediment (top 2 cm): THg Air: GEM, wet deposition, and meteorological data
3 (higher resolution)	Aquatic biota (e.g., fish using muscle and whole body samples; birds using eggs, blood and/or feathers; marine mammals using muscle samples) Terrestrial biota (e.g., sea turtles using eggs and/or scutes; birds using eggs, blood and/or feathers) Sea and freshwater Surface soil and sediment Air	In biota: Species information, body length, mass, spatial coordinates, carbon ($\delta^{13}C$), nitrogen ($\delta^{15}N$, speciated mercury ($\delta^{202}Hg$) stable isotopes, mercury ($\delta^{199}Hg$) stable isotopes in biota and suspected source-matrices of interest; other chemical tracers related to known drivers (i.e., changes in CO_2 levels and water temperature in oceans due to climate change, co-tracers from ASGM activity, etc.). Information on diet (e.g., fatty acids), stable isotopes of lower foodweb organisms (or compound specific stable isotopes of amino acids in fish), data on foodweb structure, as well as associated land cover data. From surrounding abiotic media: Water: DOM/DOC/TOC, TSS, salinity, DO, (pH), N and P, temperature, depth, phytopigments (e.g., chlorophyl *a*); Surface soil or sediment (top 2 cm): THg Air: GEM, wet deposition, and meteorological data

Ancillary measurements often collected with biota mercury data include species identification, length, weight, and spatial coordinates (low resolution level) and additionally fat levels, and stable isotopes of carbon ($\delta^{13}C$) and nitrogen ($\delta^{15}N$) and other variables can be included (higher resolution tier levels – e.g., Tier 2 and 3) (Table 4.4). Stable isotope measurements in biota assist with identifying changes in food web structure, trophic position and feeding habitat (Abeysinghe et al. 2017) and aid in evaluating causes of temporal trends in the context of abiotic factors such as changing air emissions, sediment and water chemistry, and temperature. Without these ancillary

measurements, and analyses that normalize data in the context of food web dynamics, it will be challenging to determine if the observed changes, or lack thereof, is due to changes related to the implementation of the Minamata Convention or driven by large-scale factors such as changes in food web complexity, trophic position of biota, climate change, overharvesting, and biogeochemical conditions. AMAP's mercury monitoring programs for biota include these ancillary measurements and existing surveillance efforts conducted by Canada and Norway also sometimes include stable isotope measurements of carbon ($\delta^{13}C$) and nitrogen ($\delta^{15}N$) (Braune et al. 2015, 2016).

Body length, mass, species name, and spatial coordinates (latitude/longitude) are nearly always collected as metadata in mercury monitoring programs for biota. However, some studies also collect data from other matrices including seawater and marine sediments (Azad et al. 2019) along with high resolution ancillary variables including but not limited to carbon ($\delta^{13}C$), nitrogen ($\delta^{15}N$), and mercury ($\delta^{202}Hg$ and $\delta^{199}Hg$) stable isotopes (Cransveld et al. 2017), pH, salinity, sea depth, organic carbon, dissolved oxygen, temperature, etc.

To evaluate the effectiveness of the Minamata Convention, there is also a need to quantify the relative contribution from different abiotic sources (like legacy Hg and atmospheric deposition) by supplementing biotic Hg samples with air, sediment, and water samples (Braaten et al. 2019). This is often a cost-effective way to improve the explanatory power of biota measurements in established monitoring programs (Mason et al. 2005). Measurement of mercury levels and fluxes from abiotic media, such as water and sediment/soils, should be included in Tier 2 and 3 biomonitoring to help quantify legacy sources and provide further support in understanding the drivers of temporal trends and spatial patterns of Hg in fish and wildlife. Abiotic media should not be used exclusively because of interpretative limitations and uncertain connectivity with associated biota, especially high trophic level species. This makes it important to design integrated biota monitoring programs that also include abiotic ancillary measurements.

For each terrestrial location, this should include lake and catchment morphology, riverine variables, pollution deposition patterns, and local pollution history. For each animal species data must include length, weight, sex, and age (when it can be obtained). Samples (i.e., fish muscle) for determination of total Hg concentrations, should also be analysed for stable isotopes of nitrogen ($\delta^{15}N$) and carbon ($\delta^{13}C$) for a better understanding of trophic position and energy sources. Lastly, Hg isotopes are included in Tier 3 as a relatively new tool for attributing biotic Hg body burdens to Hg source types and understanding sources of available MeHg (Lepak et al. 2018; Schudel et al. 2018; Renedo et al. 2020). Recommendations for their use with the Minamata Convention are to collect measurements of particulate-bound mercury in the atmosphere and sediment mercury isotope ratios near mercury hotspots and in fish to help evaluate effectiveness (Kwon et al. 2020).

4.4. Field sampling, laboratory analysis and data management

The timing of biota sampling at monitoring locations varies according to the objectives of the monitoring, habitats/ecosystems, and chosen bioindicators. The fraction of mercury retention in the atmosphere, soils, and waters can vary over days to centuries (Figure 4.2). Therefore, knowledge of mercury retention in habitats that biota is sampled from will be important for understanding temporal trends, spatial patterns, and linkages with mercury sources. Sample timing also depends on the rate of change in Hg concentrations in the bioindicator tissues of choice.

Information on climate variables, habitat type, and taxa ecology are generally needed for proper interpretation of temporal trends and spatial patterns. For linkages to mercury source types and understanding foodweb complexities, mercury isotopes are important (Kwon et al. 2020). To

understand mercury exposure and the potential effects on taxa, it is important to know the age category, morphometrics (e.g., weight, length, etc.), and foraging ecology.

In some cases, where total mercury body burden changes rapidly, such as in fish and birds within lakes with small watersheds (Evers et al. 2007), changes can be detected on the scale of years (Wiener et al. 2012). Biomonitoring in areas with smaller changes in environmental loadings, but with more complex ecosystems that contain varying processes that sequester, and methylate mercury require sampling annually for one or two decades (Riget et al. 2011; Eagles-Smith et al. 2016a; Sunderland et al. 2018; Evers et al. 2020). Within ocean basins, increasing mercury concentrations were detected over multiple decades for tuna in the Pacific Ocean (Drevnik et al. 2015; Drevnik and Brooks 2017).

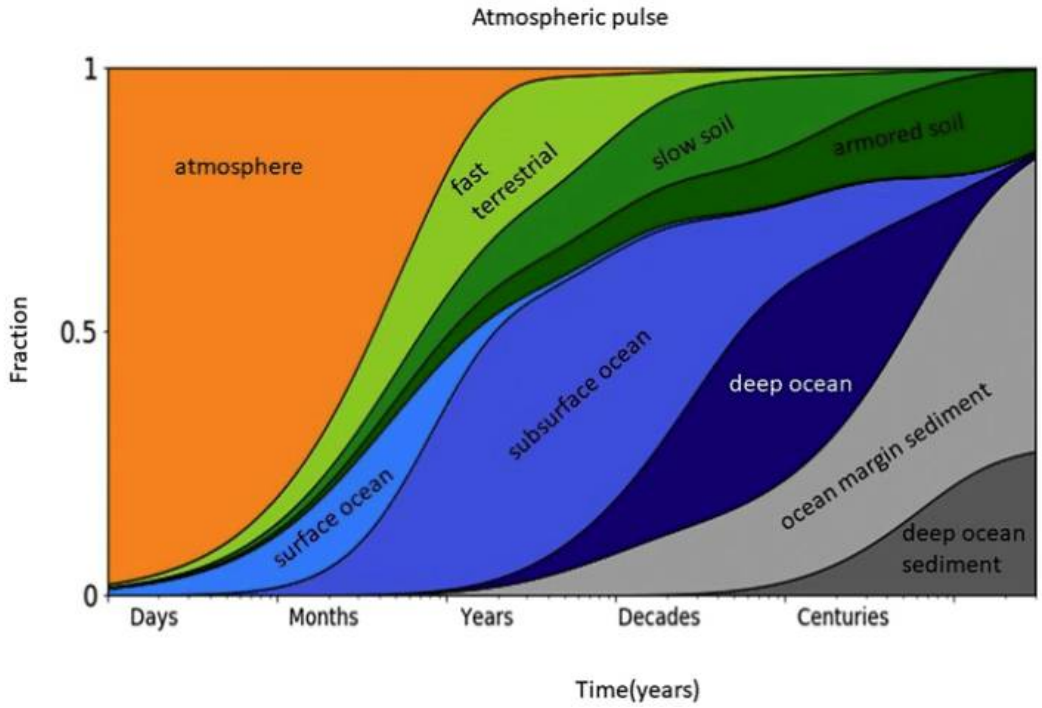

Figure 4.2. Retention of mercury fraction (0 to 100%) over time (days to centuries) in various compartments of the atmosphere, landscape (e.g., soils), and waterscape (e.g., ocean waters and sediments) (Amos et al. 2014).

4.5. Quality assurance and quality control for biota monitoring

The quality control and quality assurance of mercury concentrations analyzed from various types of animal tissue are important and require proper standard reference materials and the use of duplicate sample analyses and blanks.

While instrument calibration is important for obtaining accurate and comparable mercury data from biota, especially when comparing different types of instruments (e.g., DMA vs. CVAA), sample handling and processing are by far the greatest sources of introduced variability in observed levels of mercury in animal tissues. Therefore, protocols for sample collection, handling, shipping, and preparation will need to be carefully vetted, described, and followed.

Furthermore, similar to other matrices, analytical operations for biota monitoring will also need to follow strict chain-of-custody and standard operating procedures for laboratory analysis and data

handling. Further information on quality assurance and quality control is available as Supplementary Material to this guidance.

4.6. Statistical considerations

A wide array of statistical tests is available to evaluate temporal trends of mercury levels in biota including, but not limited to, generalized linear (GLM) and non-linear (e.g., logistic regression) models, classification and regression tree (CART), Mann-Kendall (MK) test, and Bayesian model selection and uncertainty assessment techniques including the widely used Akaike information criterion (AIC), etc. For evaluating spatial trends GLM, general additive modelling (GAM), kriging or Gaussian process regression, Cox point process and spatial covariance modelling, principal component analyses (PCA), multiple-response permutation procedures (MRPP), probability density estimations and Monte Carlo simulations are some of the approaches that can be used for existing or new biota data sets. While statistical tests may inform optimal sample sizes, power analyses combined with probability interests and variability of mercury concentration in different tissues are a more suitable basis for choosing the type of sample to be collected (see above).

Length or body mass normalization of biota will be critical for interpreting mercury and methylmercury data. Moreover, evaluating trophic position and food web structure using carbon ($\delta^{13}C$), nitrogen ($\delta^{15}N$) in conjunction with mercury ($\delta^{202}Hg$) stable isotopes and other matrices such as seawater and sediment can support more rigorous high-resolution modelling although targeted sampling may be required to achieve this goal.

As with any analysis, when dealing with biota monitoring in the context of the Effectiveness Evaluation, there are important limitations and uncertainties that need to be conveyed in a clear and transparent way. Specifically, there are major challenges linking mercury levels in biota with mercury concentration in abiotic matrices such as air and water especially considering post depositional processes, trophic position, changes in food web structure and complexity, and broad-scale drivers such as environmental chemistry factors (e.g., pH, DOC), temperature, geography, species growth rates, and climate change (Braune et al. 2015, 2016).

4.7. Conclusions

Biota monitoring data can help address the monitoring objectives and guiding questions (Table 2.1) in support of the Effectiveness Evaluation of the Convention. Historic data available from various biota monitoring programmes, databases and other resources can be used to improve our understanding of the exposures to mercury in biota before the Minamata Convention's entry into force and to help establish a baseline for the Effectiveness Evaluation. Moving forward, existing government-led national mercury monitoring programs, regional initiatives, and/or academic-led studies can provide comparable biota monitoring data for use in the Effectiveness Evaluation. New monitoring efforts may further contribute by providing comparable data on key bioindicators filling data gaps and building capacity. Biota monitoring data and associated ancillary measurements can be collected in continental and marine ecosystems designed as part of a Tiered approach for Parties and organizations who elect to develop new monitoring programs or improve existing ones.

Briefly, Tier 1 is suitable for Parties or organizations seeking to create a biota-based monitoring program, or expand a minimal program, but that may not have sufficient resources to implement the actions in Tier 2. The goal of monitoring activities under Tier 1 would be to identify temporal trends and collect total mercury measurements from trophic level 3 or 4 biota that best represent the targeted habitats. Tier 1 activities should ideally be repeated for the same species using the same size classes in the same habitat every year. Tier 2 aims to collect ancillary measurements that will contribute more meaningfully to all six monitoring objectives in support of the Effectiveness

Evaluation, improve the ability to interpret biotic Hg measurements by collecting additional ancillary measurements, and thus calls for more in-depth analysis of the Tier 1 monitoring efforts, or incorporation of mercury biomonitoring into other, in-depth mercury monitoring efforts. Tier 3 aims to increase understanding of key processes that influence the presence and movement of mercury and mercury compounds in the environment and aims at attributing the observed levels of mercury in key bioindicators to the mercury sources. In this tier, resource-intensive research methods and approaches are required. Monitoring sites and bioindicators may not all be the same across different tiers.

Key elements that are essential to all monitoring efforts, regardless of the tier under which they fall, for biota include: a) defining the target bioindicators and sample size, which usually focus on high trophic level biota that are vulnerable to relatively high methylmercury exposure; b) selecting and measuring the appropriate biomarkers (i.e., tissue types) to best interpret exposure to different sources and forms of mercury, with total mercury measurements in muscle tissue of fish and marine mammals, as well as blood, feathers or eggs of birds being most commonly used and accepted; c) identifying the monitoring locations and ancillary measurements that best reflect the objective for biomonitoring (e.g., temporal, spatial, source attribution, or estimating ecological or humans health questions) and d) managing and analyzing data as per the guiding questions for the Effectiveness Evaluation. All these aspects can use well-established standard operating procedures available in the Supplementary Material.

Chapter 5. Human Biomonitoring

5.1. Introduction

Understanding human exposures to chemical hazards through biomonitoring activities is important for scientific and regulatory purposes (WHO 2015). For mercury, in particular, human biomonitoring practices (i.e., mercury measures in hair, urine, and/or blood) are well-understood, practiced by some national governments, and can help assess the efficacy of policy actions (WHO 2018a; UNEP 2019; HBM4EU 2019).

The recent Global Mercury Assessment 2018 showcased biomonitoring efforts worldwide, and in doing so illustrated the diversity of efforts ranging from engagement of vulnerable communities situated in remote and resource-limited settings to national-level surveys implemented by government agencies involving thousands of participants (UNEP 2019). Human biomonitoring of mercury is relatively uncomplicated; these measurements are scientifically sound, technically simple with validated protocols available, and can be conducted at relatively low cost (Evers et al. 2016).

Human biomonitoring data can help address guiding questions that support the Effectiveness Evaluation (Table 2.1). First, quality measures of mercury levels in human biological samples (herein referred to as biomarkers) provide direct evidence of exposure in a given population at a given time. Second, such measures, when coupled with questionnaire data, may offer insights into possible sources and routes of exposure from which attributions may be deduced. Third, temporal changes can be gleaned if monitoring is repeated in the same population over time. Fourth, biomonitoring data can be inputted into established risk assessment frameworks to estimate health impacts including burden of disease, as well as to assess the efficacy of different risk management strategies. The guiding questions that support the Effectiveness Evaluation can provide the foundation to design a human biomonitoring study (that uses existing data and/or purposefully produces new biomonitoring data), and guidance for realizing this is detailed below.

Successful human biomonitoring activities require a multi-disciplinary team to work collaboratively across all aspects of the effort, from setting research questions that guide the design of biomonitoring activities to the interpretation and communication of results (Figure 5.1). Information in this chapter provides essential guidance (and links to key resources) for Parties and relevant organizations to consider in terms of using existing, and generating new, human biomonitoring data for the Effectiveness Evaluation. This chapter also provides a brief overview of our state of knowledge for human biomonitoring of mercury, proposes a framework by which biomonitoring data can be used for evaluating the effectiveness of the Convention, and then offers guidance on best scientific practices to: a) define the target and sample population; b) select and measure the appropriate biomarkers to help tease apart exposure to different sources and forms of mercury; c) administer surveys to gather supportive information to deepen understanding; and d) manage and analyze data as per the guiding policy question. All these aspects must be performed in a responsible and ethical manner. While the focus here is on Article 22 (Effectiveness Evaluation), many of the details below synergize with other articles of the Convention (e.g., Articles 4, 7, 14, 16-19, 21).

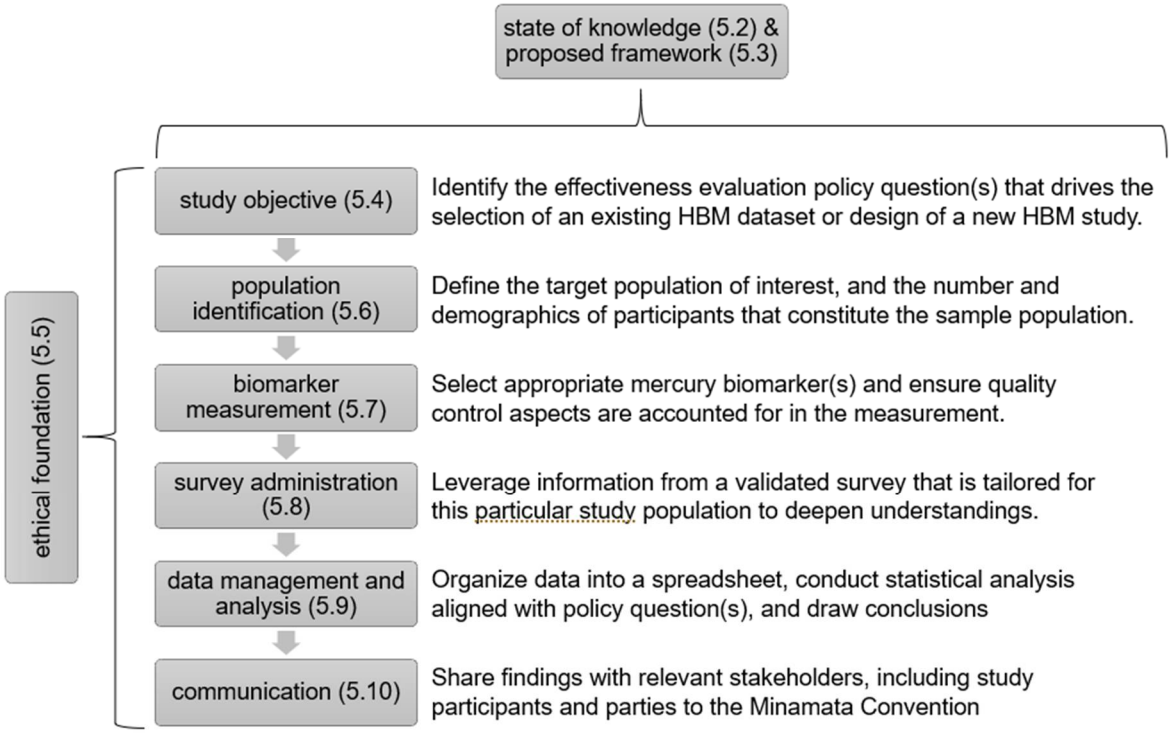

Figure 5.1. Proposed approach for using human biomonitoring (HBM) data for the purposes of the Effectiveness Evaluation. The proposed approach lists key elements that need to be considered when using existing HBM data or when planning a new HBM study. The numbers in parenthesis in the shaded boxes refer to chapter sections that offer more details.

5.2. State of knowledge

5.2.1. Existing data

To assess our current understanding of human exposures to mercury, a systematic search of the recent (2000 to 2018) literature identified 312 studies from 75 countries from which 424,858 mercury biomarker measurements from 335,991 individuals were analyzed (Basu et al. 2018). This activity was sponsored by the World Health Organization (WHO) as part of the Global Mercury Assessment 2018 (UNEP 2019). The authors of this report concluded that blood, hair, and urine mercury levels are generally less than 5 μg/L, 2 μg/g, and 3 μg/L, respectively, in background populations with no significant sources of exposure to mercury. The results also identified populations with elevated exposures. From this dataset there are two key groups of human biomonitoring data to be aware of.

First, national human biomonitoring programs exist that aim to derive information that is representative of a country or region. These are usually sponsored and/or operated by government agencies, are resource intensive, and generally cover many chemicals. These studies therefore tend to use random sampling of an adequate population size and use reference laboratories for mercury analysis. Sample sizes range from a few hundred to several thousand. The Global Mercury Assessment 2018 human biomonitoring dataset contains 192,651 biomarker measures from these programs. However, national biomonitoring programs that consider mercury exposure are only carried out in 9 countries to date, and international representation is mostly limited to higher income regions.

Second, there exist data (i.e., 232,207 biomarker measures) from cross-sectional and birth cohort studies. The design and quality of these studies vary tremendously. Further, the sample populations

usually are not representative of the target population as most rely on convenience sampling. Nonetheless, these studies are of importance as they tend to focus on vulnerable groups identified by the Minamata Convention (e.g., women of child-bearing age). Also, some of these efforts exemplify how mercury human biomonitoring may be performed successfully on a regional basis, such as the Arctic Monitoring and Assessment Programme (AMAP) and the DEMOnstration of a study to COordinate and Perform Human biomonitoring on a European Scale (DEMOCOPHES) effort.

5.2.2. Existing data gaps

Despite current understanding of human exposures to mercury worldwide, there is great variability in exposures around the world and across/within population groups. Arguably the greatest data gap concerns the many countries and regions without any mercury biomonitoring data without which evidence-based decision making is hampered. Notably, nearly 70% of the data in the Global Mercury Assessment 2018 biomonitoring dataset was represented by just 8 countries (Republic of Korea, China, Japan, United States, Brazil, Saudi Arabia, Canada, and the Russian Federation).

5.2.3. Future data sources

We can expect, with very high confidence, that mercury human biomonitoring data will be available in the future from two primary areas. First, some national human biomonitoring programs are firmly established by governments with sampling frequencies every 1-2 years (e.g., Canadian Health Measures Survey (CHMS), Czech Republic Environmental Health Monitoring System (EHMS), German Environmental Survey (GerES), Republic of Korea's National Environmental Health Survey (KoNEHS), US National Health and Nutrition Examination Survey (NHANES)), and these will be dependable programmes for evaluating the effectiveness of the Convention. Second, future data may also be expected from cross-sectional and birth cohort studies. Though, these are largely ad-hoc efforts run by academic researchers who depend on extramural funding, and as a collective they are not purposefully designed nor coordinated to address long-term effectiveness evaluation. It is also noted that many existing human biomonitoring programs, not necessarily designed for mercury exposure assessments, collect and archive blood samples (and other matrices) that may be analyzed retrospectively.

A third way forward, and in particular to help fill data gaps in a globally coordinated manner, Parties and relevant organizations without existing data sources should consider, where possible, a harmonized approach to launch new biomonitoring studies. A good starting point is the recent guidance from the WHO to characterize prenatal mercury exposure (WHO 2018a). Using this WHO protocol would enable the collection of comparable data (e.g., samples from 250 individuals per a defined study location, with minimum diversity recommended), through addressing the most vulnerable population group, i.e., the fetus. The studies would be country driven such that local ethical clearance would be required, and the studies would be conducted within the national health system. With funding from the Global Environment Facility (GEF), under the project "Develop a Plan for Global Monitoring of Human Exposure to and Environmental Concentrations of Mercury",[19] this WHO protocol was piloted between 2015 and 2017 in diverse settings and several countries. Examples of diverse human mercury exposure sources targeted in this WHO project included rice consumers (in China), seafood consumers (in Ghana and India), local industrial contamination (in India), mercury primary mining (in Kyrgyzstan), artisanal and small-scale gold mining (ASGM, in Mongolia), and freshwater fish consumers (in Russia). The GEF project showed that the generation of data using the WHO protocol in low- and middle-income countries is cost-effective, practical, and

[19] UNEP/MC/COP.3/INF/19.

feasible. The project also built local capacity to conduct relevant studies, which can therefore be repeated over time and in a range of locations to fill gaps.

5.3. Proposed framework

This section outlines a proposed framework in which monitoring programmes can provide comparable human biomonitoring data for the Effectiveness Evaluation. Driven by questions that support the Effectiveness Evaluation (Table 2.1), there are two main components to the proposed framework to bear in mind:

Pre-Minamata Convention period: 1) the use of existing biomonitoring data contained in the WHO-sponsored, Global Mercury Assessment 2018 biomonitoring dataset, or from other existing sources, can be used to help understand human exposures to mercury before the Minamata Convention's entry into force (i.e., help establish the baseline).

Effectiveness Evaluation period: 2) the use of biomonitoring data expected in the future from government-led national biomonitoring programs, regional initiatives, and/or academic-led studies; and 3) implementation of new biomonitoring studies led by Parties and relevant organizations in a harmonized way so that they are purposefully designed to fill data gaps, build capacity, and support the Effectiveness Evaluation. During the first Effectiveness Evaluation period, human biomonitoring activities may be designed according to the tiered approach outlined below.

The biomonitoring data collected from such activities: a) provide direct evidence of mercury exposure in a given population at a given time; b) when coupled with questionnaire data, offer insights into possible sources and routes of mercury exposure from which attributions may be deduced; c) can assess temporal changes in mercury exposure if monitoring is repeated in the same population over time; and d) assess potential health impacts and contribute to risk management activities.

The guidance presented below is intended to be fit for purpose, i.e., Minamata Convention stakeholders with narrow (e.g., specific country, population, or hotspot) or broad (e.g., global understandings, long-term trends) interests can generate comparable data to address the same relevant questions, albeit on different scales.

5.4. Tiered approach for human biomonitoring

Mercury human biomonitoring data can be designed as part of a Tiered approach for Parties and relevant organizations who may wish to improve existing monitoring programmes, or develop new programmes, with a view to contributing to the Effectiveness Evaluation. Details of the Tiered approach are summarized in the annex below.

Tier 1 – For Parties and organizations seeking to create a human biomonitoring program, or expand a minimal program, but that may not have sufficient resources to implement the actions in Tier 2, the goal should be to focus on a vulnerable sub-population (section 5.6) and take total mercury measurements in blood, urine, or hair (section 5.7). This activity should ideally be repeated in the same population every 2-5 years. A good starting point for Tier 1 is the recent guidance from the WHO to characterize prenatal mercury exposure (WHO 2018a).

Tier 2 – Building on Tier 1 activities, Tier 2 biomonitoring activities will perform more in-depth analysis of the Tier 1 sub-population group (e.g., measure total mercury in blood, urine and/or hair; consider measuring methylmercury and/or mercury stable isotopes in these biomarkers), or incorporate mercury biomonitoring into other, in-depth health surveys or cohort studies. These

activities are more expensive and complex than those under Tier 1, but they provide information that will address all guiding questions to support the Effectiveness Evaluation (table 2.1).

Tier 3 – To increase understanding of key processes that link mercury sources to human exposures, resource-intensive research methods and approaches are required. These include national human biomonitoring programs, or careful design of Tier 2 activities with coordinated air and biota sampling.

5.5. Ethics

It is imperative that human biomonitoring activities adhere to the World Medical Association's Helsinki Declaration, and that proper ethical approvals are in hand before any human subject research occurs. In most countries, Ministries of Health along with tertiary academic institutions, are the primary contact point for obtaining such ethical approvals. In some countries, sub-national/regional governments have self-determination of research activities and their own ethical guidelines and research licenses need to be followed, for example the National Inuit Strategy on Research.[20] Moreover, depending on the national context, specific organizations (e.g., workers unions, occupational safety boards, industry groups, dental/medical associations) may also have ethical guidelines to follow.

Given that human biomonitoring may focus on vulnerable populations, participatory engagement of pertinent stakeholders (e.g., study participants, workers, community leaders, health care providers, regional authorities) is necessary not only for ethical and safety purposes but to also help ensure that the best studies are designed, conducted, and communicated. The "International Ethical Guidelines for Health-related Research Involving Humans", prepared by the Council for International Organizations of Medical Sciences (CIOMS) in collaboration with WHO, should be consulted (CIOMS 2016). In addition, Parties and organizations may consult literature on legal, ethical, and social issues pertaining to human biomonitoring from the European Human Biomonitoring Initiative (HBM4EU 2018a), the Canadian Health Measures Survey (Day et al. 2007), the International Labour Organization's Technical and Ethical Guidelines for Workers' Health Surveillance (ILO 1998), and the World Health Organization's recent guidance on ASGM (WHO 2021a).

With regards to data ownership, human biomonitoring activities must respect the legislation of individual countries and this may vary depending on the population that is being sampled. For example, in Canada, Indigenous communities own the human biomonitoring data collected in their community (i.e., OCAP Principles – Ownership, Control, Access, and Possession), instead of the data being owned by the country or the organization responsible for generating the data. Appropriate communication and dissemination of data results back to the contributors is another important aspect of human biomonitoring. Moreover, in ethical research, all participants have the right to withdraw from studies/monitoring and have all their data and samples removed from the data set and no longer used.

5.6. Human population group

5.6.1. Identification of target population

All human populations worldwide are exposed to some amount of mercury (UNEP and WHO 2008; Basu et al. 2018). There is thus value in assessing mercury exposures in both the general population as well as in vulnerable groups. The selection of a specific target population will be guided by the interests of the Parties or relevant organizations carrying out the monitoring activities, in

[20] Available at https://www.itk.ca/national-strategy-on-research-launched.

consideration of guiding questions that support the Effectiveness Evaluation (chapter 2). For example, some initiatives may choose to focus on the general population while others may choose to focus on a specific vulnerable group (e.g., pregnant women, workers and community members living around ASGM sites, Indigenous Peoples, and local communities).

In terms of evaluating mercury exposures in the general population, the geographic scope (e.g., discrete community, entire country) and sociodemographic profile (e.g., sex, age) of this target population needs to be defined *a priori*. For guidance on studying general populations, Parties and relevant organizations can refer to aforementioned national human biomonitoring programs that tend to have detailed protocols available.

In terms of evaluating mercury exposures in population groups most vulnerable to mercury exposure, there are two broad groups to consider. First, early lifestages (i.e., fetus, newborn and children) are susceptible to mercury exposure because of the sensitivity of the developing nervous (and other physiological) system. This population group can also include pregnant women and/or women of child-bearing age. Second, some populations are vulnerable because they are exposed to higher levels of mercury. A resource document to help identify sub-populations that may be at risk of mercury exposure and health impacts was produced through a collaboration between UNEP and WHO (2008).

Human exposure to elemental and inorganic mercury may occur in occupational settings (e.g., ASGM and dentistry practices), from contact with certain products (e.g., dental amalgams, some skin-lightening creams, broken fluorescent bulbs and other waste products), and from environmental contamination (WHO 2008; Eagles-Smith et al. 2017; Ha et al. 2017; ATSDR 1999).

Human exposures to organic mercury largely arise from dietary sources. Mercury released into the environment may be converted by microorganisms to methylmercury which bioaccumulates and biomagnifies through the food web, particularly in aquatic systems (see chapter 4). Sampling of freshwater fish and seafood has found widespread methylmercury contamination, with some widely-consumed predatory species, such as tuna, swordfish, grouper, and mackerel being among the most highly contaminated.[21] Therefore, for many population groups, dietary consumption of contaminated fish, shellfish, and marine mammals is an important source of exposure. Seafood, however, is the main source of protein and nutrients for billions of people worldwide (FAO 2020). Other staple foods, such as rice, grown in sites with high concentrations of mercury may also represent a source of organic and inorganic mercury exposure for some communities (Rothenberg et al. 2014).

Well-studied population groups vulnerable to mercury because of higher exposures are listed here. From the Global Mercury Assessment 2018 report, four populations of concern were identified based on existing datasets: 1) Arctic populations (mainly Inuit) who consume high-trophic level fish and marine mammals; 2) tropical riverine communities (especially Amazonian) who consume fish, and in some cases may be exposed to mining operations; 3) coastal and/or small-island communities (including Indigenous Peoples) who rely substantially on seafood; and 4) individuals who either work or reside amongst ASGM sites. In addition to these relatively well-studied groups, other highly exposed groups for which there is awareness but relatively less data to draw firm conclusions include individuals living in mercury contaminated sites, certain occupational groups (e.g., chlor-alkali, dentistry), consumers of rice from contaminated sites, freshwater and marine fish consumers including sport fishers and Indigenous Peoples, and users of mercury-added products such as skin-

[21] Global Environment Monitoring System (GEMS) / Food Contamination Monitoring and Assessment Programme, available at: https://www.who.int/teams/nutrition-food-safety/databases/global-environment-monitoring-system-food-contamination.

lightening creams. In addition, there are certain ecosystems sensitive to mercury loading and methylation, and these may represent hotspots of biologically available methylmercury that warrant attention for those who consume local aquatic food items (see chapters 3 and 4). Coordinated studies that link human biomonitoring programs with data on environmental levels can help increase understanding of key processes that link mercury sources to human exposures.

5.6.2. Identification of sample population

Upon identifying a target population for investigation, the researchers would ideally sample all individuals from this target population, though achieving this is impractical (e.g., too many individuals to sample, it is prohibitively expensive, takes too much time and/or not everyone will agree to participate). Instead, researchers will sample a subset of the target population to realize a representative sample. Selection of the sample population needs to ensure that: 1) it is representative of the target population; and 2) there are sufficient number of people to yield valid information.

In order to select a sample population that is representative of the target population, it is necessary to understand the target population group's socioeconomic and demographic profile. In addition, it is important to understand the target population's mercury exposure profile (e.g., diet, occupation) and how this may change over time. The more specific the target population can be defined (e.g., age, sex, location, mercury exposure sources, seasonality, etc.), the easier it will be to identify a sample population with similar characteristics.

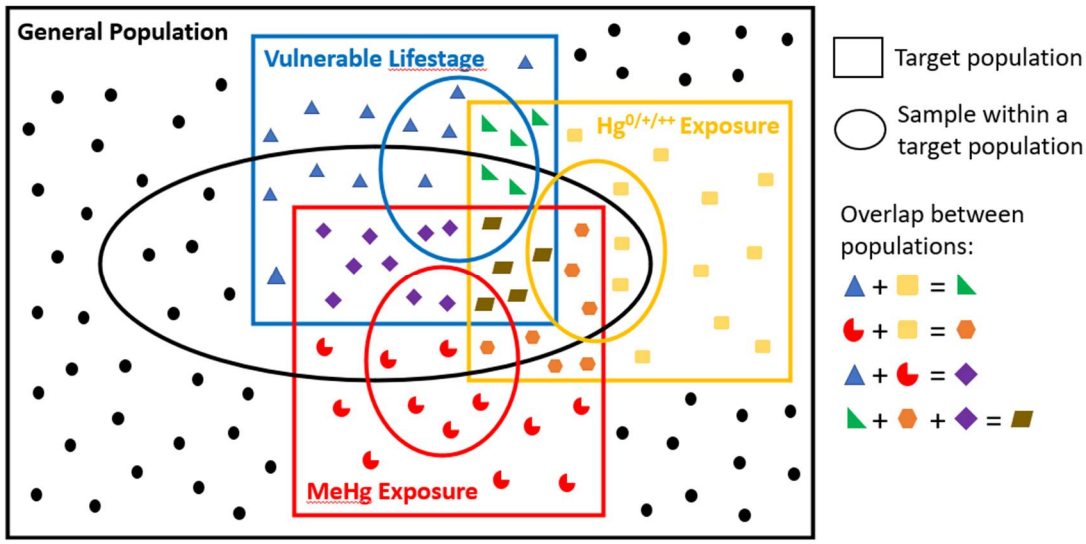

Figure 5.2. Population groups to consider. Within a country, exposures to mercury will be realized by all inhabitants (i.e., population universe), including members of the general population (outer black box) as well as members who are deemed vulnerable because of their lifestage or exposure situation (inner boxes colored blue, yellow and red). These population groups are not mutually exclusive as individuals may fall into multiple groups (e.g., those in ASGM sites may be exposed to both elemental mercury used in mining as well as methylmercury present in contaminated fish from local waterbodies as represented by the orange hexagons). Once a specific target population is selected to focus upon (driven by their interests in consideration of the guiding questions that support the Effectiveness Evaluation), steps need to be taken to help ensure that the sample population (i.e., the circles in the figure) is representative of the defined target population.

In order to select a sample population with a sufficient number of people, it is necessary to use statistical approaches that are aligned with the overarching aim of the biomonitoring effort.

Guidance on statistical approaches is covered in relevant guidance documents from WHO (2018a), HBM4EU (2017, 2018b, 2018c), along with many other resources (including online sample size calculators), and these need to be applied in a fit-for-purpose manner. To provide some additional context on possible sample sizes needed for a human biomonitoring study, the recent WHO guidance document on assessing prenatal exposures to mercury recommended a minimum of 250 pregnant women per site (WHO 2018a). In addition, the HBM4EU statistical plan (HBM4EU 2017) mentions the need for at least 120 measures to derive a biomarker reference value in a defined population (based on guidance from the International Federation of Clinical Chemistry RefVal program). A scan of national biomonitoring programs covered in the Global Mercury Assessment 2018 biomonitoring dataset reveals average sample sizes in the several thousands of people (Basu et al. 2018). While statistical approaches can help ensure that there are sufficient number of people in the study to yield valid information, other considerations will factor into sample size decision making including the size of the underlying population, financial costs, trained personnel, infrastructure, timeframe, and spatial scale. Further, during the study design phase there should also be careful consideration of whether the population can be re-sampled in the future to permit temporal trends analysis.

The nature by which participants are recruited and studied should be carefully detailed following guidance from the STROBE (Strengthening the Reporting of Observational Studies in Epidemiology) initiative (Vandenbroucke et al. 2007). Ideally the sampling process is free from any biases, and participants are selected in a random manner. All studies should include a participant flow diagram to help explain the generalizability and validity of the results obtained from the sample population.

5.7. Human biomarkers

Human exposures can be assessed through the measurement of mercury concentrations in a number of different types of biological samples, and key approaches for mercury biomonitoring (including detailed protocols on how to take samples from study participants and perform analytical measurements of mercury in the laboratory) have been recently outlined by WHO (2018b) and HBM4EU (2018d, 2019).

The most commonly used and accepted biomarkers are measures of total mercury concentrations in hair, urine, blood, and cord blood, and their selection can depend on factors such as the potential source of exposure, chemical form, and exposure lifestage. These biomarkers, in particular, were the basis for the human exposure chapter in the Global Mercury Assessment 2018 report (Basu et al. 2018). Some elaboration on these accepted biomarkers is provided below.

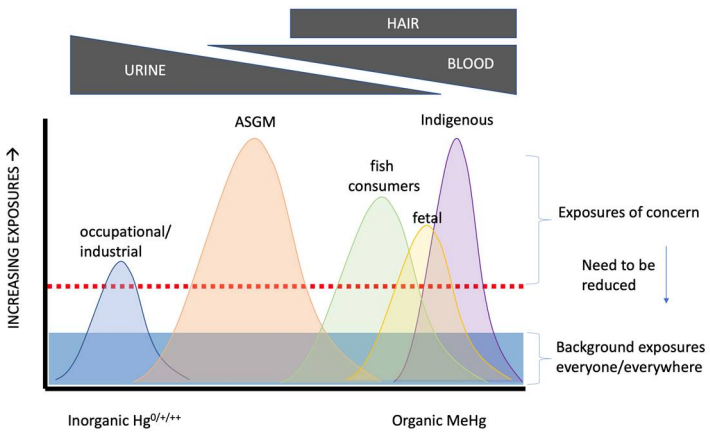

Figure 5.3. Diagram of accepted mercury biomarkers (along the top) in correspondence with the different chemical forms of mercury that these biomarkers represent exposure to (along the bottom). Key population groups identified to be of concern from the Global Mercury Assessment 2018 are outlined in the middle of the figure, along with a horizontal band along the bottom that represents general populations.

5.7.1. Human hair

Analysis of hair for total mercury concentration is commonly used to assess exposure to methylmercury (which accounts for 80–90% of the hair's total mercury content). Once incorporated, the mercury remains in the hair and this biomarker can therefore provide an integrated measurement of internal exposure to methylmercury. As hair grows at approximately 1 cm per month, exposures can be tracked over time by careful sampling (Lukina et al. 2021); for example, within person segmental hair analysis can integrate exposure data over several months, and examine differences across seasons or years.

Hair has the advantage that it is easy to collect, transport, and store, though in some communities there may be cultural objections to taking hair samples and in other groups (e.g., males, young children) short hair length may hinder proper sampling. Sampling should occur at the occipital region of the scalp for consistency and should be measured closest to the scalp to best reflect recent exposures (unless a longer temporal record is desired). In highly contaminated areas, there is a danger of external contamination of the hair, which can confound interpretation of the mercury measurement. For example, external contamination of hair by elemental mercury has been demonstrated in ASGM communities by use of mercury stable isotopes (Sherman et al. 2015). Therefore, when conducting studies in such contaminated sites care is needed in the interpretation of total mercury levels in hair. In such settings carefully analysing the hair for methylmercury, rather than total mercury, gives a better measure of dietary exposure especially when coupled with quality survey instruments, urine sampling, and biota measurements. Another potential challenge with hair monitoring in some communities may be the use of mercury-added cosmetic and beauty products. In such cases, hair total mercury levels may not accurately reflect dietary exposure to methylmercury. For this reason, when selecting individuals for hair monitoring, it is important to ascertain whether such mercury-added cosmetic products have been used. Further, when measuring hair mercury concentrations among such individuals, methylmercury (over total mercury) analysis is recommended.

5.7.2. Human urine

Analysis of urine for total mercury concentrations primarily provides information about recent (~1-2 months) exposure to inorganic and elemental mercury, although in people with high seafood consumption methylmercury may also contribute to the mercury content (Sherman et al. 2013). As the concentration of the analyte may depend on the dilution of the urine, which can vary, the measurement of mercury is often expressed in terms of its concentration per unit of creatinine or in relation to the specific gravity of the urine sample. The collection of urine, as with hair, is relatively easy, non-invasive, and cost effective, and there are good protocols available from WHO (2018b) and HBM4EU (2018d, 2019).

5.7.3. Human blood

Mercury is measured in whole blood and this provides information about recent exposures (~1-2 months) to both methylmercury and inorganic mercury. Though many human biomonitoring programs focus on blood mercury measurements, the collection is invasive and the storage and transport of blood can pose certain logistical and financial barriers particularly in resource-limited settings.

In most population groups, the measurement of total mercury levels in whole blood is an accepted biomarker for methylmercury exposure as it correlates relatively well to seafood consumption (Sheehan et al. 2014). However, in certain population groups (e.g., those who do not consume much fish and seafood, or have relatively high exposures to inorganic and elemental mercury), total

mercury may not be a good proxy for methylmercury exposure. Characterizing mercury chemical species or mercury stable isotopes in blood can provide an indication of potential sources, but these require careful sample preparation and advanced instrumentation. The US Centers for Disease Control and Prevention (CDC) now includes measures of blood methylmercury in NHANES and considers them more accurate in reflecting methylmercury exposures than measures of blood total mercury.

The measurement of total mercury levels in cord blood provides information about fetal exposure. Cord blood is collected following birth and often considered to be a non-invasive matrix, though this should be facilitated by a health care professional (e.g., nurse). Many jurisdictions have newborn screening programs in which newborn blood is sampled and archived as dried blood spots, and while mercury analysis of these dried blood spots shows promise they require careful consideration. Notably, dried blood spots are also collected in some demographic health surveys (e.g., USAID's DHS Program) which are present in over 90 countries.

5.7.4. Integrated biomarker approach

Each biomarker can provide pertinent exposure information on the type of mercury (organic vs. inorganic) and timeline of exposure (recent vs. chronic). As per above, human hair, urine, and blood are commonly used biomarkers of mercury exposure, and anyone of these three biomarkers can be selected by Parties for Tier 1 biomonitoring activities. When multiple biomarker measurements are taken from a given individual (along with mercury speciation analysis and questionnaires), a deeper exposure assessment can be performed (i.e., under Tier 2 or Tier 3 biomonitoring activities). Measurements of total mercury in hair and urine are particularly suitable (especially in resource limited settings) as they provide a relatively low-cost and non-invasive scheme to gauge exposure to the main forms of mercury. Further, with basic training, sampling and handling procedures are easy to implement, and quality assurance programs and suitable reference materials are also in place to help ensure comparability of measurement results (i.e., see good protocols from WHO (2018b) and HBM4EU (2018d, 2019) on how to take samples from study participants and perform analytical measurements of mercury in the laboratory). Biomarker measures can be further improved by also including survey instruments (see section 5.8) that collect pertinent information on the study population and exposure sources.

5.7.5. Biomarker measurements

A number of analytical methods (e.g., cold vapour atomic absorption spectrometry (CV-AAS) and cold vapour atomic fluorescence spectrometry (CV-AFS) are most widely used and accepted) are available to quantify the concentration of mercury in a given biomarker type, and these are detailed in a recent WHO guidance document (WHO 2018b) and by HBM4EU (2019). The selection of a particular analytical method will depend on factors such as availability of trained laboratory personnel and instrumentation. Regardless of the analytical method selected, it is important to practice careful quality control including the use of suitable reference materials (e.g., urine: INSPQ/Quebec; hair: NIES/Japan or IAEA/Austria; blood: NIST/US, INSPQ/Quebec) and attention to parameters such as detection limits, accuracy, and precision. It is also important to report the methods followed and QA procedures used. Analytical laboratories are encouraged to participate in quality assurance programs, such as the one run by AMAP/NCP, and these programs should be prepared to expand capabilities and provide assistance to nascent labs.

For the purposes of human biomonitoring (and as detailed above and in the included references), measures of total mercury content in a given biomarker will suffice in most cases. Such measures can be realized in under 10 minutes with minimal sample preparation using operationally simple,

commercially available benchtop instruments that integrate sample decomposition with gold amalgamation and spectrophotometry.

5.8. Survey protocol

Combining the results of mercury biomarker measurements (section 5.7) with survey questionnaire information (e.g., sociodemographic data, occupational practices, dietary habits) from the same individual provides the basis for an assessment that can deepen understanding of exposure sources and routes as well as the extent, duration, frequency, and magnitude of exposure. Survey instruments relevant to mercury are available from WHO (Annex 3 in WHO (2018a)) and HBM4EU (2020b, 2020c, 2020a).

Surveys should be tailored for the target population (e.g., culturally appropriate, language, education level, relevant food items, lifestyle, and occupation) and have undergone proper pilot testing and validation. Those conducting surveys should have received training on proper methods to help ensure that valid and complete data are captured in a standard manner, and to identify and avoid possible sources of survey bias (for example, recall bias, estimations of serving sizes and frequencies). The survey data could also be amenable for capture into an electronic format.

5.8.1. Methylmercury exposures

Most populations worldwide are exposed to methylmercury through the consumption of fish and seafood (Sheehan et al. 2014; EFSA 2012). Thus, dietary intake of mercury from these items can be estimated if information is available on the: a) types and amounts (frequency and serving size) of food ingested per unit time (day or week); b) mercury concentrations in these food items (on a wet weight basis); and c) the participant's body weight. Consumption of certain food items may vary seasonally, and mercury concentrations may vary across animal parts and be influenced by food preparation steps, and all of these need to be taken into account when conducting an exposure assessment. From a modelling perspective (chapter 6), it is also helpful to know the source of the food item (e.g., sampled locally or through international trade markets). As many of the food items that deliver mercury into human populations are also ones with high nutritional value, assessments should strive to examine risk-benefits (Mahaffey et al. 2011). Parties and relevant organizations could consider human biomonitoring efforts in geographic sites where biota are being sampled to maximize efficiencies and data quality (chapter 4). Detailed protocols for developing dietary surveys are available from the WHO (WHO 2008) and USEPA (USEPA 2016), and the HBM4EU has a comprehensive dietary questionnaire that may be adapted to fit particular needs (HBM4EU 2020a).

5.8.2. Elemental and inorganic mercury exposures

Human exposures to elemental and inorganic mercury may occur in occupational settings (e.g., in ASGM sites, chlor-alkali plants, and dentistry practices), from contact with certain products (e.g., dental amalgams, some skin-lightening creams, broken fluorescent bulbs and other waste products), and from environmental contamination (WHO 2008; Eagles-Smith et al. 2017; Ha et al. 2017; ATSDR 1999).

Identification of a target population based on these particular exposures should trigger the need to include screening level assessment surveys to deepen understanding of potential exposures. Examples of relevant screening level assessments for mercury are available from WHO (2018a) and HBM4EU (2020b, 2020c, 2020a). For the ASGM sector, guidance from WHO provides templates and tools for conducting assessments to provide an evidence base for the development of public health strategies required for National Action Plans (WHO 2021b). There is also a survey from a UNIDO/UNDP/GEF-sponsored initiative that is often used (Veiga and Baker 2004), which needs to be

applied with careful attention to tease apart different job tasks, the proximity of ASGM sites to households, and location of smelting and ore processing sites. For dentistry, a collaboration between the American Dental Association and academics yielded a survey tool to relate occupational practices with exposure biomarkers (Goodrich et al. 2016), and the HBM4EU has a survey with pertinent questions concerning personal amalgams (HBM4EU 2020a).

5.9. Management and analysis of human biomonitoring data

5.9.1. Existing and future data

Existing data, as contained initially in the WHO-sponsored, Global Mercury Assessment 2018 human biomonitoring dataset (Basu et al. 2018), can be updated to help establish a "baseline" for human biomonitoring under the Effectiveness Evaluation. In terms of future data, we can expect, with very high confidence, that biomonitoring data will be available from government-led national biomonitoring programs as well as academic-led cross-sectional and birth cohort studies. In addition, to help fill data gaps in a coordinated manner and build capacity, Parties and relevant organizations are encouraged to consider recent guidance from the WHO on a harmonized approach for conducting new biomonitoring activities (WHO 2018a).

5.9.2. Data quality

Quality practices are necessary to help ensure that biomonitoring results are valid, free of bias, and comparable across studies and regions. In terms of ensuring that field work is conducted properly, information presented earlier under sections 5.6 (Human population group) and 5.8(Survey protocol) should be consulted, along with resource documents from WHO (WHO 2008, 2018a) and the STROBE initiative (Vandenbroucke et al. 2007). It is essential that studies collect critical details on the sample population (e.g., age, sex, location, sample month/year), how they were recruited, and details on sources and routes of mercury exposure. In terms of biomarker measures, information presented earlier under section 5.7 (Human biomarkers) should be consulted so that studies use proper reference materials, participate in inter-lab comparison programs, and report on analytical parameters such as detection limits, accuracy, and precision.

Based on guidance from the US National Toxicology Program's (NTP) Office of Health Assessment and Translation (OHAT 2015), and as considered as part of the Global Mercury Assessment 2018 biomonitoring dataset, a Risk of Bias score can be derived for each study that considers: a) participant selection bias (e.g., selection method, demographics, exposure characteristics, timing of recruitment); b) exposure detection bias (e.g., quality of the methods used to measure the mercury biomarkers, recall bias); and c) statistical and other bias (e.g., biomarker distribution, reporting mercury exposure sources). Such a score can help give users of the data a frank assessment of its quality, and be used to flag potential concerns.

5.9.3. Data exchange

Paragraph 1 (d) of article 17 of the Convention calls for Parties to facilitate the exchange of epidemiological information, in close cooperation with the WHO and other relevant organizations, as appropriate. To facilitate the implementation of that article of the Convention, cooperation for the compilation and exchange of data via an appropriate knowledge-sharing platform may be considered.

For each biomonitoring study to be included in such a database there is a need for minimal essential information to help ensure that studies can be compared. These include group-level data on: sample population characteristics (population type, sample size, age, sex, education, socioeconomic status,

personal amalgams, city/region/country, day/month/year), analytical measurements (sample size, biomarker type, speciation information, quality control including detection limit, accuracy, precision, and use of reference materials), and mercury values (count (n)), percentiles including 10th, 25th, 50th, 75th, 90th, and 95th values; additional measures of central tendency (variance) including mean (SD) and geometric mean (95% CI); indication of data normality; these align with guidance from the HBM4EU statistical analysis plan (HBM4EU 2017)). Strategies for dealing with missing data and measures below detection limits are provided in the HBM4EU statistical analysis plan (HBM4EU 2017). Key information from surveys (e.g., dietary intake values; occupational practices; other exposure sources) needs to also be extracted and summarized. The data should be aggregated for the entire sample population as a primary level summary, as well as for key sub-groups (e.g., different lifestages, sexes, locations, occupational categories) as part of a secondary level summary. Finally, studies must name the ethics board that approved their work. Section 11.2.2 of the HBM4EU statistical analysis plan provides a good list of variables specific to mercury organized into exposure levels, time trends, geographical comparisons, and exposure determinants that largely align with the information listed here (HBM4EU 2017).

The focus of the human biomonitoring data should be on a population group. While compiling individual-level data may permit deeper scientific analysis, realizing this for research of human subjects is extremely challenging owing to ethical, privacy, logistical, and other concerns. The WHO guidance document provides guidance on handling individual-level data, i.e., participating countries conduct statistical analysis in-country, and then submit anonymized summarized data to a central database for international-level analyses (WHO 2018a). A similar approach may be taken for group-level data as well, with good details offered by HBM4EU on handling both individual- and group-level data (HBM4EU 2017).

5.9.4. Data analysis

Statistical analysis of human biomonitoring data may help address the questions that support the Effectiveness Evaluation (table 2.1). Detailed guidance on statistical analysis of human biomonitoring data is offered by HBM4EU, and it covers aspects such as treating missing data, time trends analysis, geographic comparisons, and uncertainty analysis (HBM4EU 2017). Five key statistical analyses are listed below that align with the monitoring objectives and guiding questions. More sophisticated aspects of data analysis (especially modelling) are provided in chapter 6, and here basic guidance is provided on how to analyze mercury human biomonitoring data.

Descriptive statistics: Descriptive statistics should be used to summarize key features of the sample population and their exposures to mercury. This information can be used, for example, to characterize spatial variability, and help identify hotspots and exposure sources. The data can also be used to indicate the percentage of those sampled with mercury biomarker values that exceed a guideline value or reference range at a certain place and point in time (these are summarized in Basu et al. (2018). Such descriptive information can then be represented visually on a map with a color scale as done for an assessment of human biomarker values from across Europe (Višnjevec, Kocman, and Horvat 2014).

Exposure assessment: To increase understanding of possible sources and routes of mercury exposure, regression-based approaches may help associate mercury biomarker measures (dependent variable) with independent variables drawn from the survey data (e.g., dietary intake, occupational practices). There are many published studies of this kind for a diverse range of mercury exposure scenarios.

Temporal analysis: Over time changes can be gleaned if repeated monitoring is performed in the same population over time. This requires that the geographic scale (local to national to global) and the target population (e.g., background, specific vulnerable, life stage, etc.) be defined, and then differences in mercury biomarker measures be compared. Depending on the context, seasonality of sampling may be an important consideration here. Section 6 of the HBM4EU statistical analysis plan provides detailed guidance on temporal trends analysis (HBM4EU 2017).

Attributive analysis: If temporal changes in mercury biomarker levels are found, stakeholders will want to know if changes are attributed to actions taken under the Minamata Convention. This will require exposure assessments and temporal analysis to be combined, and with consideration of discrete policy actions taken. Past examples of changes in human biomonitoring levels being linked with a specific action, as discussed in the Global Mercury Assessment 2018 report, include: a) decreasing blood and hair mercury levels have been reported in population groups from the United States, Denmark, the Faroe Islands, and several Arctic communities that may be linked with dietary consumption advisories and/or changing dietary habits; and b) decreasing urinary mercury levels among the general US population, German children, and some dental professionals is likely associated with the development of encapsulated amalgams, the increasing use of composite resins, and the overall awareness of occupational and environmental risks associated with mercury use.

Risk assessment: One of the ultimate goals of the Minamata Convention is to protect human health from mercury. Established risk assessment frameworks (e.g., EFSA 2012) may be used to calculate the nature and probability of mercury-associated adverse human health effects. From such data, burden of disease estimates and economic costs may be calculated, and changes over time may be explored under actual conditions and future scenarios using modelling tools.

5.10. Communication

Communication of results is a critical aspect of human biomonitoring. The HBM4EU program offers guidance on how human biomonitoring data could be organized into a report (HBM4EU 2020d), and the WHO offers guidance on how researchers should engage with stakeholders throughout the project's life course, and how biomonitoring findings should be shared with study participants, the general public, public health professionals and policy makers (WHO 2018a). In addition, particular consideration is needed with regards to contaminant research pertinent to Indigenous populations, where a partnership approach and equitable engagement ensures successful communication of monitoring and research results (see AMAP for examples of positive and negative experiences (e.g., AMAP 2021). Parties and relevant organizations may also decide on if (and how) the data is used for risk management.

5.11. Conclusions

Human biomonitoring data can help address the monitoring objectives and guiding questions that will support the Effectiveness Evaluation (see Table 2.1). The information in this chapter provides essential guidance (and links to key resources) for Parties and relevant organizations to consider in terms of using existing, and generating new, human biomonitoring data for the Effectiveness Evaluation.

In terms of using existing biomonitoring data, several databases and resources exist, and these can be used to help understand human exposures to mercury before the Minamata Convention's entry into force (i.e., help establish the baseline).

In terms of data to be realized during the Effectiveness Evaluation period, there are two sources to consider. First, biomonitoring data in the future are expected to be realized from existing

government-led national biomonitoring programs, regional initiatives, and/or academic-led studies. Second, Parties and relevant organizations can further support the Effectiveness Evaluation by implementing new biomonitoring studies in a harmonized way so that they are purposefully designed to fill data gaps, and build capacity.

Human biomonitoring data can be designed as part of a Tiered approach to inform new monitoring programmes or improve existing ones (see section 5.4 and the annex to this document). Briefly, Tier 1 is for those seeking to create a human biomonitoring programme, or expand a minimal programme, but that may not have sufficient resources to implement the actions in Tier 2. The goal of Tier 1 should be to focus on a vulnerable sub-population (section 5.6) and take total mercury measurements in blood, urine, or hair (section 5.7). This activity should ideally be repeated in the same population every 2-5 years. A good starting point for Tier 1 is the recent guidance from the WHO to characterize prenatal mercury exposure (WHO 2018a). Tier 2 aims to realize information that will help address all monitoring objectives and guiding questions in Table 2.1, and thus calls for more in-depth analysis of the Tier 1 sub-population group, or incorporation of mercury biomonitoring into other, in-depth health surveys or cohort studies. Tier 3 aims to increase understanding of key processes that link mercury sources to human exposures, and thus resource-intensive research methods and approaches are required.

There are essential elements to all human biomonitoring studies that need to be considered, and these are outlined in Figure 5.1 and elaborated upon in this chapter. Key elements include: a) defining the target and sample population (which usually focus on groups vulnerable to mercury, i.e., early lifestages or those with relatively high exposures); b) selecting and measuring the appropriate biomarkers to help tease apart exposure to different sources and forms of mercury (with total mercury measurements in hair, urine, blood and cord blood being most commonly used and accepted); c) administering surveys to gather supportive information (e.g., on socio-demographics, occupational practices, dietary habits) to deepen understanding; and d) managing and analyzing data as per the guiding policy question. All these aspects must be performed in a responsible and ethical manner.

Chapter 6. Cross-media data management and analysis

6.1. Introduction

Chapters 3, 4 and 5 provide guidance on the collection, management and analysis of data in air, biota and from human biomonitoring. By analysing monitoring data, temporal and spatial trends in the levels of mercury in specific environmental media or human matrices can be derived with confidence intervals. These trends provide a first-level indication of whether the Convention may be contributing to protecting human health and the environment from the adverse effects of mercury. Analyses of the monitoring data collected in each medium separately will be highly informative, and cross-media analysis incorporating the known mechanistic connections between media can provide further information, adding to the scientific weight of evidence that can inform the Effectiveness Evaluation. This chapter elaborates on how these monitoring data can be used in an integrated manner, where combining multiple complementary analysis approaches to answer the same question will improve robustness. This will facilitate understanding of the temporal trends and spatial patterns of mercury observed in the environment and humans, and the impact of actions motivated by the Convention.

Because the connections between monitoring media are not necessarily direct and instantaneous but do depend on physical processes or human behaviours, mechanistic models[22] explicitly representing these processes are a valuable tool for interpretation of monitoring results and can thereby contribute to the Effectiveness Evaluation. This makes cross-media analysis involving both mechanistic and statistical modelling in all relevant media an important part of the weight of evidence useful to evaluate effectiveness of the Convention. Moreover, evaluating the effectiveness of the Convention requires separating the impacts attributable to the Convention from changes that occur due to other factors, such as climate change. While monitoring data shows the impact of all of these factors, modelling can help attribute the changes to the different drivers. However, as the complexity of the modelled system increases, identifying all the relevant processes and quantifying them correctly becomes more challenging. In such cases, mechanistic models can be supplemented with different kinds of statistical models. Attribution of observed trends to specific drivers such as direct anthropogenic mercury releases, legacy mercury, natural process-driven releases, and non-mercury environmental or behavioural drivers requires the use of models which resolve the intervening processes supplemented or calibrated by empirical statistical approaches. From primary release to human exposure, mercury can undergo many physical and (bio-)chemical changes which interact with each other over a large range of timescales, and these can be influenced by human behaviour. Specific types of models are described in various locations throughout the chapter (see Model Descriptions 6.1-6.8), but each model's relevance is not limited to the subject matter with which it is first discussed.

Monitoring data and other ancillary observational data can be used in a variety of ways in concert with mechanistic and statistical models to quantify the effectiveness of Convention measures. Data from each medium can be used to evaluate that medium's model representations, and to identify situations where a given model is or is not appropriate for use. Monitoring data from one medium can also be used as input to models to explicitly connect outcomes in that medium to outcomes in other media (e.g., wet deposition can serve as input to an aquatic ecosystem model to estimate fish concentrations), or to models which can attribute those trends to specific sources or drivers. Tables

[22] Mechanistic models are based on the mathematical representation of well-known physical, chemical, or biological laws describing the behavior of constituent parts of the modeled system to make predictions of how something will play out in the real world. In contrast, empirical modeling uses observations or experiments to get statistical relationships to potential drivers.

6.1 and 6.2 summarize, for monitoring data and ancillary observational data respectively, the data, metadata, and other information that can facilitate cross-media analysis and modelling. Where available, monitoring data from ocean and freshwater, although not core media in this guidance, may provide important information to strengthen the accuracy of analysis and prediction.

The analyses discussed below fall into two main categories of approach. The first is the top-down approach, which directly uses monitoring data and statistical relationships to relevant variables to infer importance of specific drivers from the observational data. The second is the bottom-up approach, which uses mechanistic models representing physical processes to produce estimates of the quantities that are observed based on inputs to the modelled system. These two approaches can be interpreted as propagating information in opposite directions, the former from observed quantities to their drivers and the latter from the drivers to observable quantities. Both approaches can be useful and are discussed further in the following sections.

Table 6.1. Information from monitoring data. Listed for each medium and tier are the primary monitoring data, metadata, ancillary data for interpretation and to aid in analyses, and the analyses for which those data can be used.

Monitoring category	Observation Data	Metadata	Ancillary Data	Analyses
Air - Tier 1	TGM and GEM levels; Wet deposition	Latitude; longitude; elevation; Sampling time, frequency, duration; averaging methods; sampling method	Measurement/method uncertainty; proximity to known point sources; type (urban/regional/background); meteorological variables;	● Temporal trends ● Atmospheric model evaluation (for GEM) ● Spatial variations ● Input for local-scale modelling ● Back-trajectory analysis ● Bottom-up attribution analysis[23]
Air - Tier 2	Air - Tier 1 and High-resolution PBM and GOM; Estimates of dry deposition of mercury (using concentrations and site specific deposition velocities); mercury throughfall	Air - Tier 1	Air - Tier 1 and deposition of Sulfate; Land Cover; Land Use; Leaf Area Index; Air Quality Tracers (e.g., SO_2, CO_2, CO, $PM_{2.5}$, O_3)	Air - Tier 1 and ● Estimate air-ocean and air-terrestrial mercury exchange ● Covariate profiling ● Top-down attribution analysis
Air - Tier 3	Air - Tier 2 and mercury isotopes; Measurements of dry	Air - Tier 2	Air - Tier 2	Air - Tier 2 and ● Combined “top-down” and “bottom-up” attribution analyses ● Isotopic fingerprinting

[23] The term "Bottom-up" is being used to refer to a process-based analysis estimating effects of drivers on observable quantities. The term "Top-down" is being used to refer to an observation-based analysis for identification/estimation of drivers.

	deposition; additional speciation measurements			
Biota - Tier 1	Tissue/organ mercury and/or methylmercury levels; distribution statistics or quantiles	Geolocation or water body name; Spatial coverage; sampling time period; method info; tissue/organ type; habitat; wet or dry weight	Measurement/method uncertainty; population sample size; species; length/mass; trophic position/diet info; age; sex; maturity stage; carbon and nitrogen isotopic data; lake size; known point source or sediment contamination;	● Spatial variations ● Temporal trends ● Input for local exposure modelling ● Guideline value exceedance statistics
Biota - Tier 2	Biota - Tier 1	Biota - Tier 1	Biota - Tier 1 and carbon and nitrogen stable isotopes; water DOM/DOC/TOC, TSS, salinity, DO, (pH). N and P, Chl-*a*; total mercury and TOC in sediment; GEM in air; wet deposition; meteorological data	Biota - Tier 1 and ● "Top-down" biota mercury attribution ● Watershed and food web model evaluation
Biota - Tier 3	Biota - Tier 2	Biota - Tier 2	Biota - Tier 2 and speciated mercury, mercury stable isotopes in biota and suspected source-matrices; chemical tracers related to known drivers; diet information; stable isotopes of prey organisms; food web structure	Biota - Tier 2 and ● Combined "top-down" and "bottom-up" biota mercury attribution ● Isotopic fingerprinting
Human - Tier 1	Total mercury levels in hair, blood, or urine (10^{th}, 25^{th}, 50^{th}, 75^{th}, 90^{th}, and 95^{th} percentiles);	Geolocation or city/country/region; Population sample size; Spatial coverage; population type; sampling time period; method info; type of biomonitoring sample;	diet info; age; sex; known occupational and other exposures; education, socioeconomic status, amalgam status; additional measures of central tendency (variance) including mean (SD) and	● Spatial variations ● Temporal trends ● Exposure model evaluation ● Input for local health impact / risk assessment modelling ● Guideline value exceedance statistics ● "Top-down" exposure

			geometric mean (95% CI); indication of data normality; measurement/method uncertainty	attribution
Human - Tier 2	Human - Tier 1 or cord blood (10^{th}, 25^{th}, 50^{th}, 75^{th}, 90^{th}, and 95^{th} percentiles); optionally methylmercury; mercury isotopes	Human - Tier 1	Human - Tier 1, dietary intake amount and associated relevant air and biota measurements	Human - Tier 1 and ● Isotopic fingerprinting
Human Tier 3	Human Tier 2	Human - Tier 2	Human - Tier 2	Human - Tier 2 and ● Combined "Top-down" and "bottom-up" exposure attribution

Table 6.2. Example observational data from other media to support primary monitoring. Listed for each medium are the primary data, metadata, ancillary data to aid in analyses, and analyses for which the data are used.

Medium	Data	Metadata	Ancillary Data	Analyses
Soils	Mercury levels;	Latitude; longitude; depth; sampling date; averaging methods	Measurement/method uncertainty; presence of known point sources; soil horizon; land use; carbon concentrations; surface fluxes	● Terrestrial model evaluation ● Input for local-scale modelling ● Input for atmospheric modelling ● Atmospheric model evaluation
Vegetation	Mercury levels;	Latitude; longitude; sampling time; averaging methods	Measurement/method uncertainty; vegetation type; NDVI; carbon fluxes; Hg exchange fluxes; litterfall fluxes	● Terrestrial model evaluation ● Input for local-scale modelling ● Input for atmospheric modelling ● Atmospheric model evaluation

Food items and other products	Methylmercury and total mercury levels; statistical distribution information	name of product; type of food item; country/region; sampling time period	Consumer population; exposure type (diet, skin, etc.)	• Input for exposure modelling
Freshwater	Dissolved and particulate mercury and methylmercury levels;	latitude; longitude; depth; Sampling time; averaging methods; water body name	Measurement/method uncertainty; dissolved and particulate carbon concentrations; temperature	• Input for food web modelling
Ocean	Dissolved and particulate mercury and methylmercury levels;	latitude; longitude; depth; Sampling time; averaging methods; water mass name	Measurement/method uncertainty; dissolved and particulate carbon concentrations; nutrient concentrations; temperature; salinity; dissolved oxygen	• Ocean model evaluation • Input for food web modelling • Input for atmospheric modelling • Atmospheric model evaluation
Sediment	Mercury levels; methylmercury levels; mercury accumulation rates;	type of sediment; latitude; longitude; water depth; sediment depth; dating info; dating method	Measurement/method uncertainty; accumulation rates; total organic carbon; grain size	• Input for watershed modelling • Input for food web modelling • Mass balance model evaluation
Snowpack	Mercury levels;	latitude; longitude; Sampling time; averaging methods; sampling methods	Measurement/method uncertainty; snow depth; accumulation rates; snow density; exchange fluxes;	• Atmospheric model evaluation • Input for local-scale modelling

6.2. Maximizing scientific weight of evidence

Some media-specific analyses are outlined in chapters 3, 4, and 5, and can be useful tools to inform Effectiveness Evaluation via single-medium monitoring data. Chapter 3 discusses management, analysis and evaluation of atmospheric mercury data and provides tools for obtaining a more holistic picture of the state of mercury in air by adding value to the monitoring data that is collected. Chapter 4 enumerates the primary and ancillary monitoring measurements for biota which can be used for time series analysis accounting for variability associated with multiple factors, and discusses ecosystem sensitivity analysis to identify and prioritize sites for most effective use of limited monitoring resources. Chapter 5 highlights descriptive statistics and temporal analysis on human biomonitoring data used to summarize population exposures to mercury and how they change in time, exposure assessments using survey data to associate biomarker measurements to possible sources, and risk assessment to connect to human health. Such tools can also be combined into integrated analyses across media to provide further information. Chapter 2 presents guiding questions (Table 2.1) which serve as a guide to producing a continuum of evidence, and the following sections describe tools and analyses which can be used to maximize the scientific weight of evidence used to evaluate the effectiveness of the Convention, by providing multiple strong

individual lines of evidence. Furthermore, the Supplementary Material elaborates on laboratory intercomparisons for identifying biases and uncertainties in air monitoring, time trend identification, covariate analysis for source identification, and backwards trajectory models for source-receptor relationships.

6.2.1. Estimation of background and impacted levels of mercury

Analysis of monitoring data from sites chosen and categorized to represent background and impacted locations can directly estimate levels of mercury for these types of areas. Aggregating results of total air mercury and wet deposition monitoring by site type (e.g., urban/point-source-influenced/background) can show high-level source influences using data from all tiers. Similarly, biota monitoring data of all tiers at locations identified as background or affected by anthropogenic sources can be used to estimate these biota-specific levels of mercury. Summary statistics comparing some subpopulations from human biomonitoring can be used to establish mercury levels for some background and impacted locations where representative subpopulations have been sampled. Summaries from all matrices can include both mercury levels and guideline value-exceedance statistics. These most basic analyses can usually be done with a high degree of confidence using basic monitoring strategies, yet can provide valuable information for the effectiveness evaluation.

6.2.2. Identification of trends over time

The effective identification of temporal trends should yield key pieces of information: the magnitude of the trend, an associated confidence interval, and a summary measure of the statistical significance of the trend. It is more difficult to identify trends in areas with high temporal variability because the magnitude of the trend is more likely to come with a relatively wider confidence interval and lower statistical significance. Including the confidence interval and statistical significance helps to avoid over-interpretation of observed trends. Grouping sites by region or type for analysis can be done to assess general patterns in time trends.

Statistical analyses can be performed on time series data from air monitoring sites (Tier 1, 2, and 3) to identify observed trends - both at an individual site level and across groups of sites - which take into account sources of temporal variability. The significance of an upward or downward trend across a time period and its standard error can be identified using appropriate statistical tests. The magnitude of the trend and its confidence interval can be obtained using this information and statistical methods which are robust to outliers and data which deviate from linear time behaviour.

For biota monitoring (Tier 1, 2, and 3), if individual sample data are available, generalized linear modelling can account for time-variations in mercury levels in a way that controls for drivers included in the measured ancillary data and metadata. Consistency in data quantities available across monitoring locations is important for accurate application of this method.

Human biomonitoring (Tier 1, 2, and 3) differences across time can be identified on a subpopulation basis if repeated monitoring is performed in the same population over time. This requires that the spatial coverage, sampling timing and population be well defined, to be able to control for covariates such as seasonality, life-stage, and human activities (e.g., diet, occupation), etc.

Bottom-up modelling can be performed to quantify expected trends or relative trends in locations, forms of mercury, and matrices that lack direct monitoring. Spatially resolved models which show consistency with observed trends in monitoring locations can estimate expected trends in other locations with similar characteristics. Site-specific modelling can extend the observed trends in

monitoring media to other media and to exposure and health impacts through cause-effect relationships.

Modelled temporal variability is a quantity which requires careful consideration of model inputs and assumptions. Variability driven by environmental variables such as temperature and weather will only be quantifiable by a given model on the timescale that those variables are represented in the model. Often input variables are averaged over days to years, depending on the model and the input, and therefore shorter-timescale variability will be under-represented in model output.

Model Description 6.1: Atmospheric models

Atmospheric chemistry transport models represent the fate of mercury upon release to the atmosphere. They represent the chemical and physical changes in the form of the released mercury using experimentally or theoretically determined reaction rate and partitioning coefficients. Atmospheric models can be global- or regional-scale gridded models or trajectory models which trace the dispersion of air parcels forward from sources or backwards from receptors. To trace emissions to receptors, they require specification of the magnitude and spatial distribution of releases of mercury to air: as anthropogenic and geogenic direct emissions, as well as terrestrial and ocean fluxes of legacy mercury. Since these models directly simulate atmospheric concentrations and deposition, measurements of these quantities are best suited for evaluation. In comparing these quantities, it is important to consider that concentration and deposition measurements are often performed at a single point, while gridded atmospheric models represent the values over some area depending on the model grid size, and gridded and trajectory models rely on the resolution of the underlying meteorological data. Therefore, these models can be limited in their ability to resolve high local- or small-scale variability, even if their ability to do so can be improved with smaller grid size and higher observation density. The averaging time and sampling frequency of the measurements compared to those of the model output should also be considered. The relevant timescales for large-scale changes in atmospheric mercury are months to years. For simulation of trends, atmospheric models must be driven using time-varying inputs of both anthropogenic and legacy mercury to the atmosphere.

- Strengths: Bottom-up source attribution, large-scale spatial variability
- Weaknesses: Reliance on accuracy, temporal coverage, and availability of emission inventories
- Readiness: Multiple available models

Model Description 6.2: Ocean models

Global ocean models represent the marine fate of mercury deposited to the oceans from the atmosphere and entering the oceans via rivers. They require specification of the magnitude and spatial distribution of wet and dry deposition as well as river concentrations as inputs. These models simulate transport by ocean currents, mercury methylation, particle partitioning and sinking. Since ocean models directly calculate total seawater mercury and methylmercury concentrations, observations of these quantities are most comparable. In these comparisons, important considerations are comparing a near-instantaneous measurement with a longer time-averaged model value and comparing point measurements against model values representing a large area. Coastal and heavily river-influenced areas will be more sensitive to local releases via river inputs, while open-ocean measurements will be more sensitive to atmospheric inputs. The relevant timescales for ocean mercury are years to centuries. Simulation of trends of ocean mercury concentrations will require that inputs of riverine mercury releases as well as atmospheric deposition to ocean be time-varying.

- Strengths: air-sea exchange impacts, decadal time-scale changes
- Weaknesses: Propagates uncertainties in inputs from atmospheric models
- Readiness: Multiple available models

Model Description 6.3: Terrestrial models

Terrestrial models represent the exchange of mercury between atmosphere, vegetation, soil, water, sediment and groundwater reservoirs via processes associated with biochemical transformations, such as carbon processing by plants and soil microbes. These models use atmospheric mercury concentrations and deposition as inputs to calculate the plant uptake, throughfall, litterfall, and soil uptake of mercury, as well as soil evasion fluxes of mercury due to the microbial breakdown of mercury-containing carbon compounds in soils, and can include the transfer of mercury between soil sediments and water/groundwater due to soil leaching and erosion processes. The breakdown of mercury-containing compounds takes place over a wide range of time scales, meaning that terrestrial models account for mercury responses to changes ranging from seasonal to over centuries and longer. These models are useful for estimating legacy contributions to environmental mercury levels.

- Strengths: Legacy and environmental driver source attribution, long-time-scale influence
- Weaknesses: Reliance on historical input information, large amount of ancillary data required, lack of data on terrestrial ecosystems
- Readiness: Emerging applications for multi-media model coupling

6.2.3. Characterization of representative levels and spatial patterns

Total gaseous and elemental mercury concentrations from comparable measurements can be compared across monitoring sites and between types of monitoring sites to quantify spatial patterns in air mercury levels. Mapping these concentrations can show geographic patterns. Similarly, wet deposition measurements can be compared across sites and between site types. This type of analysis can be performed for measurement sites of all tiers.

Spatially resolved atmospheric models can estimate the level and form of mercury across a wide range of locations and times, including at locations and times not directly covered by monitoring. This model output can be used to supplement monitoring findings by filling the gaps between monitoring locations.

These models can also estimate how representative an observation is by quantifying the expected spatial and temporal variability in the observation's vicinity. By quantifying the representativeness of an observation, the models can improve its evaluative power. For example, in regions where spatial gradients are expected to be small according to models, a single observation site can effectively monitor a wide region. This means that models can also be used to inform monitoring locations, suggesting denser monitoring in more spatially variable areas. In regions and locations with significant local sources, finer-scale modelling could better estimate spatial variability.

Models can be used to extend the observed spatial patterns of mercury in one observed form or matrix to other forms or matrices, because they can take inputs consistent with the observations in one medium and simulate the resulting patterns in the medium/media they represent.

Subpopulation summary statistics from human biomonitoring can be used to establish baseline mercury levels and potentially broad spatial patterns depending on subpopulation locations. These

can include both mercury levels and guideline value-exceedance statistics. Comparison could be done across identified vulnerable subpopulations (with Tier 1 monitoring), or across national or other subpopulations (with Tier 2 or 3 monitoring).

A standardized method of visualization and summary analysis would facilitate communication of combined monitoring and modelling findings. The most common form of visual comparison for spatially resolved models with collections of observational data is a coloured map of modelled values with the corresponding observations of the same quantity overlaid as coloured dots using the same colour scale (e.g., Shah et al. 2021; Zhang et al. 2020). A standardized choice of colour scale and map projection would aid visual comparison between different models of the same type. An indication of the underlying model resolution in the form of a grid can aid visual interpretation of spatial variability. Colour maps should be chosen with consideration of viewers with colour vision deficiency, and be diverging for quantities that can be positive or negative. Overlaid hatching and special symbols can be used to annotate whether mapped trends are statistically significant.

Interactive web-based tools to support model data exploration and access could increase the reusability of model output. In the case of spatially resolved models, an online platform which allows for data selection and visualization, with subsetting by medium, quantity of interest, location(s), etc. would allow for maximum reuse by users for the purpose of smaller spatial-scale analysis.

Model Description 6.4: Generalized linear/additive models (GLM/GAM)

A generalized linear model (GLM) is a generalized version of linear regression which does not assume that the response variable error is normally distributed and does not assume that the response variable changes linearly with changing predictors. This added flexibility allows a GLM more explanatory power for quantities such as mercury concentrations in monitoring media which can have complicated responses to specific observable drivers. GLM can be further extended to generalized linear mixed models (GLMM), in which predictor variables additionally contain a random component to their effects, and generalized additive models (GAM), in which predictor variable coefficients are generalized to functions.

These types of models can be used with monitoring data to control for and attribute observed variability to specific independent variables. These observation-driven relationships to drivers of variability can be used as a “top-down” constraint for attribution. Because the valid range of these models is determined by their training data, monitoring data should share common comparable ancillary data across sites to most effectively implement this type of analysis. Separate training and testing data subsets should be used to avoid overfitting, and model assumptions should be checked by examining residuals.

- Strengths: Top-down attribution, application not specific to any given medium, monitoring-driven
- Weaknesses: Reliance on wide-ranging comparable data and ancillary data
- Readiness: Widely-used methodology

6.2.4. Estimation of information on source attribution

Models can not only calculate mercury levels and trends, but can also quantify the contributions to those values by specific drivers. Because emissions sources are direct inputs to atmospheric models, these models can be used to isolate emissions responses in observed and modelled trends in mercury concentrations and wet deposition (Tier 1) using a bottom-up approach. Such models can therefore be used along with observed trends to quantitatively attribute the trends to specific source types. This is true of types of sources, such as primary anthropogenic vs. legacy, as well as the relative importance of local sources vs. global sources using models that can resolve these types of sources individually.

Where ancillary air information is also available (Tier 2 and 3), a top-down approach can attribute observed trends to sources and drivers. At these locations, combining the monitoring-driven top-down approaches with bottom-up attribution from atmospheric models can balance explanatory and predictive power to provide more robust attribution estimates and a stronger weight of evidence than either method individually.

Biota monitoring and ancillary data (Tier 2 and 3) can be used in top-down modelling to estimate the contributions of different sources and large-scale drivers to biota mercury levels and trends. These sources and drivers can be further attributed by bottom-up modelling to different types of sources using a combination of watershed, mass balance, atmospheric, and/or food web models. The number of models/media required to attribute mercury levels and trends will vary from site to site and depend on the relative contributions of drivers in each medium. At intensive monitoring locations (Tier 3), top-down and bottom-up approaches can be combined to “calibrate” mechanistic model input parameters.

Quantifying the contribution of sources of natural and legacy mercury requires some level of multimedia approach. For single-medium models, inputs corresponding to these types of sources can be varied in a way that reflects the changes occurring in the source media. Coupled-media models can directly simulate the concurrent changes in legacy fluxes between media in a self-consistent manner while changing only primary releases as model inputs. For site-specific modelling, multimedia mass balance models present a tool for attribution that includes legacy sources.

The attribution of trends and changes in mercury levels in all monitoring media to environmental drivers unrelated to the Convention can also be performed using a top-down approach where the necessary ancillary data is available (Tier 2 and 3 sites), and can in some cases be performed using a bottom-up approach where those drivers can be explicitly changed in model scenarios. In the atmosphere, weather patterns and climate cycles can lead to variability in mercury levels and deposition through changes in temperature and precipitation. Atmospheric variability can translate to the surface ocean, and the ocean has analogous climate cycles that can affect observed trends. Terrestrial systems are strongly affected by land-use changes, and changes in the cryosphere can propagate effects to the atmosphere, aquatic and terrestrial environments. In biota, variability in temperatures and food web structures as through prey availability can cause changes in biota mercury levels unrelated to anthropogenic mercury emissions. These changes are unrelated to the Convention but can have impacts on observed trends, and quantifying their contribution allows a more accurate evaluation of Convention effectiveness. Variability in environmental drivers are especially relevant to site-specific and small spatial scale trends.

Changes in human biomarker levels can be attributed to drivers through exposure assessments using in-depth survey data and sophisticated biomarker analyses that include, for example, multiple biomarkers, mercury speciation analysis, and/or mercury stable isotopes (Tier 2 and 3). A top-down approach can identify contributions from changing dietary habits, occupational and other exposures

that can be estimated through the survey based on Tier 1 information. Attributions to measures influenced by the Convention can already be made at this level when considering the drivers for the changed behaviour in a careful and scientifically sound manner. When further adding ancillary monitoring and other information from Tier 2 and 3, including known dietary intake quantities due to biota mercury levels, even smaller responses can be attributed to behavioural changes influenced by the Convention or changing mercury concentrations in the diet

When adding bottom-up modelling to the above-described approach, improved explanatory power that includes even more factors influenced by the Convention can be obtained and the full pathway from source to impact can be modelled (e.g., Zhang et al. 2021). This comprehensive long-term goal of the Effectiveness Evaluation will require an accumulation of monitoring data and analyses to provide information with a high degree of confidence, but most of the attribution analysis steps will be able to provide useful information immediately based on Tier 1 data.

Model predictive and explanatory power

Mechanistic models share an overall structure whereby they are designed to simulate or represent a collection of interactive physical/biochemical processes involving mercury in one or more media and forms, and require inputs representing the flow of mercury into the scope or domain of the model as well as biogeochemical and physical environmental conditions. The mathematical representation of physical/chemical processes requires parameters such as rate constants, partition coefficients or similar experimentally measurable or computationally estimated values. The combination of these inputs, the representations of processes, and the model spatiotemporal resolution dictates the resulting model outputs. These models can have high explanatory power because of this structure and can directly relate changes in drivers to model output values in a bottom-up approach. The uncertainty in these models' output values is the accumulation of the types of uncertainty discussed in section 6.4. The model outputs are not necessarily the same quantity that monitoring efforts are measuring, but the two often overlap closely.

To conduct bottom-up analyses, estimates of primary anthropogenic emissions/releases of mercury are required as inputs for a variety of models in different media. While some inventories are currently available, they differ in methodology, represented time period, and release magnitudes. An updated, unified emission inventory which estimates both magnitudes of releases and their uncertainties would aid the Effectiveness Evaluation and provide more robust answers to questions of trend identification and attribution.

Statistical models can also be useful, especially in areas where the process-level understanding is insufficient to allow representation by mechanistic models, but where a cause-effect relationship between predictor variables and the quantity to be predicted can reasonably be justified with sound scientific explanations. When used together with mechanistic models, statistical models can be useful to determine if the process-level understanding is good enough. Such models require separate training and test data to avoid overfitting and careful determination of predictor variables to avoid confounding factors.

Statistical models trained on primary and ancillary monitoring data can have high predictive power within the range of the training (or input) data. They can identify and control for variations in drivers which can obscure an underlying mercury-specific signal, but lack the explanatory ability of process-based models. Inferring drivers based on observed statistical relationships (top-down approach) can be useful for attribution on its own, but can also be combined with bottom-up approaches to infer a best estimate which uses both observed quantities and prior estimates.

Multiple complementary models that can be used to answer the same questions should be employed together wherever possible. This can be accomplished using a Bayesian approach, meaning bottom-up analysis (representing process-level knowledge) provides an estimate independent from the monitoring data itself, and top-down analysis quantifies the evidence of those same estimates in the monitoring observations. By incorporating the quantified uncertainties of each model/analysis into this approach, a more robust estimate can be obtained. In many cases this can result in lower overall uncertainty in the quantitative answer to a given question by combining the higher predictive power of top-down approaches with the higher explanatory power of bottom-up approaches. This can be viewed as a way for statistical models to "calibrate" mechanistic models based on the observational findings from monitoring, in a way that is specific to a given question and uses all available information. This approach can be used for a single medium with multiple applicable models and/or to combine models across media.

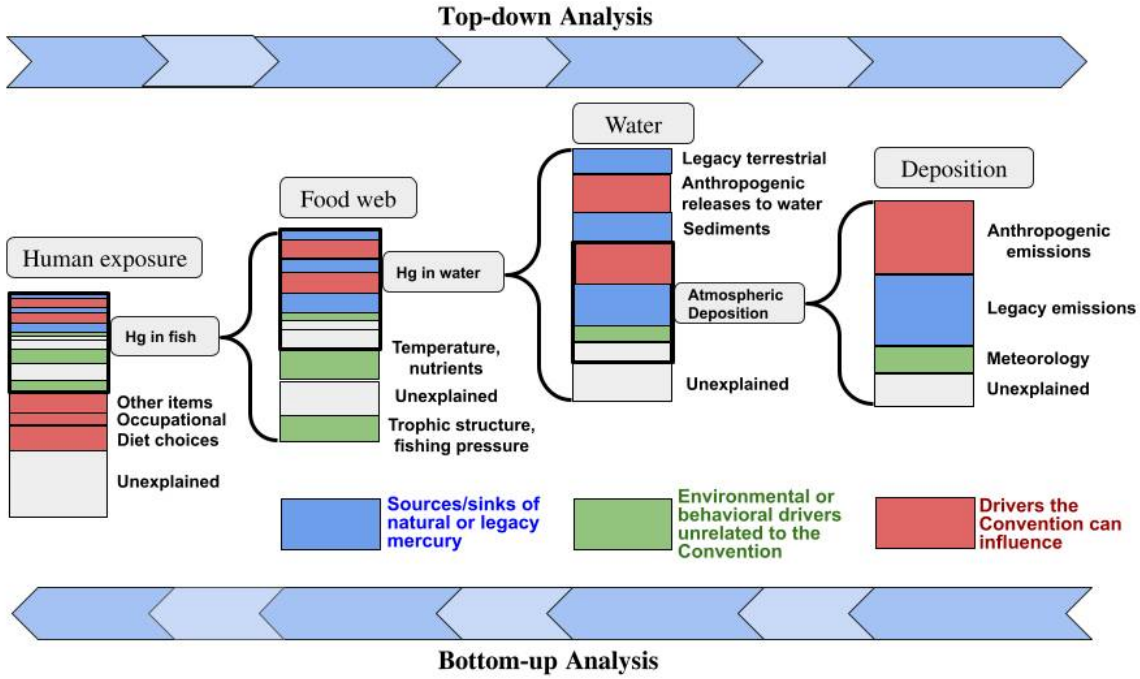

Figure 6.1. Illustration of attribution across media for hypothetical contributions of selected drivers at a hypothetical location. The coloured bars represent the fractional contributions of different drivers to observed mercury trends/variability in each medium. The drivers of variability/change in a given medium can in turn be attributable to drivers in other media (C. Thackray, unpublished).

Model Description 6.5: Mass balance models (also referred to as box models, compartment models, mass flow models)

Mass balance models represent the exchange of mercury between media, and are versatile tools which can be used on a range of spatial scales. These models use estimates of how quickly mercury is exchanged between media to self-consistently calculate mercury levels across a wide range of time scales, in a trade-off against spatially-resolved output. Inputs to these models are the releases of mercury into the model domain. Such models representing the global mercury cycle would take as inputs the total anthropogenic and geogenic releases of previously-lithospheric mercury, and represent its fate as it cycles through the atmosphere, terrestrial and ocean systems on decadal timescales and longer. In contrast, the same modelling approach could be applied to a specific location, with the inputs then being local releases of mercury as well as the transport of mercury from outside the model domain, and the model representing local mercury levels instead of global average

levels. Mass flow models can be used to evaluate local effectiveness over smaller regions by representing the processes and releases particular to those regions and using the contribution of global trends as an external input. An important consideration when comparing to point observations is the spatial aggregation implied by a single or few compartments representing each entire medium in this type of model.

- Strengths: Bottom-up attribution, consistency across timescales
- Weaknesses: Reliance on wide range of input quantities
- Readiness: Easily implemented where inputs are available

Model Description 6.6: Watershed Hg models

Watershed models combine mechanistic and empirical models that each capture the dynamics of a particular component of the local biogeochemistry to simulate mercury and methylmercury concentrations and fluxes (Golden and Knightes 2011; Knightes et al. 2014). This type of modelling is highly watershed-specific and relies on in-depth a priori knowledge of the watershed system of interest. The biogeochemical processes within the watershed contribute along with large-scale drivers such as thawing permafrost and land-use change to dictate the mercury response. Since understanding of the full collection of processes is incomplete and the local variability of the biogeochemical conditions are large, a range of ancillary parameters are therefore needed to enable statistical analysis of source-receptor relationships. This type of location-specific modelling is particularly important for sensitive environments such as the Arctic and for contaminated sites.

- Strengths: Characterize complex interactions of important processes
- Weaknesses: Intensive implementation, large uncertainty
- Readiness: Possible research implementation at intensive monitoring sites

Model Description 6.7: Food web and bioaccumulation models

Food web models represent the uptake of methylmercury to biota and the resulting bioaccumulation in freshwater and marine food webs. Inputs to these models are water concentrations of mercury and other chemical variables (e.g., DOC, TSS, pH), and parameters include water temperatures and bioenergetics parameters (and in some cases food web structures). Models either represent specific food webs and therefore simulate concentrations for species directly or simulate concentrations by trophic level. In both cases, measured tissue concentrations and trophic position are key for evaluating these models. Important comparability considerations include the age and size of sampled individuals and movement outside the represented domain (e.g., migration). The relevant timescales for food webs are years to a decade. These models can take local observations of marine or freshwater mercury levels and trends and translate them to fish concentrations to inform local exposure modelling.

- Strengths: Bottom-up attribution, specificity
- Weaknesses: Some parameters difficult to obtain (e.g., food web structure, water biogeochemistry), challenging to extend site-specific models to larger areas
- Readiness: Multiple site-specific models available. Further study is necessary to determine how current food-web model would be best used in the Effectiveness Evaluation processes

6.2.5. Estimation of exposure and adverse impacts

Possible sources and routes of human mercury exposure can be identified by regression-based approaches that can be used to relate mercury biomarkers to survey data. Survey data can give estimates of dietary intake, occupational practices, and other potential influences on exposure. Using human monitoring and survey data (Tier 2 and 3) in combination allows a top-down identification of exposure pathways for specific subpopulations. Bottom-up estimates of exposure can also be possible for certain subpopulations, using local air and/or biota monitoring data where occupational and diet information is available.

Risk assessment techniques can be used to estimate risk for populations potentially affected by variable exposure levels. Through probabilistic relationships between human biomarker mercury and adverse health effects, subpopulation burden of disease can be estimated, and extended to economic costs. With monitoring repeated on the same subpopulations over time, the changes in these expected health effects and their costs can be quantified and related to exposure pathways.

Model Description 6.8: Exposure and human health risk assessment models

This category encompasses a collection of models representing human exposure to mercury and the resulting health risks. Exposure models represent the intake of mercury by humans, and require mercury concentrations in diet items (e.g., freshwater fish, seafood, marine mammals, rice), in occupational practices (e.g., ASGM, dentistry), in certain products (e.g., skin-lightening creams, waste products), and the environment (e.g., soil). To mechanistically link exposure to mercury biomarker concentrations (i.e., levels in blood, urine, hair, and/or cord blood) in a given population (e.g., for a “bottom-up” analysis), toxicokinetic parameters describing human mercury metabolism can be used, with important uncertainties arising from differences in methylmercury uptake and elimination across individuals (Stern 2005). Regression-based models (including GAM/GLM) are also commonly used to relate human biomarker concentrations to exposure pathways in populations and subpopulations and can be used for "top-down" attributive analysis. Human biomarkers in populations may react to changes in mercury exposure in the timescale of days to months, with documented examples related to fish consumption advisories, amalgam removal, and occupational practices. Health impacts are often important for specific sub-populations, for example people who rely on local fish and marine mammals as a dominant protein source, and people with occupational exposures such as artisanal and small-scale gold miners. Health impacts are commonly modelled using statistical relationships, with acute and chronic responses to inorganic mercury exposure and longer-term impacts of dietary methylmercury exposure. These models can be applied at local/population scales to estimate effectiveness of changes in global mercury releases on limiting exposure and health impact, using observed biota mercury levels, or those from food web models, as inputs. Other site-specific applications are acute and chronic occupational exposures, such as at artisanal and small-scale gold mines.

- Strengths: Designed to interact with monitoring quantities, well-defined procedure
- Weaknesses: Potential recall bias in survey data needed to simulate exposure, inter-individual and -population variation in toxicokinetic parameters for mechanistic models
- Readiness: Well-established methodology

6.2.6. Quantification of key environmental processes

Top-down approaches using air monitoring and associated ancillary data (Tier 2 and 3) can estimate the contributions of specific environmental processes to observed mercury variability. Wide-ranging comparable ancillary data including land-use, air quality, and dry deposition parameters (Tier 2) and isotope measurements (Tier 3) can allow identification of their influence on observed mercury concentrations and wet deposition.

Large scale intercomparisons of monitoring measurements with bottom-up model output can also help identify key processes. Where monitoring shows inconsistency with mechanistic models, it indicates an area to better identify and quantify the important processes and their effects. Where the contribution of sources and sinks to levels of mercury is explicitly represented by mechanistic models, the observed levels, patterns, and trends can be used to infer changes in individual drivers.

In the atmosphere, oceans, and terrestrial system, observed spatial patterns of mercury and how they relate to environmental drivers can inform how modelling can best represent the physical and chemical processes that determine the transport and fate of mercury. Computational methods can also provide important contributions to improve process-level understanding. Theoretical computations of physical/chemical parameters add lines of evidence that are independent from monitoring or from experimentally determined values, further strengthening the total weight of evidence. Better representation of such processes will increase the applicable scope of the modelling and contribute to an iterative process where future modelling will better answer the Effectiveness Evaluation questions of interest.

6.3. Role of coupled-media modelling and analysis

Models or modelling frameworks that simulate multiple media and the flows of mercury between them in an internally consistent fashion are especially useful in light of the connections between media across a range of space and time scales. Each model discussed in this chapter represents the processes important to a specific medium, and these media are interconnected in a variety of ways. Some of these connections are effectively one-way, with one medium affecting another but not vice versa. In these situations, models can be chained together by using the output of a model for one medium as an input to a model for another. When models are chained in this way, longer simulations may be required to address the different retention times of different media.

On the other hand, some of the connections between media are effectively two-way, with both media affecting each other, possibly on different time scales. In these cases, coupled-media models which represent processes in both media in an internally consistent fashion are important for accurately attributing observed levels and trends to their drivers. The internal consistency can reduce uncertainty in situations where fewer of the possibilities for individual media are consistent across multiple media at once. The representation of coupling across multiple timescales means that these models can be more applicable for longer-term trends influenced by legacy mercury. The two-way coupling of existing single-medium models can be technically challenging, depending on the model specifics and time scales involved.

While the response of the atmosphere to changes in air emissions is relatively fast, on the order of months to years, the response in other media can be slower and lag behind those changes (see Figure 4.4). Moreover, the responses of the terrestrial and ocean systems feedback on the atmosphere, causing atmospheric trends to contain a signal contributed on these longer timescales. On the global scale, a decrease in anthropogenic emissions of mercury to air results in a fast atmospheric response proportional to the change in the total flux of mercury to air, which includes significant contributions from land and ocean legacy emissions. The immediate response of the

atmosphere is thereby dampened by the slower-equilibrating media. For example, declining atmospheric concentrations result in declining deposition to both land and oceans. This declining deposition leads to a decline in mercury levels in those media on longer time scales, which itself leads to less mercury evaded to the atmosphere and further declines in atmospheric mercury. Models which provide a coupled atmosphere-ocean-terrestrial simulation can be used for modelling trends of atmospheric, terrestrial, and ocean mercury concentrations simultaneously. This would be particularly useful for identification and attribution of trends influenced by legacy mercury.

Coupled-media models can help us to understand the implications of the trends we observe in air or other media for the eventual impacts on ecosystems and humans, which will be manifested over time. Observed decreases in air concentrations and deposition will likely contribute to decreased human exposure in the future. Even though we cannot yet observe those benefits, coupled-media models can be used to estimate them.

6.4. Model uncertainty

Some types of model uncertainty can be quantified, and represented as a distribution of the probability of specific values. This allows models to not only estimate quantities of interest, but also to provide a measure of confidence in those estimates. This is important for basing decisions and evaluations on model results, and for identifying which guiding questions (Chapter 2) have clear answers and which require further monitoring/analysis to answer with a given degree of confidence. Combining multiple modelling approaches with well-quantified uncertainties can reduce overall uncertainty by identifying areas of agreement.

The common structure of mechanistic models produces model output with three important categories of uncertainty that should be considered in model evaluation and interpretation of "bottom-up" estimates:

(a) uncertainty in the output which follows from the fact that the inputs used by the model are themselves uncertain (e.g., inventories of emissions/releases). This uncertainty can be estimated by testing a model using a range of available estimates of the inputs.

(b) uncertainty in the output which follows from the fact that the physical/chemical parameters used to represent different processes are uncertain (e.g., reaction rate coefficients and partitioning coefficients). This uncertainty can be estimated by testing a model using a range of parameter values within their uncertainty bounds.

(c) structural uncertainty due to the fact that there are processes and levels of mechanistic complexity that are not represented by the model due to incomplete knowledge about the drivers of the behaviour of modelled quantities. This type of uncertainty can be difficult to quantify because it potentially depends on unknown missing processes, but can be qualitatively assessed by experts with knowledge of the processes represented by the model.

Uncertainty for generalized linear and additive models can be estimated based on the standard error of predictions of observations in a cross-validation dataset. Confidence intervals can also be calculated using the posterior distribution (containing information from prior estimates and observational evidence) of the model parameters.

In addition to model uncertainty, the comparison of modelled and observed quantities requires consideration of uncertainty and variability in the observational data, and the uncertainty due to the comparison itself. The latter can be introduced through mismatches in the precise nature of the quantities compared, especially via spatial and temporal mismatches. The comparison of gridded model output with point observations introduces such uncertainty, because within the area of a

model grid cell some unresolved variability is to be expected. The magnitude of this uncertainty can be estimated if point observations at multiple locations within a single model cell are available, or by downscaling larger-scale variability in model output. Mismatches between model temporal resolution and observational sampling frequency and averaging time similarly need to be considered. All model outputs and observational data carry uncertainty, and the quantification of that uncertainty allows decision-making to be based not just on a given result, but also the degree of confidence in that result.

6.5. Model evaluation

Where possible, multiple applicable models can be used together for increased robustness rather than selecting a single model for a particular question. The evaluation of a model for use requires the determination of under what conditions and for what quantities/questions that model is applicable. The quantities for which the model is used should be directly calculated by the model, and the model should generate results that are consistent with directly comparable monitoring/observational data. In comparisons of models to monitoring mercury levels and trends, model-monitoring equality should not be the goal, but rather model-monitoring consistency. For quantities of interest, model and measured values are consistent if they are not statistically distinguishable from each other when accounting for the uncertainties in the model, the measurement, and the manipulation of each for the purpose of comparison. In order to draw conclusions from the applied model, the uncertainty in its results must be smaller than the magnitude of the result itself.

Model Evaluation Considerations

Applicability of a model estimating a given quantity required to answer a question of interest should be determined by:

1. Whether that quantity is estimated by the model directly, using relationships to input variables soundly based on available knowledge
2. Whether the modelled quantity is consistent with available comparable monitoring results
3. Whether the uncertainty in the modelled quantity is well-quantified and small enough to draw the conclusions necessary to answer the question of interest and/or provide a degree of confidence in that answer.

6.6. Summary of information from modelling

Table 6.3 summarizes what output models can produce to support the Effectiveness Evaluation. Model data formatted and managed for interoperability/harmonization with both monitoring data and other models, following the FAIR criteria described in chapter 2, would greatly facilitate both single-medium and multi-media analyses. A common, self-describing (not needing separate metadata) and open data format should be used for gridded model output so that data users can rely on a single set of free and open software tools for all shared model data. Shared model output should include quantities for comparison to monitoring as well as quantities that are common inputs to other models, such as fluxes across media boundaries, as well as metadata containing relevant information about the output and how it is generated. Examples of each type of model are included in the Supplementary Material (see Table S.4).

Table 6.3. Information from modelling data. For each model type, the primary model output is listed, along with the output appropriate for evaluation, metadata to accompany the model output, data for identifying model output locations, and model output to be collected for use by other types of models.

Model Type	Primary Output	Evaluation Output	Metadata	Location Data	Output For Other Models
Atmosphere	Air concentrations and temporal trends	Atmospheric concentrations, trends; wet deposition rates, trends; dry deposition to foliage/soils/ snowpack	Input sources; meteorological inputs; chemistry represented; boundary assumptions	Latitudes, longitudes, altitudes	Gross dry deposition of elemental and oxidized mercury to terrestrial and ocean locations; elemental mercury concentrations; source attribution quantification for outputs
Ocean	Seawater concentrations and temporal trends	Seawater mercury concentrations, temporal trends	Input sources; circulation source; processes represented;	Latitudes, longitudes, depths	Gross evasion fluxes to air or seawater surface elemental mercury concentration; seawater methylmercury concentrations
Terrestrial	Soil/vegetation mercury levels and temporal trends	Soil/vegetation mercury reservoirs, trends; soil-air fluxes, temporal trends	Input sources; meteorological/ climate inputs; processes represented;	Latitudes, longitudes	Gross evasion fluxes to air
Watershed	Water mercury and methylmercury concentrations and temporal trends	Freshwater mercury levels, temporal trends	Input sources; biogeochemical conditions; land-use	geolocation or represented watershed	Water mercury and methylmercury concentrations
Food web	Tissue mercury concentrations	Tissue mercury concentrations, temporal trends	Input sources; food web structure/ feeding parametrizations; bioenergetics parameters	Geolocation or represented region	Tissue mercury concentrations
Human Exposure and pathways	Population mercury intake	Human mercury biomarker concentrations, temporal trends	Input sources; diet assumptions; population parameters	Geolocation or represented city/country/ region	Mercury intake; human biomarker concentrations
Mass flow	Bulk mercury	Media-averaged	Input sources;	Represented	Time-evolution of

models	levels across media and their temporal trends	mercury levels, temporal trends	rates/timescales represented	spatial extent	mercury levels across media and fluxes between media
Integrated or coupled-media models	Concentrations across media; fluxes between media	media-specific evaluation; levels and trends of interrelations between media, and fluxes between media	input sources; media-specific modelling methods; coupling methods	Geolocations or represented spatial extent	Time-evolution of mercury levels across media and fluxes between media

6.7. Conclusions

By analysing monitoring data, temporal and spatial trends in the levels of mercury in specific environmental media or human matrices can be derived with confidence intervals. These trends provide a first-level indication of whether the Convention may be contributing to protecting human health and the environment from the adverse effects of mercury. Analyses of the monitoring data collected in each medium separately will be informative, and these monitoring data can also be used in an integrated manner, where combining multiple complementary analysis approaches to answer the same question will improve robustness and increase the scientific weight of evidence. Both for model evaluation and for analyses, model output uncertainties should be quantified. Any generated estimate should provide a detailed discussion/presentation of the associated uncertainty and factors that have determined this uncertainty.

In many cases, attribution of observed trends to specific drivers can be performed through the use of models which resolve the intervening processes, supplemented by empirical statistical approaches. Cross-media analysis involving both mechanistic and statistical modelling in all relevant media is important in order to fully evaluate effectiveness of the Convention. This evaluation requires separating the impacts attributable to the Convention from changes that occur due to other factors, and while monitoring data shows the impact of all of these factors, modelling can help attribute the changes to the different drivers. As more monitoring data and analysis tools become available, more detailed analysis can be performed.

To estimate background and impacted levels of mercury, simple analyses can be conducted on monitoring data at sites chosen for this purpose. Temporal trends can be identified at these and other locations once a long enough time record has been collected. This trend analysis should account for variability and uncertainty to obtain trend magnitudes, confidence intervals for the trends, and measures of the trends' statistical significance.

To characterize spatial patterns, several atmospheric chemical transport models can be used, supplemented with statistical models where beneficial to quantify representativeness of observed levels and trends in air, and to extrapolate ambient air concentrations and wet deposition to areas with sparse monitoring data. Spatially resolved models in air and other media can be used to interpolate levels and trends of mercury while accounting for the drivers of spatial and temporal differences.

Bottom-up analyses can be performed with atmospheric models for source attribution, and top-down analyses with GLM/GAM for air and biota attribution where sufficient ancillary data is available. Top-down analysis of changes in exposure pathways can also be performed to attribute

changes in human biomarkers to measures influenced by the Convention. At intensive monitoring sites, combined top-down and bottom-up attribution analyses can be performed for air, biota and human biomarkers. To quantify legacy impacts, coupled-media approaches should be used where possible.

Exposure can be estimated based on specific sources and exposure attribution information can be used to estimate marginal health impacts/costs of individual drivers. Trends in risk associated with trends in exposure and/or biomarker benchmark values can be estimated where the appropriate information is available.

The quantification of key environmental processes can improve our understanding of cause-effect relationships. Top-down analysis can be used to identify key environmental drivers, and large-scale measurement/model intercomparisons can be performed to identify key processes. Improved understanding can lead to a beneficial iterative approach: using the available information to improve the application of models can decrease the uncertainty for further and future analyses and evaluations.

Annex. Tiered Approach to Monitoring Mercury and Mercury Compounds

A.1. Introduction

To support Parties and organizations who may wish to develop new monitoring programs, or improve existing ones, with a view to contributing to the Effectiveness Evaluation, this document identifies a tiered approach for monitoring each of the three media (air, biota, humans):[24]

Tier 1 is intended to provide guidance on mercury monitoring under a limited set parameters for circumstances where available resources are not sufficient to implement the actions in Tier 2. Following guidance by the COP,[25] the methods in Tier 1 are cost effective, practical, feasible, and sustainable. The Tier 1 methods are intended to provide information that are useful in identifying and characterizing gaps and needs of national, regional, or local interest and to provide information that is useful to the collective effort for the Effectiveness Evaluation. While the implementation of Tier 1 actions may not fully address the questions in Table 2.1, it will contribute essential information and create a foundation for Tier 2 monitoring.

Tier 2 is intended to build upon Tier 1 methods to provide information that will address the questions identified in Table 2.1, and to create a basis for assessing source attribution at the local, national, and global scales (Figure 2.2). The methods and approaches in this tier may be more expensive or complex than those under Tier 1. The more comparable data from Tier 2 becomes available, the more robust the Effectiveness Evaluation will be.

Tier 3 identifies research methods and approaches that may play a vital role in supporting the Tier 1 and Tier 2 programs and the Effectiveness Evaluation, primarily by improving our understanding of key processes that link sources to environmental concentrations and exposures. Because Tier 3 focuses on processes, the results would likely yield insights that are broadly applicable and that should be taken into consideration in the Effectiveness Evaluation when available.

An overview of a proposed tiered approach for each matrix (air, biota and Human) is shown below.

[24] It is noted that the Convention does not impose any obligation upon Parties to conduct monitoring. As such, the tiered approach and any other activities or recommendations contained in this guidance are voluntary and presented with the sole purpose of supporting Parties who may wish to develop new monitoring programs, or improve existing ones, with a view to contributing to the Effectiveness Evaluation.

[25] Decision MC-2/10 pursuant to the terms of reference to Ad-hoc Technical Expert Group on Effectiveness Evaluation.

A.2. Atmospheric mercury monitoring

Hg Measurement	Metadata/Ancillary Measurements	Location/Spatial Distribution	Frequency	Contribution to information categories[26]	Modelling/Analysis[27]
TIER 1					
Total Gaseous Mercury or Gaseous Elemental Mercury (a range of methods may be used depending on objectives, resources, and other constraints): • automated active analysers (e.g. CV-AFS with pure gold traps) • manual active methods • passive samplers Wet Deposition, i.e., total mercury in precipitation, to the extent resources and other constraints allow: • sampler approved for use in an existing network.	• Location (latitude, longitude, elevation) • Meteorological data (where available, may be from nearby sites, including precipitation, wind direction and speed, air temperature, relative humidity, and solar radiation) • Proximity to known point sources (urban/regional/ background)	Sites should be selected in a mix of locations that include a) remote, background, b) rural, regionally representative, and c) source impacted locations (urban, industrial). Siting strategies may differ if the methods deployed are only active, only passive, or a mix of active and passive. Deploying a mix of active and passive samplers may maximize the amount of information collected given resource, infrastructure, or personnel constraints. Where possible, measurements should be collocated with other types of air quality and mercury measurements.	Varies by method: • automated active methods provide continuous sampling, often reported as hourly averages; • manual active methods provide daily or weekly integrated samples; • passive samplers integrate over 1-3 months; • wet deposition is typically collected as 7-day weekly samples.	(1) Baselines (2) Temporal Trends (3) Spatial Patterns (broadly) (5) Estimates of Exposure and Adverse Effects (initial)	• Temporal trends • Atmospheric model evaluation (for GEM) • Spatial variations • Input for local-scale modelling • Back-trajectory analysis • Bottom-up attribution analysis (from drivers and emissions)

[26] See chapter 2.

[27] See chapter 6.

Hg Measurement	Metadata/Ancillary Measurements	Location/Spatial Distribution	Frequency	Contribution to information categories[28]	Modelling/Analysis[29]
TIER 2 (adds to TIER 1)					
Speciated Reactive Mercury: • high resolution measurements of PBM, GOM using existing network SOPs Dry deposition of mercury: • Total Hg and MeHg in litterfall and throughfall (in select forest ecosystems).	• Emission inventories • Land cover and land use • Leaf area index • Atmospheric deposition of sulfate • Air quality tracers (including SO_2, CO, O_3, $PM_{2.5}$,)	Expect a few sites in each world region, surrounded by a cluster of Tier 1 sites. Sites should be a mix of a) remote, background; b) regionally representative; and c) source impacted locations and collocated with other network sites with more robust infrastructure.	Varies by method; high temporal resolution for speciation.	(1) Baselines (2) Temporal Trends (3) Spatial Distribution (4) Source Attribution (5) Estimating Exposure and Adverse Effects (6) Key Environmental Processes	• Estimate air-ocean and air-terrestrial mercury exchange • Covariate profiling • Top-down attribution analysis (from observations)
TIER 3 (adds to TIER 2)					
Mercury Isotopes: • e.g., multi-collector inductively coupled plasma mass spectrometry (MC-ICP-MS) Additional speciation methods • e.g.,cation exchange membranes. Applications of Tier 1 and Tier 2 methods in intensive research contexts to support process understanding	• Halogen and other oxidant concentrations	Expected to be opportunistic siting, collocated at long-term monitoring and research sites. Aircraft campaigns, ocean surveys, flux towers, etc.	High temporal resolution observations are often needed to characterize key processes.	(2) Temporal Trends (4) Source Attribution (5) Estimating Exposure and Adverse Effects (6) Key Environmental Processes	• Combined “top-down” and “bottom-up” attribution analyses • Isotopic fingerprinting

A.3. Biota monitoring

[28] See chapter 2.

[29] See chapter 6.

Hg Measurement	Metadata/Ancillary Measurements	Location/Spatial Distribution	Frequency	Contribution to information categories[30]	Modelling/Analysis[31]
TIER 1					
Total Hg in muscle, blood, egg, or keratin tissue of monitored fish or birds. Species selected for monitoring should have, where possible, a relatively consistent diet (and thus a narrow trophic range) that can be observed consistently over time at a given location. Trophic level 3 and 4 species are used in a number of existing programs and are a reasonable starting point.	• Location (Latitude/Longitude)depth if applicable • Species Name • Body Length and Weight • Age, Sex and Maturity Stage • Tissue type • Foraging ecology (diet) • Habitat description (e.g., size of lake, elevation, landcover and use, pollution history, water level changes, river flow and speed, or coverage of mangroves or coral reefs)	It is most important to make consistent observations at fixed locations over a long period. A mixture of background sites and locally impacted sites is recommended. With sufficient prior information, sites with well-known impact history should be chosen. These sites should be classified according to mappable site characteristics. Where little or no prior information exists, mapping of the overlay of ecosystem characteristics and mercury sources may be useful.	Annual measurements, with a consistent sampling season over time for each core fixed site.	(1) Baselines (2) Temporal Trends (5) Estimating Exposure and Adverse Effects	• Temporal trends. Note: Individual sample data are most useful for analysis rather than aggregated values. • Spatial variations, broad • Input for local exposure modelling • Guideline value exceedance statistics

[30] See chapter 2.

[31] See chapter 6.

Hg Measurement	Metadata/Ancillary Measurements	Location/Spatial Distribution	Frequency	Contribution to information categories[32]	Modelling/Analysis[33]
TIER 2 (adds to TIER 1)					
A set of focal taxa (fish or birds) would be sampled in different sites over time. While it is important to consistently sample similar taxa across locations within a region, if that is not possible, sampling several taxa in the multiple sites would help in accounting for species effects statistically. Note, monitoring novel species that have not been previously monitored elsewhere would be less informative for the Effectiveness Evaluation; but threatened species may be of more interest for national or global interests. Trophic level 3 or 4 species are preferred.	In biota: • carbon ($\delta^{13}C$) & nitrogen ($\delta^{15}N$) stable isotopes In water: • DOM/DOC/TOC, TSS, salinity, DO • (pH), N and P • phytopigments (chlorophyl-*a*) In sediment: • THg and TOC In air: • GEM • wet deposition • meteorological data Description of local hydrologic catchment.	Sites added in this tier would be sampled to cover a wider range of landscapes and geochemical characteristics. The additional sites may be selected, for example, according to the habitat type and then either rotated or randomly sampled within each habitat type. If the data sets from additional locations are paired with those from fixed sites monitoring similar covariates over time, the combined data sets will inform each other and contribute to source attribution. If possible, air and deposition measurements should also be carried out for the same sites.	Yearly monitoring rotating across sites added at Tier 2 (in such a manner that each particular site would only be monitored every few years).	(1) Baselines (2) Temporal Trends (3) Spatial Distribution (4) Source Attribution (5) Estimating Exposure and Adverse Effects (6) Key Environmental Processes	• "Top-down" biota mercury attribution • Watershed and food web model evaluation

[32] See chapter 2.

[33] See chapter 6.

Hg Measurement	Metadata/Ancillary Measurements	Location/Spatial Distribution	Frequency	Contribution to information categories[34]	Modelling/Analysis[35]
TIER 3 (adds to TIER 2)					
Sampling as above, but consideration may be given to all species (e.g., fish, sea turtles, birds, and marine mammals) and even lower trophic level taxa. Species at lower trophic levels may provide useful information to attribution of changes as they are more likely to respond more quickly to changes in Hg exposure and show changes earlier.	• Speciated Mercury, Mercury (δ^{202}Hg and δ^{199}Hg) stable isotopes in biota and suspected source-matrices of interest • Other chemical tracers related to known drivers (i.e., changes in CO_2 levels and water temperature in oceans due to climate change, co-tracers from ASGM activity, etc.) • Information on diet (e.g., fatty acids) • Stable isotopes of lower foodweb organisms (or compound specific stable isotopes of amino acids in fish) • Food web structure • Movement patters of focal taxa	Intensively monitor selected areas (e.g., catchments and coastal areas), with a primary site (supersite) for co-located measurements and secondary (or satellite) sites to capture variability across the catchment. Catchments or coastal selected for this strategy may be either background locations (mostly influenced by long range transport) or locally impacted locations (that are likely to see changes due to mitigation efforts).	Sampling may be more frequent than annual.	(1) Baselines (2) Temporal Trends (3) Spatial Distribution (4) Source Attribution (5) Estimating Exposure and Adverse Effects (6) Key Environmental Processes	• Combined "top-down" and "bottom-up" biota mercury attribution • Isotopic fingerprinting.

[34] See chapter 2.

[35] See chapter 6.

A.4. Human biomonitoring

Hg Measurement	Metadata/Ancillary Measurements	Location/Spatial Distribution	Frequency	Contribution to information categories[36]	Modeling/Analysis[37]
TIER 1					
Blood, urine, or hair THg depending on sampled population. Essential data for mercury values include count (n), percentiles including 10th, 25th, 50th, 75th, 90th, and 95th values; additional measures of central tendency (variance) including mean (SD) and geometric mean (95% CI); indication of data normality.	WHO Survey or HBM4EU Instruments. Relevant survey information (e.g., dietary, occupational, sociodemographic), where possible. Sample population characteristics (population type, sample size, age, sex, education, socioeconomic status, personal amalgams, city/region/country, day/month/year), analytical measurements (sample size, detection limit, accuracy, precision, and use of reference materials). Ethics board that approved work.	Vulnerable sub-populations should be identified based on exposure or risk that is most critical for them (i.e., dietary exposures, occupational groups, or high risk lifestage (e.g., pregnant women)).	Every 2-5 years for the same population, with monitoring activities staggered for different populations in different years. Timing of sampling should take into account possible seasonal changes in exposure.	(1) Baselines (2) Temporal Trends (3) Spatial Patterns (5) Estimating Exposure and Adverse Effects	Data should be aggregated for the entire sample population as a primary level summary, as well as for key sub-groups (e.g., different lifestages, sexes, locations, occupational categories) as part of a secondary level summary. • Spatial variations • Temporal trends • Exposure model evaluation • Input for local health impact / risk assessment modelling • Guideline value exceedance statistics • “Top-down” exposure attribution

[36] See chapter 2.

[37] See chapter 6.

Hg Measurement	Metadata/Ancillary Measurements	Location/Spatial Distribution	Frequency	Contribution to information categories[38]	Modelling/Analysis[39]
TIER 2 (adds to TIER 1)					
Blood/cord blood, urine, and/or hair THg depending on sampled population and survey. Methyl mercury and isotopes may also be considered.	WHO Survey or HBM4EU Instruments, or incorporation of Hg sampling into other health surveys or cohort studies. Relevant survey information (e.g., dietary, occupational, sociodemographic), and where possible coordinated measures in air and/or biota.	Two strategies: 1) Perform more in-depth analysis of sub-populations with high-exposure or classified as a vulnerable lifestage; 2) Incorporation of Hg sampling into other, in-depth health surveys or cohort studies.	Same as above.	(1) Baselines (2) Temporal Trends (3) Spatial Patterns (4) Source Attribution (5) Estimating Exposure and Adverse Effects (6) Key Environmental Processes	• Isotopic fingerprinting
TIER 3 (adds to TIER 2)					
Same as above for Tier 2	WHO Survey or HBM4EU Instruments or National/Regional population survey instruments. Relevant survey information (e.g., dietary, occupational, sociodemographic), and where possible coordinated measures in air and/or biota.	Two strategies: 1) National population survey (ideally leveraging other surveys/samples, and inclusion of vulnerable sub-groups) 2) Sampling of sub-populations with coordinated air and biota sampling.	Same as above.	(1) Baselines (2) Temporal Trends (3) Spatial Patterns (4) Source Attribution (5) Estimating Exposure and Adverse Effects (6) Key Environmental Processes.	• Combined “Top-down” and “bottom-up” exposure attribution

[38] See chapter 2.

[39] See chapter 6.

References

Aas W, Bohlin-Nizzetto P (2019) Heavy metals and POP measurements, 2017, EMEP/CCC- Report 3. Norwegian Institute for Air Research.

Abeysinghe KS, Qiu G, Goodale E, Anderson CW, Bishop K, Evers DC, Goodale MW, Hintelmann H, Liu S, Mammides C, Quan RC (2017) Mercury flow through an Asian rice-based food web. Environmental Pollution 229: 219-228.

Ackerman JT, CA Eagles-Smith, MP Herzog, CA Hartman, SH Peterson, DC Evers, AK Jackson, JE Elliott, SS Vander Pol, and CE Bryan (2016) Avian mercury exposure and toxicological risk across western North America: A synthesis. Science of the Total Environment 568: 749-769.

Ackerman JT, Herzog MP, Schwarzbach SE (2013) Methylmercury is the predominant form of mercury in bird eggs: a synthesis. Environ. Sci. Technol. 47: 2052-2060. DOI: 10.1021/es304385y.

Albert C, Helgason HH, Brault-Favrou M, Robertson GJ, Descamps S, Amélineau F, Danielsen J, Dietz R, Elliott K, Erikstad KE, Eulaers I (2021) Seasonal variation of mercury contamination in Arctic seabirds: a pan-arctic assessment. Science of the Total Environment 750: 142201.

AMAP (2011) AMAP Assessment 2011: Mercury in the Arctic. Arctic Monitoring and Assessment Programme (AMA). Oslo, Norway.

AMAP (2021) Mercury in the Arctic. Arctic Monitoring and Assessment Programme (AMAP). Oslo, Norway.

Amos HM, Jacob DJ, Holmes C, Fisher JA, Wang Q, Yantosca RM, Corbitt ES, Galarneau E, Rutter A, Gustin M, Steffen A, Schauer J, Graydon JA, St. Louis VL, Talbot R, Edgerton E, Zhang Y, Sunderland EM (2012) Gas-particle partitioning of atmospheric Hg (II) and its effect on global mercury deposition. Atmos. Chem. Phys. 12: 591-603.

Amos HM, Jacob DJ, Kocman D, Horowitz HM, Zhang Y, Dutkiewicz S, Horvat M, Corbitt ES, Krabbenhoft DP, Sunderland EM (2014) Global biogeochemical implications of mercury discharges from rivers and sediment burial. Environmental Science & Technology 48: 9514-9522.

Angot H, Barret M, Magand O, Ramonet M, Dommergue A (2014) A 2-year record of atmospheric mercury species at a background Southern Hemisphere station on Amsterdam Island. Atmos. Chem. Phys. 14: 11461-11473.

Angot H, Dastoor A, de Simone F, Dastoor A, De Simone F, Gårdfeldt K et al. (2016) Chemical cycling and deposition of atmospheric mercury in polar regions: review of recent measurements and comparison with models. Atmos. Chem. Phys. 16: 10735-10763. DOI: 10.5194/acp-16-10735-2016.

Arnold RD and Wade JP (2015) A definition of systems thinking: A systems approach. Procedia Computer Science 44: 669-678.

ATSDR (1999) Toxicological Profile for Mercury (Update). Agency for Toxic Substances and Disease Registry, Atlanta, Georgia 617 (March).

Azad AM, Frantzen S, Bank MS, Johnsen IA, Tessier E, Amouroux D, Madsen L, Maage A (2019) Spatial distribution of mercury in seawater, sediment, and seafood from the Hardangerfjord ecosystem, Norway. Science of The Total Environment 667: 622-637. DOI: 10.1016/j.scitotenv.2019.02.352.

Basu N, Horvat M, Evers DC, Zastenskaya I, Weihe P, Tempowski J (2018) A State-of-the-Science Review of Mercury Biomarkers in Human Populations Worldwide between 2000 and 2018. Environmental Health Perspectives 126: 1-14. DOI: 10.1289/EHP3904.

Bjerregaard P, Schmidt TG, Mose MP (2020) Elevated mercury concentrations in biota despite reduced sediment concentrations in a contaminated coastal area, Harboøre Tange, Denmark, Environmental Pollution 260: 113985.

Bloom NS (1992) On the Chemical Form of Mercury in Edible Fish and Marine Invertebrate Tissue. Canadian Journal of Fisheries and Aquatic Sciences. DOI: 10.1139/f92-113.

Blum JD, Bergquist BA (2007) Reporting of variations in the natural isotopic composition of mercury. Anal. Bioanal. Chem. 388: 353-359.

Blum JD, Sherman LS, Johnson MW (2014) Mercury isotopes in earth and environmental sciences. Annu. Rev. Earth Planet. Sci. 42: 249-269. DOI: 10.1146/annurev-earth-050212-124107.

Braaten HF, de Wit HA, Larssen T, Poste AE (2018) Mercury in fish from Norwegian lakes: The complex influence of aqueous organic carbon. Science of The Total Environment 2018: 341-348.

Braaten HFV, Akerblom S, Kahilainen KK, Rask M, Vuorenmaa J, Mannio J, Malinen T, Lidersen E, Poste AE, Arnundsen PA, Kashulin N, Kashulina T, Terentyev P, Christensen G, de Wit HA (2019) Improved Environmental Status: 50 Years of Declining Fish Mercury Levels in Boreal and Subarctic Fennoscandia. Environ. Sci. Technol. 53: 1834-1843.

Braune B, Chételat J, Amyot M, Brown T, Clayden M, Evans M, Fisk A, Gaden A, Girard C, Hare A, Kirk J (2015) Mercury in the marine environment of the Canadian Arctic: Review of recent findings. Science of the Total Environment 509: 67-90.

Braune BM, Gaston AJ, Mallory ML (2016) Temporal trends of mercury in eggs of five sympatrically breeding seabird species in the Canadian Arctic. Environmental Pollution 214: 124-131.

Broadley HJ, Cottingham KL, Baer NA, Weathers KC, Ewing HA, Chaves-Ulloa R, Chickering J, Wilson AM, Shrestha J, Chen CY (2019) Factors affecting MeHg bioaccumulation in stream biota: the role of dissolved organic carbon and diet. Ecotoxicology 28: 949-963.

Brooks SB, Saiz-Lopez A, Skov H, Lindberg SE, Plane JM, Goodsite ME (2006) The mass balance of mercury in the springtime arctic environment. Geophys. Res. Lett. 33: L13812.

Brown RJC, Pirrone N, Van Hoek C, Horvat M, Kotnik J, Wangberg I, Corns WT, Bieber E, Sprovieri F (2010) Standardisation of a European measurement method for the determination of mercury in deposition: Results of the field trial campaign and determination of a measurement uncertainty and working range. Accreditation and quality assurance 15: 359-366.

Brunke E-G, Walters C, Mkololo T, Martin L, Labuschagne C, Silwana B, Slemr F, Weigelt A, Ebinghaus R, Somerset V (2016) Mercury in the atmosphere and in rainwater at Cape Point, South Africa. Atmos. Environ. 125: 24-32.

Burger J, Gochfeld M (1993) Heavy metal and selenium levels in feathers of young egrets and herons from Hong Kong and Szechuan, China. Arch. Environ. Contam. Toxicol. 25: 322–327. DOI: 10.1007/BF00210724.

Carvan MJ, Kalluvila TA, Klingler RH, Larson JK, Pickens M, Mora-Zamorano FX, Connaughton VP, Sadler-Riggleman I, et al. (2017) Mercury-induced epigenetic transgenerational inheritance of abnormal neurobehavior is correlated with sperm epimutations in zebrafish. PLoS ONE 12: e0176155.

Cerveny D, Roje S, Turek J, Randak T (2016) Fish fin-clips as a non-lethal approach for biomonitoring of mercury contamination in aquatic environments and human health risk assessment. Chemosphere 163: 290-295. DOI: 10.1016/j.chemosphere.2016.08.045.

Chen CY, Borsuk ME, Bugge DM, Hollweg T, Balcom PH, Ward DM, Williams J, Mason RP (2014) Benthic and pelagic pathways of methylmercury bioaccumulation in estuarine food webs of the Northeast United States Plos One 9: 1-11.

Chételat J, Richardson MC, MacMillan GA, Amyot M, Poulain AJ (2018) Ratio of methylmercury to dissolved organic carbon in water explains methylmercury bioaccumulation across a latitudinal gradient from North-Temperate to Arctic Lakes. Environmental Science and Technology 52: 79-88.

CIOMS (2016) International Ethical Guidelines for Health-Related Research Involving Humans, 4th Edition. Council for International Organizations of Medical Sciences.

Cizdziel JV, Hinners TA, Heithmar EM (2002) Determination of Total Mercury in Fish Tissues using Combustion Atomic Absorption Spectrometry with Gold Amalgamation. Water, Air, & Soil Pollution 135: 355–370. DOI: 10.1023/A:1014798012212.

Cole AS, Steffen A, Eckley CS, Narayan J, Pilote M, Tordon R, Graydon JA, St. Louis VL, Xu X, Branfireun BA (2014) A survey of mercury in air and precipitation across Canada: Patterns and trends. Atmosphere 5: 635-668.

Cransveld A, Amouroux D, Tessier E, Koutrakis E, Ozturk AA, Bettoso N, Mieiro CL, Bérail S, Barre JP, Sturaro N, Schnitzler J (2017) Mercury stable isotopes discriminate different populations of European seabass and trace potential Hg sources around Europe. Environmental Science & Technology 51: 12219-12228.

Cristol D, Evers DC (2020) The impact of mercury on North American songbirds: effects, trends, and predictive factors. Ecotoxicology 29: 1107-1116. DOI: 10.1007/s10646-020-02280-7.

Day B, Langlois R, Tremblay M, Knoppers BM (2007) Canadian Health Measures Survey: Ethical, Legal and Social Issues. Health Reports / Statistics Canada, Canadian Centre for Health Information = Rapports Sur La Santé / Statistique Canada, Centre Canadien d'information Sur La Santé 18 Suppl (82): 37-51.

De Simone F, Cinnirella S, Gencarelli CN, Hedgecock IM, Pirrone N (2016) A modeling comparison of mercury deposition from current anthropogenic mercury emission inventories. Environ. Sci. Technol. 50: 5154-5162.

De Simone F, Gencarelli CN, Hedgecock IM, Pirrone N (2014) Global atmospheric cycle of mercury: a model study on the impact of oxidation mechanisms. Environ. Sci. Pollut. Res. 21: 4110-4123.

Depew DC, Basu N, Burgess NM, Campbell LM, Devlin EW, Drevnick PE, Hammerschmidt CR, Murphy CA, Sandheinrich MB, Wiener JG (2012a) Toxicity of dietary methylmercury to fish: derivation of ecologically meaningful threshold concentrations. Environmental Toxicology and Chemistry 31: 1536-1547.

Depew DC, Basu N, Burgess NM, Campbell LM, Evers DC, Grasman KA, Scheuhammer AM (2012b) Derivation of screening benchmarks for dietary methylmercury exposure

for the common loon (Gavia immer): rationale for use in ecological risk assessment. Environmental Toxicology and Chemistry 31: 2399-2407.

Depew DC, Burgess NM, Anderson MR, Baker R, Bhavsar SP, Bodaly RA, Eckley CS, Evans MS, Gantner N, Graydon JA, Jacobs K (2013) An overview of mercury concentrations in freshwater fish species: a national fish mercury dataset for Canada. Canadian Journal of Fisheries and Aquatic Sciences 70: 436-451.

Dietz R, Letcher RJ, Desforges JP et al. (2019) Current state of knowledge on biological effects from contaminants on arctic wildlife and fish. Sci. Total Environ. 696: 133792.

Dietz R, Mosbech A, Flora J, Eulaers I (2018) Interactions of climate, socio-economics, and global mercury pollution in the North Water. Ambio 47: 281-295.

Dietz R, Sonne C, Basu N, Braune B, O'Hara T, Letcher RJ, Scheuhammer T, Andersen M, Andreasen C, Andriashek D, Asmund G (2013) What are the toxicological effects of mercury in Arctic biota? Science of the Total Environment 443: 775-790.

Dolgova S, Crump D, Porter E, Williams K, Hebert CE (2018) Stage of development affects dry weight mercury concentrations in bird eggs: Laboratory evidence and adjustment method. Environ. Toxicol. Chem. 37: 1168-1174. DOI: 10.1002/etc.4066.

Dommergue A, Sprovieri F, Pirrone N, Ebinghaus R, Brooks S (2010) Overview of mercury measurements in the Antarctic troposphere, Atmos Chem and Phys. 10: 3309-3319.

Drevnick PE, Brooks BA (2017) Mercury in tunas and blue marlin in the North Pacific Ocean. Environmental Toxicology and Chemistry 36: 1365-1374.

Drevnick PE, Lamborg CH, Horgan MJ (2015) Increase in mercury in Pacific yellowfin tuna. Environmental Toxicology and Chemistry 34: 931-934.

Driscoll CT, Mason RP, Chan HM, Jacob DJ, Pirrone N (2013) Mercury as a global pollutant: sources, pathways and effects. Environ. Sci. Technol. 47: 4967-4983.

Eagles-Smith CA, Ackerman JT, Willacker JJ, Tate MT, Lutz MA, Fleck JA, Stewart AR, Wiener JG, Evers DC, Lepak JM, Davis JA (2016a) Spatial and temporal patterns of mercury concentrations in freshwater fish across the Western United States and Canada. Science of the Total Environment 568: 1171-1184.

Eagles-Smith CA, Silbergeld EK, Basu N, Bustamante P, Diaz-Barriga F, Hopkins WA, Kidd KA, Nyland JF (2017) Intrinsic and Extrinsic Modulators Mercury Exposure, Bioaccumulation , and Adverse Effects in Wildlife and Humans in the Context of Rapid Global Change. 13th International Conference on Mercury as a Global Pollutant. DOI: 10.1007/s13280-017-1011-x.

Eagles-Smith CA, Silbergeld EK, Basu N, Bustamante P, Diaz-Barriga F, Hopkins WA, Kidd KA, Nyland JF (2018) Modulators of mercury risk to wildlife and humans in the context of rapid global change. Ambio 47: 170-197.

Eagles-Smith CA, Wiener JG, Eckley CS, Willacker JJ, Evers DC, Marvin-DiPasquale M, Obrist D, Fleck JA, Aiken GR, Lepak JM, Jackson AK (2016b) Mercury in western North America: A synthesis of environmental contamination, fluxes, bioaccumulation, and risk to fish and wildlife. Science of The Total Environment 568: 1213-1226.

Eagles-Smith CA, Willacker JJ, Nelson SJ, Flanagan Pritz CM, Krabbenhoft DP, Chen CY, Ackerman JT, Grant EHC, Pilliod DS (2020) A national-scale assessment of mercury bioaccumulation in United States national parks using dragonfly larvae as biosentinels

through a citizen-science framework. Environmental Science & Technology 54: 8779-8790.

Eckley CS, Luxton TP, Knightes CD, Shah V (2021) Methylmercury Production and Degradation under Light and Dark Conditions in the Water Column of the Hells Canyon Reservoirs, USA. Environ. Toxicol. Chem. 40: 1827-1837.

Edmonds ST, Evers DC, Cristol DA, et al. (2010) Geographic and Seasonal Variation in Mercury Exposure of the Declining Rusty Blackbird. The Condor 112: 789-799. DOI: 10.1525/cond.2010.100145

EFSA (2012) Scientific Opinion on the Risk for Public Health Related to the Presence of Mercury and Methylmercury in Food. EFSA Journal 10 (12). DOI: 10.2903/j.efsa.2012.2985.

Evans RD, Addison EM, Villeneuve JY, MacDonald KS, Joachim DG (2000) Distribution of inorganic and methylmercury among tissues in mink (Mustela vison) and otter (Lutra canadensis). Environ. Res. 84: 133-139. DOI: 10.1006/enrs.2000.4077.

Evers DC (2018) The Effects of Methylmercury on Wildlife: A Comprehensive Review and Approach for Interpretation. Encyclopedia of the Anthropocene. Elsevier.

Evers DC, Han YJ, Driscoll CT, Kamman NC, Goodale MW, Lambert KF, Holsen TM, Chen CY, Clair TA, Butler T (2007) Identification and Evaluation of Biological Hotspots of Mercury in the Northeastern U.S. and Eastern Canada. Bioscience 57: 29-43.

Evers DC, Keane SE, Basu N, Buck D (2016) Evaluating the effectiveness of the Minamata Convention on Mercury: Principles and recommendations for next steps. Science of the Total Environment 569: 888-903.

Evers DC, Mason RP, Kamman NC, Chen CY, Bogomolni AL, Taylor DL, Hammerschmidt CR, Jones SH, Burgess NM, Munney K, Parsons KC (2008a) Integrated mercury monitoring program for temperate estuarine and marine ecosystems on the North American Atlantic coast. EcoHealth 5: 426-441.

Evers DC, Sauer AK, Burns DA, Fisher NS, Bertok DC, Adams EM, Burton MEH, Driscoll CT (2020) A synthesis of patterns of environmental mercury inputs, exposure and effects in New York State. Ecotoxicology 29: 1565-1589.

Evers DC, Savoy LJ, DeSorbo CR, Yates DE, Hanson W, Taylor KM, Siegel LS, Cooley JH, Bank MS, Major A, Munney K (2008b) Adverse effects from environmental mercury loads on breeding common loons. Ecotoxicology 17: 69-81.

Evers DC, Schmutz JA, Basu N, DeSorbo CR, Fair J, Gray CE, Paruk JD, Perkins M, Regan K, Uher-Koch BD, Wright KG (2014) Historic and contemporary mercury exposure and potential risk to Yellow-billed Loons (Gavia adamsii) breeding in Alaska and Canada. Waterbirds 37: 147-159.

Evers DC, Sunderland E (2019) Technical Information Report on Mercury Monitoring in Biota: Proposed components towards a strategic long-term plan for monitoring mercury in fish and wildlife globally. UN Environment Programme, Chemicals and Health Branch, Geneva, Switzerland. 40 pp.

FAO (2020) The State of World Fisheries and Aquaculture 2020. Sustainability in Action. DOI: 10.4060/ca9229en.

Fielding R, Kiszka JJ, Macdonald C, McCormack MA, Dutton J, Ollivierre AD, Arnett JA, Elkins M, Darby NA, Garcia HM, Skinner S (2021) Demographic and geographic patterns of cetacean-based food product consumption and potential mercury

exposure within a Caribbean whaling community. Human and Ecological Risk Assessment 2021: 1-25.

Fu X, Feng X, Sommar J, Wang S (2012) A review of studies on atmospheric mercury in China. Sci. Total Environ. 421: 73-81.

Fu X, Zhang H, Yu B, Wang X, Lin C, Feng X (2015) Observations of atmospheric mercury in China: a critical review. Atmos. Chem. Phys. 15: 9455-9476.

Gabriel MC, Howard N, Osborne TZ (2014) Fish mercury and surface water sulfate relationships in the Everglades Protection Area. Environmental Management 53: 583-593.

Gerson JR, Szponar N, Almeyda Zambrano A, Bergquist B, Broadbent E, Driscoll CT, Erkenswick G, Evers DC, Fernandez LE, Hsu-Kim H, Inga G, Lansdale K, Marchese MJ, Martinez A, Moore C, Pan W, Silman M, Ury EA, Vega C, Watsa M, and Bernhardt ES (2021) Amazon forests capture high levels of atmospheric mercury pollution from artisanal gold mining. Nature Communications, in press.

Gichuki SW, Mason RP (2013) Mercury and metals in South African precipitation. Atmos. Environ. 79: 286-298.

Golden HE, Knightes CD (2011) Simulated watershed mercury and nitrate flux responses to multiple land cover conversion scenarios. Environmental Toxicology and Chemistry 30: 773-786. DOI: 10.1002/etc.449.

Goodrich JM, Chou H-N, Gruninger SE, Franzblau A, Basu N (2016) Exposures of Dental Professionals to Elemental Mercury and Methylmercury. Journal of Exposure Science & Environmental Epidemiology 26: 78-85. DOI: 10.1038/jes.2015.52.

Gorski PR, Armstrong DE, Hurley JP, Krabbenhoft DP (2008) Influence of natural dissolved organic carbon on the bioavailability of mercury to a freshwater alga. Environmental Pollution 154: 116-123.

Gustin M, Evers DC, Bank M, Hammerschmidt CR, Pierce A, Basu N, Blum J, Bustamante P, Chen C, Driscoll CT, Horvat M, Jaffe D, Pacyna J, Pirrone N, Selin N (2016) Importance of integration and implementation of emerging and future mercury research into the Minamata Convention. Environmental Science and Technology 50: 2767-2770.

Gustin M, Jaffe D (2010) Reducing the Uncertainty in Measurement and Understanding of Mercury in the Atmosphere. Environ. Sci. Technol. 44: 2222-2227.

Gustin MS, Amos HM, Huang J, Miller MB, Heidecorn K (2015) Measuring and modeling mercury in the atmosphere: A critical review. Atmos. Chem. Phys. 15: 5697-5713.

Gustin MS, Bank MS, Bishop K, Bowman K, Branfireun B, Chételat J, Eckley CS, Hammerschmidt CR, Lamborg C, Lyman S, Martínez-Cortizas A, Sommar J, Tsui MT, Zhang T (2020) Mercury biogeochemical cycling: A synthesis of recent scientific advances. Sci Total Environ. 737: 139619-139632. DOI: 10.1016/j.scitotenv.2020.139619. PMID: 32783819.

Gustin MS, Dunham-Cheatham SM, Huang J, Lindberg S, Lyman SN (2021) Development of an Understanding of Reactive Mercury in Ambient Air: A Review. Atmosphere 12: 73-91.

Gustin MS, Huang J, Miller MB, Peterson C, Jaffe DA, Ambrose J, Finley BD, Lyman SN, Call K, Talbot R, Feddersen D, Mao H, Lindberg SE (2013) Do We Understand What the

Mercury Speciation Instruments Are Actually Measuring? Results of RAMIX, Environ. Sci. Technol. 47: 7295-7306.

Ha E, Basu N, Bose-O'Reilly S, Dórea JG, McSorley E, Sakamoto M, Chan HM (2017) Current Progress on Understanding the Impact of Mercury on Human Health. Environmental Research 152: 419-433. DOI: 10.1016/j.envres.2016.06.042.

Harris R, Krabbenhoft DP, Mason R, Murray MW, Reash R, Saltman T (2007) Ecosystem responses to mercury contamination: Indicators of change. CRC Press, Boca Raton, FL.

HBM4EU (2017) Statistical Analysis Plan. Available at: https://www.hbm4eu.eu/work-packages/deliverable-10-5-statistical-analysis-plan.

HBM4EU (2018a) Legal and Ethics Policy Paper Update August 2018. Available at: https://www.hbm4eu.eu/?mdocs-file=4263.

HBM4EU (2018b) SOP 1 : Selection of Participants and Recruitment. Available at: https://www.hbm4eu.eu/mdocs-posts/sop-1-selection-of-participants-and-recruitment.

HBM4EU (2018c) SOP 2 : Quality Assurance for Recruitment and Fieldwork. Available at: https://www.hbm4eu.eu/mdocs-posts/sop-2-quality-assurance-for-recruitment-and-fieldwork.

HBM4EU (2018d) SOP 3: Procedure for Obtaining Human Samples. Available at: https://www.hbm4eu.eu/mdocs-posts/sop-3-procedure-for-obtaining-human-samples.

HBM4EU (2019) Prioritised List of Biomarkers, Matrices and Analytical Methods for the 2 Nd Prioritisation Round of Substances. Available at: https://www.hbm4eu.eu/upload-work-package/get/file/Deliverable-9.5-Prioritised-list-of-biomarkers-matrices-and-analytical-methods-for-the-2nd-prioritisation-round-of-substances.pdf.

HBM4EU (2020a) 2nd Prioritisation Report on Survey Design: Study Protocols , SOPs and Guidelines, Tailored and Transferred Questionnaires for Recruitment and Sampling. Available at: https://www.hbm4eu.eu/upload-work-package/get/file/Deliverable-7.6-2nd-prioritisation-Report-on-survey-design-Study-protocols-SOPs-and-Guidelines-tailored-and-transferred-questionnaires-for-recruitment-and-sampling.pdf.

HBM4EU (2020b) Interviewer Manual to the Basic Questionnaire for 2nd Round Priority Substances. Available at: https://ec.europa.eu/research/participants/documents/downloadPublic?documentIds=080166e5cffe1a02&appId=PPGMS.

HBM4EU (2020c) Matrix-Specific Questionnaires to Accompany the Sampling of Urine and Blood. Available at: https://www.hbm4eu.eu/?mdocs-file=6084.

HBM4EU (2020d) WP10 Data Management and Analysis Task 10.4 Statistical Analysis. Available at: https://www.hbm4eu.eu/?mdocs-file=8353.

Hedgecock IM, Pirrone N (2004) Chasing quicksilver: Modeling the atmospheric lifetime of Hg-(g)(0) in the marine boundary layer at various latitudes, Environ. Sci. Technol. 38: 69-76. DOI: 10.1021/es034623z.

Hedgecock IM, Pirrone N, Trunfio GA, Sprovieri F (2006) Integrated mercury cycling, transport, and air-water ex-change (MECAWEx) model, J. Geophys. Res. 111: D20302. DOI: 10.1029/2006JD007117.

Horowitz HM, Jacob DJ, Zhang Y, Dibble TS, Slemr F, Amos HM, Schmidt JA, Corbitt ES, Marais EA, Sunderland EM (2017) A new mechanism for atmospheric mercury redox chemistry: Implications for the global mercury budget, Atmos. Chem. Phys. 17: 6353-6371.

Hsu-Kim H, Eckley CS, Achá D, Feng X, Gilmour CC, Jonsson S, Mitchell CP (2018) Challenges and opportunities for managing aquatic mercury pollution in altered landscapes. Ambio 47: 141-169.

Hsu-Kim H, Kucharzyk KH, Zhang T, Deshusses MA (2013) Mechanisms regulating mercury bioavailability for methylating microorganisms in the aquatic environment: a critical review. Environmental Science and Technology 47: 2441-2456.

ILO (1998) Technical and ethical guidelines for workers' health surveillance Occupational Safety and Health Series, No. 72. International Labour Organization, Geneva, Switzerland.

Jung G, Hedgecock IM, Pirrone N (2009) ECHMERIT V1.0: A new global fully coupled mercury-chemistry and transport model. Geoscientific Model Development 2: 175-195.

Klapstein SJ, Ziegler SE, Risk DA, O'Driscoll NJ (2017) Quantifying the effects of photoreactive dissolved organic matter on methylmercury photodemethylation rates in freshwaters. Environ. Toxicol. Chem. 36: 1493-1502.

Knightes CD, Golden HE, Journey CA, Davis GM, Conrads PA, Marvin-DiPasquale M, Brigham ME, Bradley PM (2014) Mercury and methylmercury stream concentrations in a Coastal Plain watershed: A multi-scale simulation analysis. Environmental Pollution 187: 182-192. DOI: 10.1016/j.envpol.2013.12.026.

Knightes CD, Sunderland EM, Barber MC, Johnston JM, Ambrose RB, Jr. (2009) Application of ecosystem-scale fate and bioaccumulation models to predict fish mercury response times to changes in atmospheric deposition. Environmental Toxicology and Chemistry 28: 881-893. DOI: 10.1897/08-242R.1.

Kocman D, Wilson SJ, Amos HM, Telmer KH, Steenhuisen F, Sunderland EM, Mason RP, Outridge P, Horvat M (2017) Toward an Assessment of the Global Inventory of Present-Day Mercury Releases to Freshwater Environments. International Journal of Environmental Research and Public Health 14: 138.

Kronberg RM, Schaefer JK, Björn E, Skyllberg U (2018) Mechanisms of Methyl Mercury Net Degradation in Alder Swamps: The Role of Methanogens and Abiotic Processes. Environmental Science & Technology Letters 5: 220-225.

Kwon SY, Blum JD, Yin R, Tsui MTK, Yang Y-H, Choi JW (2020) Mercury stable isotopes for monitoring the effectiveness of the Minamata Convention on Mercury. Earth Sci. Rev. 203: 103111. DOI: 10.1016/j.earscirev.2020.103111.

Lee CS, Lutcavage ME, Chandler E, Madigan DJ, Cerrato RM, Fisher NS (2016) Declining mercury concentrations in bluefin tuna reflect reduced emissions to the North Atlantic Ocean. Environmental Science and Technology 50: 12825-12830.

Lepak RF, Janssen SE, Yin R, Krabbenhoft DP, Ogorek JM, DeWild JF, Tate MT, Holsen TM, Hurley JP (2018) Factors affecting mercury stable isotopic distribution in piscivorous fish of the Laurentian Great Lakes. Environmental Science & Technology 52: 2768-2776.

Lescord GL, Johnston TA, Branfireun BA, Gunn JM (2018) Percentage of methylmercury in the muscle tissue of freshwater fish varies with body size and age and among species. Environ. Toxicol. Chem. 37: 2682-2691. DOI: 10.1002/etc.4233.

Li F, Ma C, Zhang P (2020) Mercury Deposition, Climate Change and Anthropogenic Activities: A Review, Front. Earth Sci. 8: 316. DOI: 10.3389/feart.2020.00316.

Liu M, Zhang Q, Maavara T, Liu S, Wang X, Raymond PA (2021) Rivers as the largest source of mercury to coastal oceans worldwide. Nature Geoscience 14: 672-677.

Lukina AO, Fisher M, Khoury C, Than J, Guay M, Paradis JF, Arbuckle TE, Legrand M (2021) Temporal Variation of Total Mercury Levels in the Hair of Pregnant Women from the Maternal-Infant Research on Environmental Chemicals (MIREC) Study. Chemosphere 264: 128402. DOI: 10.1016/j.chemosphere.2020.128402.

Lyman SN, Cheng I, Gratz LE, Weiss-Penzias P, Zhang L (2020) An updated review of atmospheric mercury. Sci Total Environ. 707: 135575-135595.

Macagnano A, Papa P, Avossa J, Perri V, Marelli M, Sprovieri F, Zampetti E, De Cesare F, Bearzotti A, Pirrone N (2018) Passive sampling of gaseous elemental mercury based on a composite TiO2NP/AuNP layer. Nanomaterials 8: 798.

Mahaffey KR, Sunderland EM, Chan HM, Choi AL, Grandjean P, Marien K, Oken E, et al. (2011) Balancing the Benefits of N-3 Polyunsaturated Fatty Acids and the Risks of Methylmercury Exposure from Fish Consumption. Nutrition Reviews 69: 493-508. DOI: 10.1111/j.1753-4887.2011.00415.x.

Manceau A, Azemard S, Hédouin L, Vassileva E, Lecchini D, Fauvelot C, Swarzenski PW, Glatzel P, Bustamante P, Metian M (2021) Chemical Forms of Mercury in Blue Marlin Billfish: Implications for Human Exposure. Environmental Science & Technology Letters 8: 405-411.

Martin LG, Labuschagne C, Brunke E-G, Weigelt A, Ebinghaus R, Slemr F (2017) Trend of atmospheric mercury concentrations at Cape Point for 1995–2004 and since 2007. Atmos. Chem. Phys, 17, 2393–2399, 2017.

Marumoto K, Suzuki N, Shibata Y, Takeuchi A, Takami A, Fukuzaki N, Kawamoto K, Mizohata A, Kato S, Yamamoto T, Chen J, Hattori T, Nagasaka H, Saito M (2019) Long-Term Observation of Atmospheric Speciated Mercury during 2007–2018 at Cape Hedo, Okinawa, Japan. Atmosphere 10: 362. DOI: 10.3390/atmos10070362.

Mason RP, Abbott ML, Bodaly RA, Bullock OR, Jr, Driscoll CT, Evers D, Lindberg SE, Murray M, Swain EB (2005) Monitoring the response to changing mercury deposition. Environmental Science & Technology 39: 14A-22A.

McLagan DS, Mitchell CPJ, Huang H, Abdul Hussain B, Lei YD, Wania F (2017) The effects of meteorological parameters and diffusive barrier reuse on the sampling rate of a passive air sampler for gaseous mercury. Atmos. Meas. Tech. 210: 3651-3660.

McLagan DS, Mitchell CPJ, Huang H, Lei YD, Cole AS, Steffen A, Hung H, Wania F (2016) A high precision passive air sampler for gaseous mercury. Environ. Sci. Technol. Lett. 3: 24-29.

McLagan DS, Mitchell CPJ, Steffen A, Hung H, Shin C, Stupple GW, Olson ML, Luke WT, Kelley P, Howard D, Edwards GC, Nelson PF, Xiao H, Sheu G-R, Dreyer A, Huang H, Abdul Hussain B, Lei YD, Tavchunsky I, Wania F (2018) Global evaluation and calibration of a passive air sampler for gaseous mercury. Atmos. Chem. Phys. 18: 5905-5919.

MOEJ (2011) Manual of Measurement Method of Hazardous Air Pollutants, Environment Management Bureau of Water & Air Environment Fields, Ministry of Environment Japan.

Monteiro LR, Furness RW (2001) Kinetics, Dose–Response, and Excretion of Methylmercury in Free-Living Adult Cory's Shearwaters. Environmental Science & Technology 35: 739-746. DOI: 10.1021/es000114a.

Munthe J, Waengberg I, Pirrone N, Iverfeldt Å, Ferrara R, Ebinghaus R, Feng X, Gårdfeldt K, Keeler G, Lanzillotta E, Lindberg SE, Lu J, Mamane Y, Prestbo E, Schmolke S, Schroeder WH, Sommar J, Sprovieri F, Stevens RK, Stratton W, Tuncel G, Urba A (2001) Intercomparison of methods for sampling andanalysis of atmospheric mercury species. Atmos. Environ. 35: 3007-3017.

Naccarato A, Tassone A, Martino M, Moretti S, Macagnano A, Zampetti E, Papa P, Avossa J, Pirrone N, Nerentorp M, Munthe J, Wängberg I, Stupple GW, Mitchell CPJ, Martin AR, Steffen A, Babi D, Prestbo EM, Sprovieri F, Wania F (2021) A field intercomparison of three passive air samplers for gaseous mercury in ambient air. Atmos. Meas. Tech. 14: 3657-3672.

Obrist D, Kirk J, Zhang L, Sunderland E, Jiskra M, Selin NE (2018) A review of global environmental mercury processes in response to human and natural perturbations: changes of emissions, climate and land use. Ambio 47: 116-140.

OHAT (2015) Handbook for Conducting a Literature-Based Health Assessment Using OHAT Approach for Systematic Review and Evidence Integration. Office of Health Assessment and Translation (OHAT), Division of the National Toxicology Program, National Institute of Environmental Health Sciences, USA.

Olk TR, Karlsson T, Lydersen E, Økelsrud A (2016) Seasonal variations in the use of profundal habitat among freshwater fishes in Lake Norsjø, Southern Norway, and subsequent effects on fish mercury concentrations. Environments 3: 29.

Peterson SA, Van Sickle J, Hughes RM, et al. (2004) A Biopsy Procedure for Determining Filet and Predicting Whole-Fish Mercury Concentration. Arch Environ Contam Toxicol 48: 99-107. DOI: 10.1007/s00244-004-0260-4.

Peterson SH, Ackerman JT, Costa DP (2016). Mercury correlations among blood, muscle, and hair of northern elephant seals during the breeding and molting fasts. Environ. Toxicol. Chem. 35: 2103-2110. DOI: 10.1002/etc.3365.

Peterson SH, Ackerman JT, Eagles-Smith CA, Hartman CA, Herzog MP (2017) A critical evaluation of the utility of eggshells for estimating mercury concentrations in avian eggs. Environ. Toxicol. Chem. 36: 2417-2427. DOI: 10.1002/etc.3777.

Peterson SH, Ackerman JT, Toney M, Herzog MP (2019) Mercury Concentrations Vary Within and Among Individual Bird Feathers: A Critical Evaluation and Guidelines for Feather Use in Mercury Monitoring Programs. Environ. Toxicol. Chem. 38: 1164-1187. DOI: 10.1002/etc.4430.

Pirrone N, Cinnirella S, Feng X, Finkelman RB, Friedli HR, Leaner J, Mason R, Mukherjee AB, Stracher G, Streets DG, Telmer K (2009) Global Mercury Emissions to the Atmosphere from Natural and Anthropogenic Sources, Springer, New York, USA, chap. 1: 3-49.

Pirrone N, Ferrara R, Hedgecock IM, Kallos G, Mamane Y, Munthe J, et al. (2003) Dynamic processes of mercury over the Mediterranean region: results from the

Mediterranean Atmospheric Mercury Cycle System (MAMCS) Project. Atmospheric Environment 37: S21-S39.

Poste AE, Braaten HFV, de Wit HA, Sørensen K, Larssen T (2015) Effects of photodemethylation on the methylmercury budget of boreal Norwegian lakes. Environ. Toxicol. Chem. 34: 1213-1223.

Prestbo EM, Gay DA (2009) Wet deposition of mercury in the U.S. and Canada, 1996–2005: Results and analysis of the NADP mercury deposition network (MDN). Atmospheric Environment 43: 4223-4233.

Renedo M, Bustamante P, Cherel Y, Pedrero Z, Tessier E, Amouroux D (2020) A "seabird-eye" on mercury stable isotopes and cycling in the Southern Ocean. Science of the Total Environment 742: 140499.

Riget F, Braune B, Bignert A, Wilson S, Aars J, Born E, Dam M, Dietz R, Evans M, Evans T, Gamberg M (2011) Temporal trends of Hg in Arctic biota, an update. Science of the Total Environment 409: 3520-3526.

Rimmer CC, McFarland KP, Evers DC, et al. (2005) Mercury Concentrations in Bicknell's Thrush and Other Insectivorous Passerines in Montane Forests of Northeastern North America. Ecotoxicology 14: 223-240. DOI: 10.1007/s10646-004-6270-1.

Risch MR, DeWild JF, Gay DA, Zhang L, Boyer EW, Krabbenhoft DP (2017) Atmospheric mercury deposition to forests in the eastern USA, Environmental Pollution 228: 8-18.

Rodriguez CAB, Bezerra MF, Rezende CE, Bastos WR, Lacerda LD (2019) Mercury and methylmercury in carapace of the marine turtle Caretta caretta, in northeastern Brazil and its potential for environmental monitoring. Anais da Academia Brasileira de Ciências 91. DOI: 10.1590/0001-3765201920180672.

Ronald K, Tessaro SV, Uthe JF, Freeman HC, Frank R (1997) Methylmercury poisoning in the harp seal (Pagophilus groenlandicus). Science of The Total Environment 8: 1-11.

Rothenberg SE, Windham-Myers L, Creswell JE (2014) Rice Methylmercury Exposure and Mitigation: A Comprehensive Review. Environmental Research 133: 407-23. DOI: 10.1016/j.envres.2014.03.001.

Rudd JW, Bodaly RA, Fisher NS, Kelly CA, Kopec D, Whipple C (2018) Fifty years after its discharge, methylation of legacy mercury trapped in the Penobscot Estuary sustains high mercury in biota. Science of the Total Environment 642: 1340-1352.

Saiz-Lopez A, Travnikov O, Sonke J, Thackray C, Jacob D, Carmona-García J, Francés-Monerris A, Roca-Sanjuán D, Acuña AA, Dávalos J, Cuevas C, Jiskra M, Wang F, Bieser J, Plane J, Francisco J (2020) Photochemistry of oxidized Hg (I) and Hg (II) species suggests missing mercury oxidation in the troposphere. Proceedings of the National Academy of Sciences 117: 30949-30956.

Schartup AT, Qureshi A, Dassuncao C, Thackray CP, Harding G, Sunderland EM (2018) A model for uptake and trophic transfer of methylmercury by marine plankton. Environ. Sci. Technol. 52: 654-662.

Scheuhammer A, Braune B, Chan HM, Frouin H, Krey A, Letcher R, Loseto L, Noël M, Ostertag S, Ross P, Wayland M (2015) Recent progress on our understanding of the biological effects of mercury in fish and wildlife in the Canadian Arctic. Science of the Total Environment 509: 91-103.

Scheuhammer AM, Lord SI, Wayland M, Burgess NM, Champoux L, Elliott JE (2016) Major correlates of mercury in small fish and common loons (Gavia immer) across four large study areas in Canada. Environmental Pollution 210: 361-370.

Schmeltz D, Evers DC, Driscoll CT et al. (2011) MercNet: a national monitoring network to assess responses to changing mercury emissions in the United States. Ecotoxicology 20: 1713-1725.

Schudel G, Miserendino RA, Veiga MM, Velasquez-López PC, Lees PS, Winland-Gaetz S, Guimarães JRD, Bergquist BA (2018) An investigation of mercury sources in the Puyango-Tumbes River: Using stable Hg isotopes to characterize transboundary Hg pollution. Chemosphere 202: 777-787.

Selin NE (2018) A proposed global metric to aid mercury pollution policy. Science 360: 607-609.

Sergio F, Caro T, Brown D, Clucas B, Hunter J, Ketchum J, McHugh K, Hiraldo F (2008) Top Predators as Conservation Tools: Ecological Rationale, Assumptions, and Efficacy. Annual Review of Ecology, Evolution, and Systematics 39: 1-19.

Shah V, Jacob DJ, Thackray CP, Wang X, Sunderland EM, Dibble TS, Saiz-Lopez A, Černušák I, Kellö V, Castro PJ, Wu R, Wang C (2021) Improved Mechanistic Model of the Atmospheric Redox Chemistry of Mercury. Environ. Sci. Technol. Article, in press. DOI: 10.1021/acs.est.1c03160.

Sheehan MC, Burke TA, Navas-Acien A, Breysse PN, McGready J, Fox MA (2014) Global Methylmercury Exposure from Seafood Consumption and Risk of Developmental Neurotoxicity: A Systematic Review. Bulletin of the World Health Organization 92: 254-269F. DOI: 10.2471/blt.12.116152.

Sherman LS, Blum JD, Basu N, Rajaee M, Evers DC, Buck DG, Petrlik J, Digangi J (2015) Assessment of Mercury Exposure among Small-Scale Gold Miners Using Mercury Stable Isotopes. Environmental Research 137: 226-234. DOI: 10.1016/j.envres.2014.12.021.

Sherman LS, Blum JD, Franzblau A, Basu N (2013) New Insight into Biomarkers of Human Mercury Exposure Using Naturally Occurring Mercury Stable Isotopes. Environ. Sci. Technol. 47: 3403-3439. DOI: 10.1021/es305250z.

Sheu G, Gay D, Schmeltz D, Olson M, et al. (2019) A New Monitoring Effort for Asia: The Asia Pacific Mercury Monitoring Network (APMMN). Atmosphere 10: 481-497.

Si L, Ariya PA (2018) Recent advances in atmospheric chemistry of mercury. Atmosphere 9: 76. DOI: 10.3390/atmos9020076

Skov H, Brooks S, Goodsite M, Lindberg S, Meyers T, Landis M, Larsen M, Jensen B, McConville G, Christensen J (2006) Fluxes of reactive gaseous mercury measured with a newly developed method using relaxed eddy accumulation. Atmos. Environ. 40: 5452-5463. DOI: 10.1016/j.atmosenv.2006.04.061, 2006.

Skov H, Christensen JH, Goodsite ME, Heidam NZ, Jensen B, Wahlin P, Geernaert G (2004) Fate of elemental mercury in the arctic during atmospheric mercury depletion episodes and the load of atmospheric mercury to the arctic. Environ. Sci. Technol. 38: 2373-2382.

Skov H, Hjorth J, Nordstrøm C, Jensen B, Christoffersen C, Poulsen MB, Liisberg JB, Beddows D, Dall'Osto M, Christensen JH (2020) Variability in gaseous elemental mercury at Villum Research Station, Station Nord, in North Greenland from 1999 to 2017. Atmos. Chem. Phys. 20: 13253-13265. DOI: 10.5194/acp-20-13253-2020.

Slemr F, Martin L, Labuschagne C, Mkololo T, Angot H, Magand O, Dommergue AD, Garat P, Ramonet, M, Bieser J (2020) Atmospheric mercury in the Southern Hemisphere – Part 1: Trend and Inter-annual variations in atmospheric mercury at Cape Point, South Africa, in 2007–2017, and on Amsterdam Island in 2012–2017, Atmos. Chem. Phys. 20: 7683-7692.

Sprovieri F, Pirrone N, Bencardino N, et al. (2016) Atmospheric mercury concentrations observed at ground-based monitoring sites globally distributed in the framework of the GMOS network. Atmos. Chem. Phys. 16: 11915-11935.

Sprovieri F, Pirrone N, Bencardino N, et al. (2017) Five-year records of mercury deposition flux at GMOS sites in the Northern and Southern hemispheres. Atmos. Chem. Phys. 17: 2689-2708.

Sprovieri F, Pirrone N, Ebinghaus R, Kock H, Dommergue A (2010) A review of worldwide atmospheric mercury measurements. Atmos. Chem. Phys. 10: 8245-8265.

Sprovieri F, Pirrone N, Landis MS, Stevens RK (2005) Oxidation of gaseous elemental mercury to gaseous divalent mercury during 2003 polar sunrise at Ny-Alesund. Environ. Sci. Technol. 39: 9156-9165.

Steffen A, Bottenheim J, Cole A, Ebinghaus R, Lawson G, Leaitch WR (2014) Atmospheric mercury speciation and mercury in snow over time at Alert, Canada, Atmos. Chem. Phys. 14: 2219-2231.

Steffen A, Douglas T, Amyot M, et al. (2008) A synthesis of atmospheric mercury depletion event chemistry in the atmosphere and snow. Atmos. Chem. Phys. 8: 1445-1482.

Steffen A, Lehnherr I, Cole A, Ariya P, Dastoor A, Durnford D, Kirk J, Pilote M (2015) Atmospheric mercury in the Canadian Arctic. Part I: a review of recent field measurements. Sci. Total Environ. 509-510: 3-15. DOI: 10.1016/j.scitotenv.2014.10.109.

Steffen A, Scherz T, Olson M, Gay D, Blanchard P (2012) A comparison of data quality control protocols for atmospheric mercury speciation measurements. Environ. Monit. 14: 752-765.

Stern AH (2005) A Revised Probabilistic Estimate of the Maternal Methyl Mercury Intake Dose Corresponding to a Measured Cord Blood Mercury Concentration. Environmental Health Perspectives 113: 2. DOI: 10.1289/ehp.7417.

Sunderland EM, Driscoll Jr CT, Hammitt JK, Grandjean P, Evans JS, Blum JD, Chen CY, Evers DC, Jaffe DA, Mason RP, Goho S, Jacobs W (2016) Benefits of regulating hazardous air pollutants from coal and oil-fired utilities in the United States. Environmental Science and Technology 50: 2117-2120.

Sunderland EM, Li M, Bullard K (2018) Decadal Changes in the Edible Supply of Seafood and Methylmercury Exposure in the United States. Environmental Health Perspectives. DOI: 10.1289/EHP2644.

Taylor VF, Buckman KL, Seelen EA, Mazrui NM, Balcom PH, Mason RP, Chen CY (2019) Organic carbon content drives methylmercury levels in the water column and in estuarine food webs across latitudes in the Northeast United States. Environmental Pollution 246: 639-649.

Trasande L, DiGangi J, Evers DC, Petrlik J, Buck DG, Šamánek J, Beeler B, Turnquist MA, Regan K (2016) Economic implications of mercury exposure in the context of the global

mercury treaty: hair mercury levels and estimated lost economic productivity in selected developing countries. Journal of Environmental Management 183: 229-235.

Travnikov O, Angot H, Artaxo P, et al. (2017) Multi-model study of mercury dispersion in the atmosphere: atmospheric processes and model evaluation. Atmos. Chem. Phys. 17: 5271–5295. DOI: 10.5194/acp-17-5271-2017, 2017

UNEP (2016) Global Review of Mercury Monitoring Networks. United Nations Environment (UNEP), Geneva, Switzerland.

UNEP (2019) Global Mercury Assessment 2018. United Nations Environment (UNEP), Geneva, Switzerland.

UNEP, WHO (2008) Guidance for Identifying Populations At Risk From Mercury Exposure. United Nations Environment Programme (UNEP) and World Health Organization (WHO), Geneva, Switzerland.

USEPA (2006) Method 1631 Revision E. Mercury Laboratory Analysis. United States Environmental Protection Agency, Athens, Georgia.

USEPA (2016) Guidance for Conducting Fish Consumption Surveys. EPA 823-B-16-002. United States Environmental Protection Agency, Washington, DC.

Valdes C, Black FJ, Stringham B, Collins JN, Goodman JR, Saxton HJ, Mansfield CR, Schmidt JN, Yang S, Johnson WP (2017) Total Mercury and Methylmercury Response in Water, Sediment, and Biota to Destratification of the Great Salt Lake, Utah, United States. Environmental Science & Technology 51: 4887-4896.

Vandenbroucke JP, Von Elm E, Altman DG, et al. (2007) Strengthening the Reporting of Observational Studies in Epidemiology (STROBE): Explanation and Elaboration. PLoS Medicine. DOI: 10.1371/journal.pmed.

Vander Zanden MJ, Casselman JM, Rasmussen JB (1999) Stable isotope evidence for the food web consequences of species invasion in lakes. Nature 401: 464-467.

Veiga MM, Baker R (2004) Protocols for Environmental and Health Assessment of Mercury Released by Artisanal and Small-Scale Gold Miners. United Nations Industrial Development Organization (UNIDO), Vienna, Austria.

Višnjevec AM, Kocman D, Horvat M (2014) Human Mercury Exposure and Effects in Europe. Environmental Toxicology and Chemistry 33: 1259-70. DOI: 10.1002/etc.2482.

Wang F, Saiz-Lopez A, Mahajan A, Martín JG, Armstrong D, Lemes M, Hay T, Prados-Roman C (2014) Enhanced production of oxidised mercury over the tropical Pacific Ocean: a key missing oxidation pathway. Atmos. Chem. Phys. 14: 1323-1335. DOI:10.5194/acp-14-1323-2014.

Wang R, Feng X-B, Wang W-X (2013) In Vivo Mercury Methylation and Demethylation in Freshwater Tilapia Quantified by Mercury Stable Isotopes. Environ. Sci. Technol. 47: 7949–7957. DOI: 10.1021/es3043774.

Wang X, Bao ZD, Lin C-J, Yuan W, Feng X (2016) Assessment of global mercury deposition through litterfall. Environmental Science and Technology 50: 8548-8557.

Wang X, Yuan W, Lin C-J, Zhang LM, Zhang H, Feng XB (2019) Climate and vegetation as primary drivers for global mercury storage in surface soil. Environmental Science and Technology 53: 10665-10675.

Wängberg I, Hageström U, Sommar J, Ferm M (2016) Development and Testing of a Passive Sampler for Measurement of Gaseous Mercury. Atmos. Chem. Phys. Discuss [preprint]. DOI: 10.5194/acp-2016-528.

Wangman R, Trebacz E, Boila G, Lockhart WL (1998) Methylmercury and total mercury in tissues of arctic marine mammals. Sci. Total Environment 218: 19-31.

Weiss-Penzias P, Amos H, Selin N, Gustin M, Jaffe D, Obrist D, Sheu G-R, Giang A (2015) Use of a global model to understand speciated atmospheric mercury observations at five high-elevation sites. Atmos. Chem. Phys. 15: 1161-1173.

Weiss-Penzias PS, Gay DA, Brigham ME, Parsons MT, Gustin MS, ter Schure A (2016) Trends in mercury wet deposition and mercury air concentrations across the US and Canada. Sci. Total Environ. 568: 546-556. DOI: 10.1016/j.scitotenv.2016.01.061.

Whitney MC, Cristol DA (2017) Impacts of Sublethal Mercury Exposure on Birds: A Detailed Review. Reviews of Environmental Contamination and Toxicology 244: 113-163. DOI: 10.1007/398_2017_4.

WHO (2015) Human biomonitoring: facts and figures. World Health Organization. Regional Office for Europe. https://apps.who.int/iris/handle/10665/164588.

WHO (2018a) Assessment of prenatal exposure to mercury: human biomonitoring survey. The first survey protocol. World Health Organization, Regional Office for Europe, Copenhagen, Denmark.

WHO (2018b) Assessment of prenatal exposure to mercury: standard operating procedures. World Health Organization, Regional Office for Europe, Copenhagen, Denmark.

WHO (2021a) Human biomonitoring in artisanal and small-scale gold mining: ethical and scientific principles. World Health Organization, Geneva, Switzerland.

WHO (2021b) A step-by-step guide for developing a public health strategy for artisanal and small-scale gold mining in the context of the Minamata Convention on Mercury. World Health Organization, Geneva, Switzerland.

Wiener JG, Evers DC, Gay DA, Morrison HA, Williams KA (2012a) Mercury contamination in the Laurentian Great Lakes region: Introduction and overview. Environmental Pollution 161: 243-251.

Wiener JG, Sandheinrich MB, Bhavsar SP, Bohr JR, Evers DC, Monson BA, Schrank CS (2012b) Toxicological significance of mercury in yellow perch in the Laurentian Great Lakes region. Environmental Pollution 161: 350-357.

Wilkinson M, Dumontier M, Aalbersberg I, et al. (2016) The FAIR Guiding Principles for scientific data management and stewardship. Sci Data 3, 160018. DOI: 10.1038/sdata.2016.18.

WMO (2009) Operations Handbook - Ozone Observations with a Dobson Spectrophotometer: revised 2008. WMO/GAW Report 183.

WMO (2010) Guidelines for the Measurement of Atmospheric Carbon Monoxide. WMO/GAW Report 192.

WMO (2012) Standard Operating Procedures (SOPs) for Air Sampling in Stainless Steel Canisters for Non-Methane Hydrocarbons Analysis. WMO/GAW Report 204.

WMO (2014) Quality Assurance and Quality Control for Ozonesonde Measurements in GAW. WMO/GAW Report 201.

Wright PL, Zhang L, Marsik F (2016) Overview of mercury dry deposition, litterfall, and throughfall studies. Atmos. Chem. Phys 16: 13399-13416.

Wyn B, Kidd KA, Burgess NM, Curry RA (2009) Mercury biomagnification in the food webs of acidic lakes in Kejimkujik National Park and National Historic Site, Nova Scotia. Canadian Journal of Fisheries and Aquatic Sciences 66: 1532-1545.

Xu H, Sonke JE, Guinot B, Fu X, Sun R, Lanzanova A, Candaudap F, Shen Z, Cao J (2017) Seasonal and annual variations in atmospheric Hg and Pb isotopes in Xi'an, China. Environ. Sci. Technol. 51: 3759-3766.

Yang Y, Yanai RD, Schoch N, Buxton VL, Gonzales KE, Evers DC, Lampman GG (2020) Determining optimal sampling strategies for monitoring mercury and reproductive success in common loons in the Adirondacks of New York. Ecotoxicology 29: 1786-1793.

Zhang L, Wu Z, Cheng I, Wright LP, Olson ML, Gay DA, Risch MR, Brooks S, et al. (2016) The estimated six-year mercury dry deposition across North America. Environmental Science and Technology 50: 12864-12873.

Zhang L, Zhou P, Cao S, Zhao Y (2019) Atmospheric mercury deposition over the land surfaces and the associated uncertainties in observations and simulations: a critical review. Atmos. Chem. Phys. 19: 15587-15608.

Zhang Y, Soerensen AL, Schartup AT, Sunderland EM (2020) A Global Model for Methylmercury Formation and Uptake at the Base of Marine Food Webs. Global Biogeochemical Cycles 34, e2019GB006348. DOI: 10.1029/2019GB006348.

Zhang Y, Song Z, Huang S, et al. (2021) Global health effects of future atmospheric mercury emissions. Nat. Commun. 12: 3035. DOI: 10.1038/s41467-021-23391-7.