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

Draft of 15 April 2021

Chapter 1 (Acknowledgements) and **Chapter 2** (List of abbreviations and glossary of terms) are to be developed.

Chapter 3: Introduction and objectives

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 fulfill 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 populations, are of particular interest to COP in the context of the effectiveness evaluation.

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:

- Explain the role of monitoring in the effectiveness evaluation and set realistic expectations about what can be learned over time.
- Provide guidance to Parties and organizations that are currently conducting monitoring programs on what data and accompanying information would inform the effectiveness evaluation.
- Provide guidance to Parties 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 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 4 builds on the four overarching policy questions proposed for the effectiveness evaluation and establishes five categories of information that comparable monitoring data can provide to address these questions. It then provides the rationale for selecting air, biota and

human as core matrices for monitoring activities, and presents general guidance for ensuring comparability and quality of monitoring data that is common to all the matrices.

Following chapters address monitoring of mercury in specific matrices: air (Chapter 5), biota (Chapter 6) and humans (Chapter 7). 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 8 discusses how these matrix-specific data can be compiled, analyzed and synthesized, how those data can be used in mechanistic and statistical models, and how the 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.

[Reference to be added] present 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. These [reference to be added] will be "living documents" that might be updated to support the COP in identifying available monitoring information for effectiveness evaluation, as well as to support Parties and stakeholders to consider monitoring activities to fill 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 4: Use of comparable monitoring data for the effectiveness evaluation

4.1 Questions to be addressed by comparable monitoring data

Paragraphs 2 and 3 of Article 22 of the Minamata Convention require 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" and to use this information, along with other information, to inform a periodic effectiveness evaluation. It has been proposed that an effectiveness evaluation should address the following four policy questions:[1]

1) Have the Parties taken actions to implement the Minamata Convention?
2) Have the actions taken resulted in changes in mercury supply, use, emissions and releases into the environment?
3) 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?
4) 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. Attributing observed changes in mercury levels to actions influenced by the Minamata Convention or estimating the human or ecosystem health impacts

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

of those changes requires 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.

The overarching policy questions can be operationalized through monitoring activities that are organized across the following five information categories:

1) Characterization of representative levels and spatial patterns from local to global scales.
2) Identification of temporal trends.
3) Quantification of key environmental processes to understand cause-effect relationships.
4) Modelling source attribution.
5) Estimation of exposure and adverse impacts.

Each of these information categories may be elaborated in the form of questions, as outlined in **Table 4.1** below, which may support the collection and analysis of the relevant monitoring data. The answers to these questions can form an evidentiary basis for addressing questions 3 and 4 of the effectiveness evaluation framework. Different types of observations or observation contexts may be most appropriate for addressing different questions.

Table 4.1 Types of Information from Monitoring and Associated Analysis Questions

1) Characterization of representative levels and spatial patterns from local to global scales

- What are the level and form of mercury in the observed matrix (air, biota, human) at a given location and time?
- Taken together, what does the available data suggest about:
  - Spatial variability?
  - Ecosystem/Population variability?
- Do the observed spatial variations and patterns or gradients differ among:
  - Forms (chemical species) of mercury?
  - Environmental matrices?
- How do the observed spatial variations and patterns or gradients compare to those of:
  - Mercury emissions and releases?
  - Related pollutants/emissions or environmental variables?

2) Identification of temporal trends

- Do the level and form of mercury in the observed matrix (air, biota, human) at a given location change over time? What is the temporal variability?
- How do observed temporal variations and trends differ spatially?
- How do observed temporal variations and trends in mercury compare to or co-vary with variations and trends of:
  - Mercury in different forms (chemical species) or other within matrices?
  - Mercury emissions and releases?

- Related pollutants/emissions or environmental variables?

3) *Quantification of key environmental processes to understand cause-effect relationships*

- What does the level, spatial pattern, or temporal trends of mercury suggest about:
  - Sources (primary anthropogenic, legacy, and natural) and sinks of mercury?
  - Relative importance of environmental processes and parameters driving transport and fate?
  - Exposure and adverse impacts to humans or biota?
- How consistent are the observed levels, spatial patterns and temporal trends with the modelled estimates and what lessons can be learned from them to improve the existing models ?

4) *Modelling source attribution*

- Using models and statistical analyses consistent with observational data, how can the observed levels, spatial patterns, and temporal trends in the environment, biota, and humans be attributed to changes in:
  - Sources or sinks of natural and legacy mercury?
  - Anthropogenic sources (local, regional, global) of mercury?
  - Changes influenced by the Convention?
  - Changes not influenced by the Convention?

5) *Estimation of exposure and adverse impacts*

- How do the observed levels of mercury in air, biota, and humans compare to established benchmark levels associated with adverse effects on human health and the environment?
- How are the changes in exposure attributed to mitigation measures or behavioral changes influenced by the Convention?

4.2 Monitoring matrices

Mercury has been recognized as 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 significant adverse effects on human health and the environment. For the purpose of evaluating the effectiveness of the Convention's primary objective (i.e., "to protect human health and the environment from mercury"), it is important to track 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 in the context of the effectiveness evaluation of the Convention.

Air: Mercury levels in the atmosphere are linked to mercury emissions to air. 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. Many of the Convention measures to control mercury supply, use, emissions, storage, and disposal are expected to reduce levels of mercury in the

atmosphere. Article 22 of the Convention requires the COP to establish arrangements to provide monitoring data on the trends in mercury levels in the environment.

Biota: Mercury released into the environment may be methylated and accumulated in fish and wildlife, and it can negatively impact human health through the consumption of food. Mercury may also cause significant negative impacts to the environment, for example by adversely impacting ecosystem services and leading to the loss of 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 (specifically 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 specific areas, such as those addressing ASGM. Levels of mercury in the oceans' surface can 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 6. Review of environmental monitoring methods that may be used to assess ASGM sites is also being developed[2] to support the implementation of Article 7 of the Convention. Existing monitoring information, across a range of matrices, is reviewed in [refence to be added]. Some of this information may be considered for use in effectiveness evaluation.

4.3 Tiered approach for developing and improving monitoring programmes

To support Parties 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):

- Tier 1 is intended to provide guidance to Parties that are seeking to create a monitoring program, or expand a minimal program, but that may not have sufficient resources to

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

implement the actions in Tier 2. Following guidance by the COP,[3] the methods in Tier 1 are cost effective, practical, feasible, and sustainable. The Tier 1 methods are intended to provide information that are useful to Parties in identifying and characterizing gaps and needs of national, regional, or local interest and to provide information that is useful to the collective effectiveness evaluation effort. While the implementation of Tier 1 actions may not fully address the four policy questions (see Chapter 3), 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 all policy questions identified by the COP and create a basis particularly for source attribution at the local, national, and global scales. The methods and approaches may be more expensive or complex than those under Tier 1. It would not be necessary that all Parties implement the Tier 2 actions for the effectiveness evaluation to be performed at a global level, but the basis for the effectiveness evaluation will be more robust with greater implementation.

- Tier 3 identifies research methods and approaches that may play a vital role in supporting the Tier 2 programs and the effectiveness evaluation, primarily by improving our understanding of key processes that link sources to environmental concentrations and exposures. Not all Parties would be expected to employ Tier 3 methods or approaches, but where Tier 3 efforts and results are available, the information may be taken into consideration in the effectiveness evaluation.

4.4 Quality of monitoring data

The effectiveness evaluation of the Convention will require monitoring data that is comparable, as well as credible and of sufficient quality. In order to ensure the suitability of data used to understand the presence and movement of mercury in the environment and humans, the quality assurance and quality control (QA/QC) protocols, which are employed by the different monitoring programmes, will need to be well documented and implemented in a transparent manner to enable the use of data from across different monitoring programmes for the purposes of effectiveness evaluation. As such, identifying the minimum requirements for QA/QC will help the assessment of the quality and comparability of data generated from various monitoring programmes.

The quality of mercury measurements can often be assessed on the basis of:

- Selection bias – describe the location/population/setting (Is the data representative and sampled in a consistent manner (e.g. following standard operating procedures)?);

- Exposure detection (Is the measurement accurate and precise? Are additional metadata provided to do exposure/attribution assessment?);

- Statistical parameters – describe the adequate sample size, and whether basic and essential data is present, complete and well summarized (Is the data useful to address the policy questions?).

Chapters 5, 6 and 7 describe QA/QC considerations for specific media or matrices[4].

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

[4] The terms “media” and “matrix/matrices” are being used interchangeably throughout this document.

4.5 Data management

The effectiveness evaluation may employ a system that adheres to the FAIR principles (findable, accessible, interoperable and reusable[5]) for data management and stewardship including the following elements, as appropriate:

Findable:

- A central and searchable database acting as the repository of available data;
- Unique identification systems (e.g. “Digital Object Identifiers” or “DOIs”) and controlled vocabulary to facilitate searching and retrieval of information;
- 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 all Governments, Indigenous Peoples and local communities 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 about how the data were collected/produced to facilitate reproducibility and further analyses.

As monitoring data become findable, accessible, interoperable and reusable, it can be analyzed and interpreted, for example through the use of mechanistic and statistical modelling to answer some of the driving policy questions. Further details are discussed in Chapter 8.

Major international and national monitoring programmes and networks have their own data management systems, including data repositories, portals, or catalogs. Some of these may be used as primary repositories to gather and exchange information so that monitoring data, relevant metadata, ancillary data and QA/QC information from different programs can be used for 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. These data management systems and networks are reviewed in [reference to be added].

[5] Wilkinson et al. (2016).

Chapter 5: Monitoring Mercury Pollution in Air

This chapter will highlight the importance of measuring and monitoring mercury in the atmosphere in the context of evaluating the effectiveness of the Minamata Convention. The development of a scientifically sound strategy, based on transparent processes, harmonized methodologies, and reliable and comparable existing data is critical. Answering basic questions of how, where, what, and why will be increasingly important to facilitate the effectiveness evaluation of the Minamata Convention and assure data quality and comparability.

5.0 Introduction

To assist the Convention in reaching its objectives, it is necessary to make use of sound science-based information and comparable data that show how mercury concentrations in the environment are changing; and specifically how these changes relate to measures introduced under the Convention. In this connection, monitoring of atmospheric mercury has been identified as one of the primary and most appropriate types of monitoring to assist the Convention in determining its effectiveness (Evers et al., 2016, Gustin et al. 2016). Monitoring of atmospheric mercury generally has the following objectives: 1) to gather information on the spatial and temporal trends variations of mercury concentrations in the atmosphere; 2) assessing atmospheric mercury inputs to aquatic and terrestrial ecosystem and 3) to provide data for the development and improving of transport and chemistry models (Obrist et al., 2018). The aim of monitoring under the Convention, is that the data generated be used for the effectiveness evaluation, but in order for this to transpire, the data collected should be comparable. This Chapter aims to provide guidance to Parties on the different methods to monitor mercury in air, what vital procedures need to be in place to continue or start with mercury in air monitoring to generate comparable monitoring data. Certain Parties have monitoring stations that's have been operational for many years, then there are Parties that have just started out recently by joining a monitoring network and then there are Parties with no monitoring activities. To accommodate the different needs, a tiered approach, such as the one presented in the **Annex** to this document, offers a way forward to support Parties in starting or expanding on their monitoring program.

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 concentrations may be attributable to controls on mercury emissions and releases affected 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.

To Parties, the tiered approach offers an array of standard options for methods, recognizing that Parties have different mercury monitoring needs, capabilities, and experience. The guidance offered in this chapter describes criteria to consider when deciding which measurement methods use, the frequency of measurements, and where to potentially locate sites depending on the Parties monitoring needs and objectives. For mercury in air monitoring the Tiered approach can be seen as follows:

Tier 1 – Documents mercury trends and spatial distribution in air (TGM/GEM) and in wet deposition over broad geographic areas and provides information to inform atmospheric modeling (statistical and mechanistic). Following the COPs guidance to the TEG, the measurement methods in this tier are cost effective, practical, feasible, and sustainable. The options reflect a basic expectation that the Parties will pursue one or more, (preferably a

combination) of these basic sampling options (e.g., continuous, manual, passive), where feasible, to undertake their obligation for the effectiveness evaluation.

Tier 2 – Explains temporal trends and attributes mercury sources to mercury concentrations in biota. Sampling is more intensive than Tier 1 sites and also relies more on ancillary measurements and models to gain insight into the links between emissions and concentration trends in biota. All Parties will not be required to implement the Tier 2 recommendations for a global effectiveness evaluation to be performed, but the basis for the effectiveness evaluation will be more robust with greater implementation.

Tier 3 – Improves representativeness of the measurements and understanding of key processes (e.g., related to transformation and deposition) using new, advanced measurement techniques and sophisticated research. Not all Parties are expected to employ Tier 3 methods or approaches, but where Tier 3 efforts and results are available, the information should be taken into consideration in the effectiveness evaluation.

5.1 Why mercury in air should be considered as a matrix

Mercury is a naturally occurring element and is emitted to the atmosphere from a variety of natural, legacy emissions and anthropogenic sources (Driscoll et al., 2013). Atmospheric deposition represents the major pathway of mercury input to terrestrial and aquatic ecosystems outside areas of direct contamination. Land and ocean processes play an important role in the redistribution of mercury in terrestrial, freshwater, and marine ecosystems and the production of methylmercury (MeHg), which can enter the food web and adversely affect human health (Eagles-Smith et al., 2018). Monitoring efforts and modeling studies have shown that the atmosphere responds relatively quickly (though not proportionally due to legacy emissions) to decreasing mercury emissions.

Mercury is emitted to the atmosphere in three main forms, gaseous elemental mercury, reactive mercury species and particle associated mercury. The residence time for gaseous elemental mercury (GEM) in the atmosphere is relatively long — approximately 3 to 12 months compared to the other forms (gaseous oxidized mercury and particle-bound mercury) with a short residence time in the atmosphere ranging from a few hours to weeks (Horowitz et al., 2017). This, together with the influence of re-emissions of previously deposited mercury mean that there will be a delay between abatement and observation of a response in atmospheric mercury levels (Sprovieri et al., 2016). The lag time is expected to be much shorter than the time response for mercury 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 (Layman et al., 2020). However, in the case for human biomonitoring, studies have shown that the effects due to mercury exposure to vulnerable populations, will also occur relatively fast (Basu et., al 2018). 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).

5.1.1 Mercury in air

As noted, 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. Atmospheric mercury is measured in the following forms:

- Gaseous Elemental Mercury (Hg^0; GEM),
- Gaseous Oxidized Mercury (Hg^{II}; GOM) (operationally defined[6])
- Total Gaseous Mercury (TGM) = GEM + GOM
- Particle-bound Mercury (PBM)

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 therefore be transported and deposited to remote locations such as the Arctic and Antarctic (Skov et al., 2004).

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, Layman et al., 2020). Measurements of mercury species such as oxidized and particle-bound mercury compounds are therefore important as they help to improve the understanding of short-term oxidation processes regarding the removal of mercury from the atmosphere (Pirrone et al. 2000; Fu et al., 2015; Weiss-Penzias et al., 2016; De Simone et al. 2016). The speciation method is well defined, used in various networks, there are SOP's available and comparison studies between different monitoring networks has been done delivering satisfactory results (Steffen et al., 2012, Gustin et al., 2015; Sprovieri et al., 2016). However, speciation measurements although very important in helping to understand the global mercury cycle as well as improving model output, is quite complex, costly, requires very skilled operators and therefore not considered as a priority species to monitor at this stage for the effectiveness evaluation of the Minamata Convention. It should however be noted that several monitoring networks and research groups do perform continuous mercury speciation measurements and are therefore encouraged to continue with this monitoring as the data collected from this technique will be helpful in answering questions listed in Table 4.1, and should be taken into consideration for the effectiveness evaluation.

5.1.2 Atmospheric mercury wet deposition

After mercury is emitted into the atmosphere, it eventually returns to the Earth's surface via a process called atmospheric deposition. On a global scale, net atmospheric mercury deposition (the difference between re-emissions and deposition) is determined by the overall emissions, whereas deposition at the local scale is more influenced by atmospheric processes and the speciation of regional and local emissions. Understanding these processes is therefore critical for the development of accurate predictive models that can inform global mercury policy. The pathway for wet deposition occurs when mercury is deposited with 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.

The amount of precipitation is the main driver of wet mercury deposition to aquatic and terrestrial habitats (Prestbo et al., 2009; Weiss-Penzias et al. 2016; Sprovieri et al. 2017). GOM and PBM are water soluble and the primary atmospheric forms responsible for wet deposition of mercury in precipitation.

Wet deposition of mercury in precipitation represents an important pathway for a better understanding of mercury exchange at the land-water-air interface. Monitoring mercury in

[6] 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).

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 mercury monitoring, wet deposition monitoring is relatively easy to start off with and is considered a very reliable method to achieve comparable data (Shue et al., 2019).

Recommendations

- Continuation of mercury wet deposition networks and the intercomparison studies on wet deposition methods, equipment and SOPs, is encouraged.
- Parties that do consider this monitoring activity join an existing network making use of that networks guidelines to achieve comparable results.

5.2 Where should atmospheric mercury measurements be collected

Site characteristics can seriously affect the concentration levels of mercury in air; as a result, site selection is a critical part of any monitoring network's design (Schmeltz et al., 2011) To support the effectiveness evaluation of the Convention, a variety of sites are needed, each with a different mercury footprint. Selection of monitoring sites for effectiveness evaluation of the Minamata Convention should be based on the site's potential to: provide insight into changing atmospheric mercury, help assess the contributions to sensitive ecosystems, and help validate atmospheric models, all of which aim is to help answer the different policy question as explained in the previous chapter.

Such types of sites should include a) background or remote, b) rural, c) urban and d) contaminated or industrial sites. Site type A will be able to provide information on determining long-term global trends and provide data for evaluating and refining transport models, B will provide information on mercury concentrations that is regionally representative and is limited to some degree by the influence of significant local pollution sources, C can provide information on non-point sources such as, local fossil fuel combustion and Hg transported in cities. Data from urban areas will be useful to improve emissions inventories, site type D will assist with determining the effect on human exposure of communities living close to point sources (e.g ASGM activities) or when surveys needs to be performed. For more site specific criteria, the following documents provide guidelines and are widely used within existing mercury networks (WMO/GAW Report 160, NADP Site Selection and Installation Manual).

Therefore, sites are needed that can observe the changes of mercury in air that are occurring at background sites (far from Hg sources) as well as more impacted or vulnerable sites (close to point sources). This will create an observing system that will help track changes that are occurring locally, regionally and eventually globally. When establishing new sites, it is important that the following practical points be taken into consideration before a final decision is made, a) the availability of stable electricity, b) access to the sampling site for personnel and equipment, c) air conditioning (for automated instruments), d) internet to download data remotely and access equipment, e) building and storage facilities/security, etc.

Recommendations

- New sites be established in conjunction with current monitoring activities (air quality sites and weather stations) to make use of available infrastructure and co-measurements with reactants important for Hg.

- If existing air monitoring stations do not have the capacity to measure mercury, technical upgrades may have to be considered.
- Considering that site consistency has an important impact on mercury concentration data, mercury monitoring data submitted by Parties be used for the effectiveness evaluation process.
- In establishing new sites, a combination of the various sites (A-D) is recommended as this will allow Parties to address several monitoring questions.

5.2.1 Data gaps

In the past 2 decades a number of Hg monitoring sites have been established (in Europe, Canada, USA and Asia) as part of regional networks and/or projects (Pirrone et al., 2003; Steffen et al., 2008, Sprovieri et al.,2016), 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. The main reason is to make consistent a globally distributed Hg observations available that can be used to validate regional and global-scale models for assessing global patterns of Hg concentrations and deposition fluxes. Several different initiatives have done this, some of which are still operational. Building on these and expanding on this can help filling in the gaps and improve the process of evaluating the effectiveness of the Convention.

From the current global picture, Europe and North America are spatially very well covered but there exist gaps in the Asia-Pacific region and Russia. In the southern hemisphere large data gaps exists in Africa, the tropics and South America.

Recommendations

- Establishing of new sites in the southern hemisphere, Tropics, Russia and Asia-Pacific region be prioritized and sufficient resources be made available for this.
- Data collected within existing and new networks be used to fill in the gaps by integrating the data using sufficient techniques and processes.
- For establishing new sites, the focus could be on sites that are expected to be sensitive to Hg input.

5.3 How to measure mercury in air: Sampling and measurement methods

Several different methods are available for monitoring of air Hg. Selection of methods should be based on purpose of monitoring. All methods employed in a monitoring program need to be tested, intercompared and validated to ensure quality of data used for effectiveness evaluation, research or other purposes. This section is to help Parties to select monitoring techniques that would best meet their monitoring needs and requirements. However, flexibility shall be allowed in the selection of techniques.

5.3.1 Active air sampling

Active air sampling methods involve air pulled through a pump at a constant flow rate through on an active material, otherwise known as a trap. 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 trap for a set amount of time, the mercury trapped is removed from the trap using thermal desorption and spectroscopic detection.

Active sampling of this kind can be undertaken in an instrument that does both the sample and analysis in situ (automatically) or can be done 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 Continuous Mercury Air Measurements; and (b) Manual Non-Continuous Mercury Air Measurements.

(a) Automated continuous mercury air measurements

Currently, continuous mercury measurements (GEM/TGM and speciation measurements) are recorded with different commercial instruments available from various manufacturers that are capable of detecting mercury at very low concentrations in air (sub ppm to ppt). These instruments have high temporal resolution, low limit of detection, established and proven quality assurance and quality control protocols (Sprovieri et al., 2016).

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

- Cold Vapour Atomic Fluorescence Spectroscopy (CVAFS)
- Cold Vapour Atomic Absorption Spectroscopy (CVAAS)

GEM/TGM concentration is presented as nano gram (ng) of Hg^0 per m^3, using volumes at standard pressure and temperature. The CVAFS instruments are more sensitive in comparison to CVAAS, but require pure Ar or He gas during the desorption and detection step, whereas CVAAS instruments use mercury free air or nitrogen.

It should be noted that the mercury air community has by far chosen to use pure gold dual trap method with CVAFS for national, international networks of background and regional stations and the majority of published research papers report using this method (Sprovieri et al., Martin et al., 2017, Slemr et al., 2020). At the same time, for many other applications where low detection limits are not required, less concern about interference and a lower QA threshold is set, the use of different instrument methods like single gold-sand trap-CVAAS are good choices e.g fence line monitoring, artisanal gold mining, industrial locations, etc. 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 analyzers are substantial 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. Automated continuous monitoring will be able to address several monitoring questions (e.g identify spatial trends, data input to models etc.) as outlined in Table 4.1 and is ideally suited for background and regional monitoring to meet the effectiveness evaluation requirements.

Recommendations

- Instruments capable of detecting ambient mercury (GEM/TGM) concentrations with CVAAS or CVAFS be used for continuous mercury monitoring which are applicable to the selected site requirements.

- SOP's from available networks and programs (AMNet, APMMN, CAMNet, EMEP, GMOS) be compared and where possible and needed integrated to ensure that there is sufficient intercomparability.
- New monitoring stations make use of SOP's that are currently used in established networks.
- Field intercomparisons are also possible for continuous measurements by co-locating equipment at sites. This method of intercomparison could be expensive but, can be seen as a better way to ensure comparability than requiring the exact same SOP.
- As new methods become available SOP's for them undergo the same review process to ensure sufficient backward comparability.

(b) Manual non-continuous mercury air measurements

With this technique, mercury in the atmosphere is collected manually on an adsorbent material (gold tube) over a 24h period at a constant flow rate using a pump. Analysis of TGM in ambient air is possible when the sample is analysed after exposure using thermal desorption and spectroscopic detection in a laboratory. With this method, analysis and collection of TGM in the ambient air is possible. 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. The temporal resolution is lower than it is for continuous mercury monitoring, but 24h-average is suitable for effectiveness evaluation purposes. It should also be noted that achieving quality results with this method, will require an operator with trace-clean technique experience and consistent low blanks. Interference from Ozone and high humidity can also influence the performance. (Environment Management Bureau of Water & Air Environment Fields, MOEJ, 2011).

Recommendations

- For this method the gold traps have proved to be the most reliable adsorbent material.
- The review of manual method guidelines and SOP's used in different networks and research groups be conducted.
- Networks and research groups using this method, start an intercomparison study or round robin exercise.
- This method be used if consistent low blanks and replicates are achieved and local weather conditions favorable for sampling.
- The existing global networks of Manual Non-Continuous measurement be put forward as a method to be used in the monitoring process of the effective evaluation.

5.3.2 Passive sampling

Passive Air Sampling (PAS) for gaseous mercury has recently had an increase in method developments. 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, through isotopic methods, can contribute to Hg source characterization. PAS are easy to deploy, are low cost,

require no electricity, can be deployed at background, remote, urban, hotspots and there is no worry about media failure as a new sampler with active material is used each time. After exposure, the PAS are easily analyzed with well-documented and credible methods in any analytical laboratory (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 poor temporal resolution and can be affected by meteorology (Mclagan et al., 2017). Despite this, PAS are useful when sampling close to contaminated sites or hot spots and can assist with understanding contaminated sites emissions to air, specifically describing concentration gradients. PAS have shown to be a useful tool to monitor at contaminated locations as other methods such as gold traps are too sensitive for these locations. For Parties that have no mercury air monitoring program or previous mercury air experience, passive sampling is considered an excellent method to start with. This Tier 1 approach will enable Parties to contribute in answering monitoring program objectives 1, that speaks to (spatial variability), 2 (trends and emissions) and 5 (impacts on local ASGM communities) if passive are exposed at relevant locations.

A comparative evaluation of three types of PASs was carried out during late winter and early spring of 2019. The study involved the side-by-side deployment of PASs at two monitoring sites, located in Italy and Canada, over a three-month period. The PASs involved in this study were developed by the Italian Institute of Atmospheric Pollution Research (CNR-IIA) (Macagnano et al., 2018), the Swedish Environmental Research Institute (IVL) (Wängberg et al., 2016) and the University of Toronto (McLagan et al., 2016a). Data were submitted for compilation to a blind third party in order to control for bias. The performances of the PASs were assessed for accuracy through comparison with active sampling data, for precision and for sensitivity (e.g. the method detection limit, MDL), as well as in terms of the linearity of uptake over extended deployment periods. The major finding has been presented and discussed in Naccarato et al. 2021 (under revision, preprint on-line available).

Recommendations

- Passive samplers that have undergone inter-comparison studies and are commercially available can be used.
- Mercury passive networks that are currently functioning on a global scale be put forward as a method to be used in the monitoring process of the effectiveness evaluation.
- Parties that do not have any monitoring activities or infrastructure start monitoring with passive samplers first, then consider other monitoring methods.
- As new passive samplers become available their sufficient intercomparisons to existing samplers is established through intercomparison studies.

5.3.3 Wet deposition sampling

Mercury wet deposition sampling can be carried out using a variety of commercially available precipitation collectors for either wet only or bulk collection (Sprovieri et al.,2017). Wet only samples are collected using Teflon or borosilicate glass bottles. When operating a wet only sampler, the precipitation sensor is activated during a rain event, and the lid moves off the funnels and comes to rest. When a rain 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 (Sheu et al., 2019). The wet-only samplers have the advantage of avoiding particle dry deposition

511 although the contribution to the measured wet deposition fluxes from gaseous or particulate
512 mercury species is probably not large in non-industrialised or non-urban areas. For extended
513 sampling periods it is also necessary to prevent significant gas phase diffusion of Hg^0 to the
514 surface of the collected sample where it could contribute to the mercury content of the sample
515 via oxidation to water-soluble forms. This can be easily done using a capillary tube between the
516 funnel and the bottle. Shielding of the sample bottles from light is also necessary to avoid photo-
517 induced reduction of the mercury in the precipitation sample.

518 When starting with wet deposition monitoring, the following factors should be considered when
519 choosing a sampling location; a) availability of stable electricity when using a wet only collector
520 (bulk collectors requires no electricity) b) laboratory facilities to prepare samples and glassware
521 that needs to be treated with acids and other chemicals, c) portable sample trays and coolers
522 for collecting and transporting field samples, d) skilled operator to conduct analysis, e)
523 availability of meteorological data at sampling site, f) fridge or other storage facility for samples.

Recommendations

- Samples should preferably be collected as wet only samples and Parties use wet only collectors that have undergone comparison studies (e.g. Japan, Taiwan study in APMMN and USA)
- Parties can also use wet collectors used in other established networks (Canada, China, EMEP, GMOS, NADP, Republic of Korea) that's been published in literature.
- Another alternative is to collect samples using bulk (continuously open to the atmosphere) deposition sampling will also be useful and should be collected if there's no electricity at sampling sites or no wet only equipment available but Parties should beware that comparability across wet/bulk networks will not always be possible.
- Data from a bulk samplers are acceptable if that bulk samples are collected in locations sufficiently remote from pollution and in areas of high precipitation, so that contamination from and proportion of dry deposition is negligible.
- For, parties that are interested in starting a wet deposition program it is advisable to align with established networks (APMMN, EMEP, GMOS, NADP) and use their SOP's as this will minimize variability.
- The establishment of a variety (different locations) of global wet deposition sites by Parties or networks is strongly advised where field intercomparison between the different wet collectors used can take place. This process will help to achieve comparable data as multi sites will reflect different conditions.

524 **5.4 Quality Assurance and Quality Control (QA/QC)**

525 High quality accurate results are possible by following comparable analytical chemical methods,
526 using sensitive instrumentation and strict QA/QC procedures throughout the sampling, storage
527 and analysis procedure. Implementing these crucial steps will ensure that reliable information
528 and comparable mercury data is available on which to conduct research and make policy
529 decisions. This section will highlight key QA/QC procedures that should be followed for air
530 measurements to be comparable throughout the different monitoring networks.

531 ***5.4.1 Quality assurance and quality control for field operations***

Described below are QA/QC activities and procedures recommended for use at mercury measurement sites.

(a) Siting and instrument placement

To ensure collection of spatially- and temporally-representative data which will help to address different monitoring questions, all monitoring sites should satisfy the siting criteria presented in Chapter 5.2 for the various monitoring sites. Monitoring sites should be assessed regularly with respect to the siting criteria to determine whether they continue to meet the guidance siting requirements.

(b) Instrumentation

It is important to ensure that the correct installation and operation of automated/manual instruments and wet deposition collectors are followed. Where appropriate, then manufactures guidelines should be followed or relevant SOP's within networks. Instruments must meet minimum requirements for e.g., sensor sensitivity chemical inertness. Repaired and newly purchased instruments should be pre-tested in the laboratory before being sent for operational use.

(c) Sample collection and handling

Specific quality control procedures that prevent contamination from occurring during sample collection and handling include:

- Inspecting the collection vessels for visible signs of contamination before they are placed in the collector;
- Using deionized water in the field to clean collectors before use;
- Wearing disposable plastic gloves whenever handling precipitation collectors, passive samplers and transferring samples from field sites;
- 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.

(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:

- Maintaining samples (precipitation) in cooled containers while in transit and when stored in laboratory;
- Weighing samples to determine sample volume at the station and at the laboratory in order to detect leaks in transit;
- If properly stored and preserved, precipitation samplers can be stored for up to 6 months before it needs to be analysed (Gay et al.,2009, Sprovieri et al., 2017);
- Passive samples should be stored in a cool dry place after collected from the field in double bags and sealed tightly.

(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 per month For weekly sampling, 1 to 2 blanks per month. This should be done at every site. 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 sampler is exposed with the other samples but it's sealed that it's closed to the atmosphere.

5.4.2 QA/QC for laboratory operations

Laboratory operations cover a wide range of activities including sample reception, field sampling support, sample transfer, storage and data reporting. Specific quality control procedures in the laboratory where analysis will be performed as part of the effectiveness evaluation process should include but are not limited to:

- Good laboratory practice and sample handling practice;
- Documentation of analytical procedures;
- Appropriate safety measures;
- Traceability of calibration standards;
- Accuracy checks (calibration controls, spikes, blinds, reagent blanks).

5.4.3 Data quality control and reporting

Data QA/QC procedures needs to be applied routinely to ensure that analytical results are accurate and complete. Specific QA/QC procedures include:

- Verification that all extreme values and below detection limit values are correct;
- Identification and re-analysis of samples that failed criteria.

5.4.4 Laboratory intercomparison studies

All laboratories that will analyze mercury samples for the purpose of the effectiveness evaluation of the Minamata Convention should participate in intercomparison studies. This is an excellent way to not only gauge laboratory

performance, but also to detect problems in analysis procedures.Standard samples containing known mercury concentrations from accredited facilities or departments (e.g. U.S. Geological Survey) can be used to determine each lab's analytical capability and can highlight any bias in the analytical protocol. Participation in laboratory intercomparisons through different networks are encouraged. Unsatisfactory performance by a laboratory reduces the overall quality of the results. Corrective action will help to improve the performance of an under-performing laboratory to a satisfactory level. Suggested actions include:

- Improving the internal laboratory quality control programme
- Routine system of analysing and control-charting Certified Reference Materials
- Arranging for an expert visit/audit by another mercury laboratory that is part of a recognized network.

5.4.5 Data flagging

The purpose of the data flagging process is to further improve the overall quality of the data collected .This process involves the reviewing of field and laboratory metadata for problems (e.g data gaps, error messages) and inaccuracies, Several networks (e.g AMNet, CAMNet) and databases have this capability and satisfactory results were achieved when different networks automated flagging criteria and data were used for GEM measurements (Steffen et al., 2012). Data collected under the convention will greatly benefit from having an automated QA/QC procedure in place for continuous GEM/TGM data. Within GMOS, an online G-DQM system modelled after the AMNet/CAMNet QA program was specifically developed to screen continuous ambient data from the various monitoring sites for QA/QC purposes to achieve comparable and harmonized data from all the ground-based monitoring sites within the network (D'Amore et al., 2015). This online portal is still available and could be used to serve the Convention's needs. Advances in cyberinfrastructure and sensor networks now provide enormous quantities of data, even in near real-time. Therefore, the problem is no longer how much data we have, but what is its quality. The following flags are examples of set criteria used within the various networks and will be useful for data collected under the Convention (MOEJ, 2011, Steffen et al., 2012, D'Amore et al., 2015, Sprovieri et al., 2017, McLagan et al., 2018):

- **Automated CV-AFS:** A/B ratio of gold cartridge should be less than 10%, multiple peaks detected.
- **Automated CV-AAS:** Lamp voltage to low, cell temperature outside of range of 10 – 40°C.
- **Passive sampling:** Blank contains more than 2 ng of Hg.
- **Manual Method:** Differences of absorbance for constructing the calibration curve higher then ± 10%.
- **Wet Deposition:** No precipitation during exposure period/ too little sample less than 50ml.

5.4.6 Calibration and Traceability

Proper calibration of field and laboratory equipment is essential. All calibrations should be traceable to accepted international standards and only certified reference material used. Calibration procedures be documented as part of the QA/QC protocol. In the laboratory, calibrations must be applied to all analytical procedures. For automated instruments, measuring TGM/GEM automated calibration is advisable following the manufactures guidelines or SOP of a recognized network. For manual measurements, calibration curves should be constructed regularly.

5.5 Timing (frequency and duration) of measurements

The following sampling and exposure times are guidelines for the various sampling methods and sites locations depending on how the data will be used.

(a) Continuous measurement with CV-AFS/AAS

- A sampling time of 5 min, or 15 min where the average concentration is lower than 1.0 ng/m^3 (mostly sites located in the SH).
- Hourly averages can be used for the data analysis
- Each time interval should be reported with an appropriate uncertainty estimate.

(b) Manual Measurements

- A sampling time of 24 hours and a constant flow rate is recommended as several networks are currently using this parameter and available data can be used as comparison.
- One sample per week is acceptable but if personnel and equipment are available, more samples a week will be beneficial.

(c) Wet deposition

- Weekly sampling is recommended, e.g., Tuesday to Tuesday to provide an integrated - 7-day sample as this is the preferred procedure used within most existing wet deposition networks currently
- If the station is remote and man-power and equipment is limited, exposure of the wet collector should be done on a bi-weekly basis.
- What is of importance is that once an exposure period has been decided on, the site operators should keep to that schedule.

(d) Passive sampling

- To generate data with PAS, an exposure time of 3-months is suitable for effectiveness evaluation for a global network as several other monitoring programs (e.g GAPS network for POP's) operates on this time frame (Rauert et al., 2016)
- If sufficient resources and manpower are available, the sampling time can be increased to monthly exposure which will increase data points and very helpful in the effectiveness evaluation process.

5.6 Data collection

Observations form the basis of the effectiveness evaluation. Together with statistical and model based analysis the observed changes of Hg levels in environmental matrices will contribute in determining the outcome of the effectiveness evaluation. Therefore, to ensure their comparability, data needs to be sufficiently harmonized and validated. All data used for the evaluation could be made available directly or indirectly via a FAIR compliant common data portal. The latter is important to ensure that all conclusions are based on an agreed set of observations. Especially the data quality process will play an integral part in achieving data that is harmonized, comparable, and accessible throughout the different mercury networks and measurement campaigns. Past experience has shown that lots of officially available data is in reality effectively inaccessible. The following aspects will be helpful for data collection.

5.6.1 Data evaluation, format, flagging and control

Data management, processing and evaluation is an integral part of any monitoring plan. Data processing involves both automatic quality control and manual review components. The purpose of applying these measures, is to arrive at a data set that represents the ambient concentration of mercury and clearly distinguish against any artefacts. Periods of calibration (manual or automatic) as well as maintenance should be clearly made according to the protocol recommended in this document. All raw data must be kept in electronic form in at least two physically different and safe locations and must not be altered in any way. For rapid delivery of data, automatic processing and data quality control is required.

5.6.2 Metadata

Metadata are additional information collected about the specifics on the how, when, and where data are gathered. Mercury data without metadata are incomplete, and by themselves of limited value; at the same time metadata are of no use for data interpretation on their own. Metadata describe a dataset, what was collected, where and when the data, the methods employed, the measurements made, and other critical information. They also include information where relevant on data ownership and attribution. Key information that should be documented include the following:

- Stations Name/Code
- Country
- Latitude & Longitude
- Altitude above sea level
- Height above ground
- Station start date (YY:MM:DD)
- Collection start (YY:MM:DD)
- Time zone in UTC
- Species collected (GEM, TGM, etc)
- Unit (ng/m3)
- Time Resolution/ Exposure Period
- Integration Method (Hourly mean etc)
- Instrument (wet only or bulk for wet deposition)
- Calibration interval
- Flow rate & sampling volume
- Operator (name, email)

5.6.3 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: (1) meteorological variables such as temperature, pressure, relative humidity, wind direction and wind speed and (2) chemical variables such as carbon monoxide (biomass burning), sulphur dioxide (volcanic activity) or ozone (Arctic Mercury Depletion Events), 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-2008, 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.

5.7 Data management, analysis and evaluation

The following tools will provide Parties with a more holistic picture of the state of mercury in air by adding value to the monitoring data that is collected.

5.7.1 Local, regional and hemispheric trend analysis

- Atmospheric concentrations and wet deposition data.
- Mann-Kendall trend analysis or machine learning methods (e.g Empirical Wavelet Methods)

- For trend analysis, it could make sense to look at 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 and minimum monthly coverage of 60% (with the values giving an idea of what is deem sufficient when trends are reported in literature)

5.7.2 Data based analysis

- e.g. PMF (Probability Mass Function) analysis of sources using other measured species indicative of major Hg sources (e.g. SO_2, CO_2,).
- 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.

5.7.3 Source Receptor Relationship (SRR) based on footprints and trajectory analysis

- Analysis of source regions using backward modeling datasets generated by Lagrangian models like HYSPLIT or FLEXPART. Backward trajectories or 3-dimensional footprints of air parcels released at the measurement location and followed up to 5-20 days or 30-50 for Polar regions 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) measurements. 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.

5.7.4 Data archiving, distribution and availability

It is recognized that different networks have their own database and archiving processes and may choose to implement the guidelines and examples therein in different ways. However, databases that will be used in the effectiveness evaluation process should all contain a set of base criteria for example.

- The data producer is responsible for design and maintenance of an appropriate data Archive.
- Data and corresponding metadata should be recorded and archived with reference to Universal Time Coordinated (UTC).
- All raw data must be kept in electronic form in at least two different locations and must not be altered in any way.
- Databases should allow the data to meet the standards of FAIR principles:
    - Findable: most importantly that they are indexed somewhere that users can look them up, and that they have persistent and unique identifiers (e.g. DOIs).
    - Accessible: all potential users have access to the data using free and open methods.
    - Interoperable: formatted and with necessary metadata to allow further analyses to be performed.
    - Reusable: with enough explanatory information about how the data were collected/produced to be useful for future users.

Recommendations

- A full review of the different databases currently available to determine if it's on par to meet the needs of the Convention.
- A key consideration in data management solutions is that any data infrastructure used should be adequately resourced (both financially and in relation to thematic and IT expertise available at the datacenter) and have a reasonable prospect for secure and sustainable operations over the long-term (on the order of decades).
- Newly established stations/programs make use of existing data centres for data submission *(e.g. EBAS,* GOS4M a GEO flagship program)
- Access to all data should be facilitated, if necessary, by linking existing hubs rather than creating a new structure.
- Data submitted for the effectiveness evaluation process should be open access / free to use for public and non-commercial enterprises (e.g. GPL license) and follow the FAIR principals (Wilkerson et al., 2016)
- Another key consideration for any data management solution is that data housed in it has a clearly defined data policy.
- Data that's collected that follows sound methods and SOP's used in established networks, be considered.
- Existing networks with their own data infrastructure whose data meets the standards of the guidance, provide access for the purpose of the effectiveness evaluation to the mercury research community.

5.7.5 Chemistry transport model-based analysis

More in-depth analysis can be done using global and regional complex three-dimensional chemistry transport models (CMAQ, DEHM, ECHMERIT, GEM-MACH, GEOS-Chem, GLEMOS, WRF-CHEM). The advantage of these models is that, based on basic principles, they can be used to test hypothesis on the drivers of changes in atmospheric Hg cycling by means of scenario analysis (e.g. emission perturbations, process studies, what-if-scenarios).

Moreover, by tagging individual emissions or by means of emission perturbations these models can be used for detailed source apportionment studies. This kind of study is invaluable to determine the main sources for mercury pollution in certain regions and to find the most effective and efficient ways to reduce the environmental mercury burden. A more in-depth discussion will follow in Chapter 8 of the Guidance Document.

Chapter 6: Biota Mercury Monitoring

6.1 Introduction

Mercury emitted to the air and released to water and land can be retained in the environment for years 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 6.1**). 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 phytoplankton in aquatic systems (Lee and Fisher, 2016). As a result, meso and top predators in a food web, such as fish, reptiles, birds, and mammals, may have MeHg concentrations in their tissues that are many orders of magnitude higher than the concentrations found in the surrounding environment (often > 10^6 to 10^7 higher).

Studies show that the biomagnification and bioaccumulation of MeHg adversely affect the reproductive success of many wildlife species, impacting multiple taxa across many habitats and geographic areas of the world. 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 (Fielding et al. 2021).

Monitoring of mercury in biota can inform policy-making at various levels and across sectors. It can also help support the three objectives of the effectiveness evaluation of the Minamata Convention through the five information categories identified in chapter 4 (spatial gradients, temporal trends, modelling of cause-effect relationships, source attribution, and exposure and adverse impacts).

Ongoing biomonitoring programs provide valuable information for the effectiveness evaluation and the breadth of existing data will undoubtedly provide a basis for establishing the best bioindicators, geographic areas of interest and a baseline for estimating change, as well as for identifying data gaps. However, there is a great challenge in using existing data that were not necessarily designed for describing spatial gradients or standardized tracking of temporal trends or linking with anthropogenic mercury sources. The extent to which the existing information may be used in the effectiveness evaluation will depend on the approach agreed upon by the Parties, and the degree of flexibility in the use existing biotic mercury concentrations as baselines may be an important determinant for establishing time series and developing spatial patterns. Continuous monitoring programs that are designed to produce long times series with comparable methods and ancillary data will be particularly valuable in this effort. Furthermore, linkages of biotic Hg concentrations to 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 different source types (Dietz et al. 2019).

Fostering international collaboration and coordination among national projects will be crucial to create harmonized regional approaches and to strive, where possible, to integrate biomonitoring activities into an interdisciplinary framework to assess ecological and human health risk that can be merged to represent regional and eventually global spatio-temporal patterns.

Chapter 6 provides a brief overview of our state of knowledge with regards to existing data in fish and wildlife, monitoring programmes and databases, and challenges and data gaps, and then proposes a framework by which biomonitoring data can be used to support the effectiveness evaluation, and offers scientific and technical considerations for the selection of bioindicators and monitoring sites, as well as on ecosystem sensitivity and its importance to assessing threat and the need for ancillary measurements. Finally, this chapter explores a possible tiered approach to monitoring mercury in biota with a view to supporting the effectiveness evaluation of the Convention.

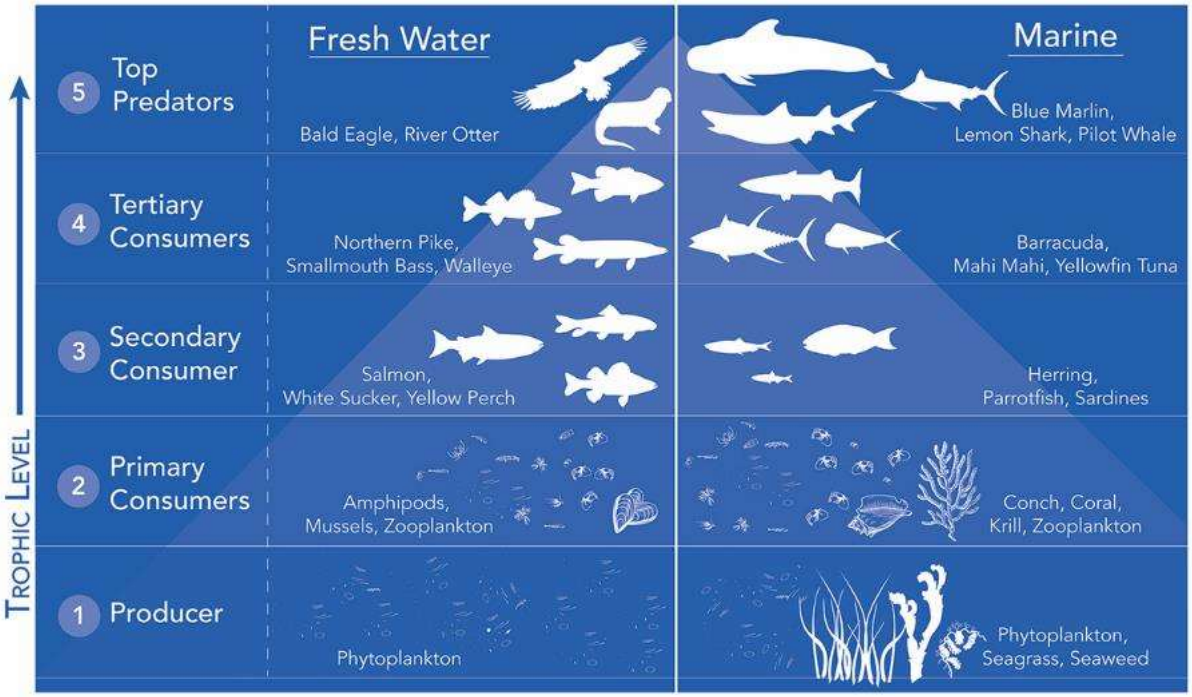


Figure 6.1. Examples of methylmercury biomagnification in freshwater and marine ecosystems.

6.2 State of Knowledge

6.2.1 Existing data in fish and wildlife

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 are summarized and exceed 530,000 data points and represent the world's oceans and continents (Appendix 1). Biotic mercury concentrations are most robust for fish, for both marine and freshwater ecosystems, and the number of analysed samples are known to be much greater when including unpublished governmental and other datasets.

Numerous studies, particularly recent ones, document adverse impacts across many fish species. In fish, adverse impacts of MeHg exposure include reproductive behavioral, and immunological impairment (Depew et al. 2012a, Scheuhammer et al. 2015; Carvan et al. 2017). Elevated methylmercury concentrations likely have significant impacts to recreational and commercial fisheries by reducing the viability and sustainability of fish populations, especially those in areas of high ecosystem sensitivity and at trophic level 4 – due to biomagnification effects.

In birds, numerous studies document reduced reproductive success, behavioral 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).

In mammals, elevated MeHg concentrations can result in biochemical changes in the brain, ataxia, and reduced reproductive output (Ronald et al. 1977; 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).

6.2.2 Existing biota monitoring programmes and databases

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.

However, one excellent example is the circumpolar Arctic Monitoring and Assessment Programme (AMAP 2011, 2018, 2021) and its associated national programs including the Northern Contaminants Program (NCP) in Canada. Based on a recent UNEP report, national programs for monitoring mercury in biota exist or were recently established in Japan, Norway, Sweden, and the United States.

The WHO Global Environment Monitoring System – Food Contamination Monitoring and Assessment Programme (GEMS/Food) includes the GEMS/Food contaminants database which collects data from national authorities on levels of contaminants, including mercury, in food. Methylmercury concentrations in these food items can be related to consumption guidelines created by countries, or globally by the Food and Agriculture Organization and the of the United Nations (FAO) and the World Health Organization Codex Alimentarius, or “Food Code”, which is a collection of standards, guidelines and codes of practice adopted by the Codex Alimentarius Commission.

The Global Biotic Mercury Synthesis (GBMS) database provides information on the global distribution of biotic mercury concentrations that affect human health and the environment. This database includes data on a wide range of species including freshwater fishes, marine fishes, birds, and marine mammals. The data included in GBMS may be useful in integrating mercury science into important policy decisions related to the Minamata Convention on Mercury and the protection of human health and the environment from the risks of mercury exposure. The GBMS database can assist with long-term monitoring of mercury in the environment. In addition, this database identifies species that are at risk of high mercury exposure and that also might represent a risk to humans due to fish consumption.[7]

Additional information on existing monitoring programmes and networks are reviewed is available in [reference to be added] and UNEP (2016).

6.2.3 Challenges and existing data gaps

The usefulness of the data on mercury concentrations in fish and wildlife that has been generated from existing programmes varies for the purposes of the Minamata Convention according to the objectives of those programmes and the associated understanding needed for those objectives.

[7] https://www.briloon.org/uploads/BRI_Documents/Mercury_Center/UNEP%20Projects/FOR%20WEB%20Mercury%20in%20Global%20Environment%20102614.pdf.

Current programmes (other than the GEMS/Food) do not cover all geographical regions and even though there may be existing biotic Hg data for many areas, those data are generally not useful for monitoring Hg because they are not the proper bioindicators, are not in specific areas of interest, are not collected with a standardized approach, or are not of interest from a temporal standpoint (depending on the objective). Information and monitoring efforts are particularly sparse in developing countries within tropical biomes and across ocean basins.

While assessments of the fate of Hg in the environment are necessary to allow for effectiveness evaluations of mercury emission reduction measures, they are often hindered by lack of data and limited understanding of the Hg cycling processes (Braaten et al., 2018). Knowledge gaps relate to re-emission of deposited mercury, catchment inputs of mercury to headwaters and retention along the aquatic continuum. Whether (or at what rate) reduced emissions of mercury to the atmosphere will result in lower loading of mercury to surface waters (and therefore reduced exposure of aquatic ecosystems to mercury) is difficult to assess, especially given the large stores of legacy mercury present in soils and lake sediments and the complexity of mercury mobilization and transport mechanisms. Given these limitations, it may be useful for monitoring programmes to include water catchment and river monitoring of aqueous mercury (and supporting water chemistry, especially dissolved organic carbon) in addition to air quality and deposition monitoring in order to aid input and output budgets of mercury as a tool in the effectiveness evaluation. The quantification of key elements in the mercury budget has proven useful for highlighting available data sources and data gaps, in addition to the highlighted limited ability to compare independent estimates of mercury fluxes obtained from various sources and with different approaches (Braaten et al., 2018).

While mercury (Hg) accumulation in biota and its impact at the individual level are well-documented, less is known about the impact on biodiversity and ecosystem services. The observed effects of direct exposure to Hg in areas of high biodiversity value and species endemism (e.g. near or downstream artisanal and small-scale gold mining operations) as well as biomagnification and bioaccumulation of methylmercury (MeHg) in species of global concern – for example, seabirds and many marine mammals – suggest that mercury may exacerbate the impact of habitat degradation and overharvesting, and contribute to global loss of biodiversity by forcing vulnerable species closer to collapse and possible extinction (e.g. Evers and Sunderland 2019; Palacios-Torres et al. 2017; Wintle et al. 2011). Studies also suggest that elevated MeHg concentrations can have significant impacts on the market value of some ecosystem services, such as recreational and commercial fisheries, by reducing the viability and sustainability of fish populations, especially at higher trophic levels due to biomagnification effects (e.g. Lusk et al. 2005).

6.3 Proposed framework for monitoring mercury in biota

The development of a monitoring framework to support effectiveness evaluation will require the understanding of ecosystem processes, the incorporation of models, and the identification of existing data on mercury exposure and how those data are interpreted in primary scientific literature on effects.

The development of a sustainable and long-term global biomonitoring framework linking existing data on mercury levels in biota to effectiveness evaluation of the Minamata Convention could focus on (i) expanding existing monitoring programs to support trend analysis and modelling, and establish linkages to Hg source types and (ii) identifying areas that have regional data gaps so new programs can generate data that meet statistical power for confidence in understanding spatial gradients, including those that incorporate ecosystem sensitivity (Evers et al., 2011b) and temporal trends (Bignert et al., 2004; AMAP 2011).

To this end, an understanding is needed of (i) key environmental attributes for mercury transport, methylation and bioaccumulation, (ii) exposure pathways and effects of mercury on target bioindicators, and (iii) contributing factors that influence mercury transformation, bioaccumulation and biomagnification which can be used to understand ecosystem sensitivity. By addressing these needs, monitoring activities can contribute to answering the overarching policy questions identified in Chapter 4 by:

1) Describing spatial gradients of biotic mercury concentrations with an emphasis to understand ecosystems sensitive to mercury input and methylation;
2) Tracking temporal trends of biotic mercury concentrations through standard bioindicators;
3) Quantifying environmental processes that affect the mobility, circulation, and methylation of mercury to inform modelling of cause-effect relationships;
4) Establishing linkages between Hg source types and key bioindicators; and
5) Estimating global exposure and associated adverse effects on biota.

By setting these actions in parallel with the identification of existing biotic Hg data and associated effects thresholds, and current biomonitoring programs at varying geographic scales, a conceptual model for integrated global biomonitoring under the Minamata Convention may be drawn (**Figure 6.2**).

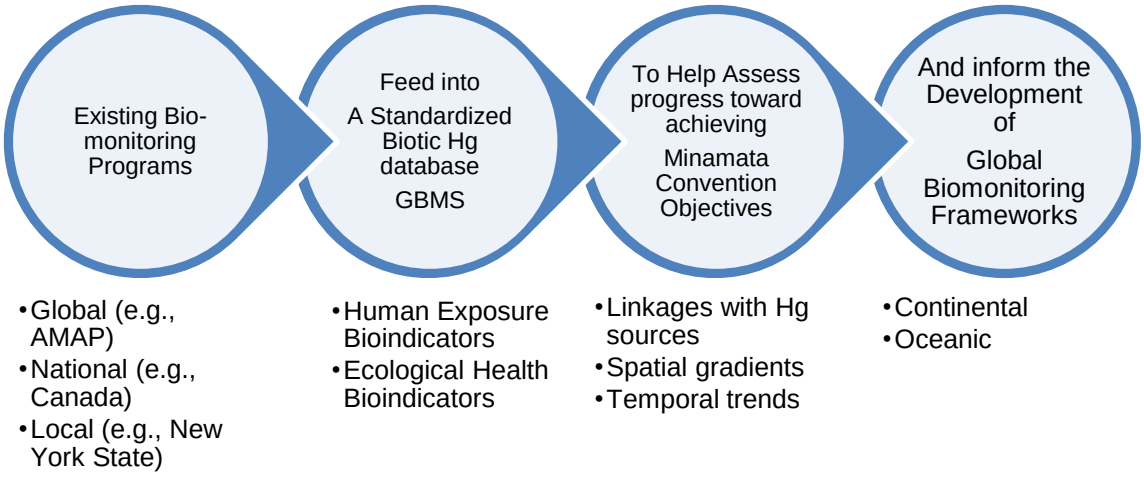


Figure 6.2. A simplistic Theory of Change for how integrated global biomonitoring frameworks can be developed to assess the effectiveness of the Minamata Convention. It begins with building directly from knowledge and information gathered from existing data and biomonitoring programs, responds to the overarching objectives of the Minamata Convention, and would provide the tools needed for standardized global Hg monitoring with acknowledgement of differences in oceanic and continental coverage.

6.3.1 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 the three key objectives of identifying spatial gradients, temporal trends, and linking with mercury source types.

A key initial step in bioindicator selection is to decide whether an organism is linked to ecological health endpoints or is relevant for human exposure assessment – which can often be combined for both purposes if carefully considered. Biota identified to best fit these two categories are well described and are categorized within their respective biomes and associated aquatic ecosystems (**Table 6.1**; Evers et al. 2016). The taxa of interest for the Minamata Convention, based on Article 19 (b), include 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, provides a strong basis for the selection of bioindicators for monitoring. Informed selection can ensure cost-effective comparability at regional and global scales.

An important aspect in combining monitoring efforts for documentation of the effectiveness evaluation of the Convention would be to define regional biological species for monitoring, to minimize the effects of species-specific physiological differences. 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. 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.

Based on the knowledge of existing biotic Hg data and only using comparable data (e.g., trophic level 4 sampling unit) for relevant terrestrial biomes and associated marine areas, a matrix of available data that can respond to overarching questions related to temporal trends, spatial gradients, and linkages with mercury source types (**Table 6.1**). While monitoring mercury in trophic level 4 biota are best for assessing impacts on the environment and potential for human exposure, understanding temporal trends is more complex. Rich supporting information on the processes that affect bioaccumulation and mercury availability is therefore needed to conduct robust trend analysis and separate anthropogenic and androgenic influences from each other. Such information is readily available from several long-term programs in the Northern Hemisphere and they have already been used for trend assessment on a pan-regional scale (AMAP 2021). 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 (<2 years of age) where age and therefore bioaccumulation is not such a big confounding factor. Trophic level 4 fish may still be used for this objective, but younger rather than older individuals should be sampled.

Table 6.1. Examples of trophic level 4 biota that could serve as bioindicators with major biomes and associated nearshore areas (based on Evers et al. 2016). Note: trophic level 4 young individual fish (<2 years) can be used for the objective of tracking temporal trends).

Terrestrial Biomes and Associated Marine Areas	Ecological Health Bioindicators			Ecological Health Bioindicators relevant for assessment of potential human exposure		
	Freshwater Birds	**Marine Birds**	**Marine Mammals**	**Freshwater Fish**	**Marine Fish**	**Marine Mammals**
Arctic Tundra and Arctic Ocean	Loons, Songbirds	Fulmars, Murres	Polar Bears, Seals	Arctic Char, Arctic Grayling	Arctic Char, Halibut, Cod	Beluga, Narwhal
Boreal Forest-Taiga and North Pacific and Atlantic Ocean	Loons, Eagles, Osprey, Songbirds	Osprey, Petrels	Otter, Seals	Pike, Walleye	Bluefish, Tuna	Pilot Whale
Temperate Mixed Forest and Pacific and Atlantic Ocean	Loons, Egrets, Herons, Eagles, Osprey, Terns, Songbirds	Osprey, Terns	Otter, Seals	Bass, Walleye	Barracuda, Mackerel, Mahi mahi, Sharks, Tuna	Pilot Whale
Tropical Rainforest and South Pacific and Atlantic and Indian Ocean	Egrets, Herons, Kingfishers, Songbirds	Albatrosses, Frigatebirds, Shearwaters, Terns, Tropicbirds	Otter, Seals	Catfish, Cichlids, Snook	Barracuda, Grouper, Mahi mahi, Sharks, Swordfish, Tuna	Pilot Whale

6.3.2 Monitoring sites

Certain ecosystem conditions (primarily those with an aquatic component, especially wetlands) can encourage the production and bioavailability of MeHg in the environment. Bacteria often produce more MeHg when moderate amounts of sulphate and low oxygen (hypoxic or anoxic) conditions are present to provide optimal conditions for the metabolic processes of the bacteria (Hsu-Kim et al. 2013, 2018).

Environmental factors such as pH, dissolved organic carbon, total suspended solids, and sulphur concentrations are important in influencing inorganic Hg input, transport, and methylation potential (Wyn et al. 2009; Gabriel et al. 2014;Gorski et al. 2008; ; Chaves-Ulloa et al. 2016; Chételat et al. 2018; Rudd et al. 2018; Schartup et al. 2018). The complex chemical conversions and cycling of Hg make it particularly challenging to predict the concentration of MeHg in fish and wildlife from concentrations of inorganic mercury in air, water, and sediments (Gustin et al. 2016, Sunderland et al. 2016; Eagles-Smith et al. 2018). Even in areas where Hg

deposition is low, concentration in biota may be disproportionately high if conditions are conducive to MeHg production and biomagnification.

The selection of monitoring locations (which could include the multi-level approach of primary and secondary sites) will need to account for the broad geographic range of methylation abilities in oceanic and continental areas, while also responding to the primary five objectives. The response from one site is not necessarily relatable to the response of a neighboring site that has different habitat characteristics.

6.3.3 Ecosystem sensitivity and its importance to assessing threat

Ecosystems are extremely 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).

As is well known, mercury is a unique contaminant, as the amount of total mercury in any given location does not necessarily correlate to ecosystem impact - largely due to the complexities of the methylation process. Ecosystems that are highly sensitive to mercury methylation may require only limited amounts of mercury to cause damage. Similarly, ecosystems with little to no sensitivity to mercury methylation may experience high levels of mercury inputs with limited impacts to nature and people. To credibly assess the potential threat of mercury to biota, biodiversity, ecosystem services and people, it is important to combine the two factors of mercury risk (from multiple inputs) and ecosystem sensitivity (to converting available mercury into its more toxic, bioavailable form). The structure of this approach to mercury threat assessment and the associated information is provided in **Table 6.2**.

Table 6.2. Description of mercury threat assessment categories, key variables, and comments on indicators being used for each. Mercury inputs are categorized according to UNEP (2019) nomenclature. Degradation caused unsustainable land management practices as well as by agriculture have been included as they also have a direct influence on the availability of mercury in the environment.

Threat assessment category	Key variable	Indicator comments
Ecosystem Sensitivity	Land cover types	There are multiple land cover types – each will be weighted according to their relative methylation potential. For example, deserts have little to no methylation potential in comparison to tropical forests.
	Wetland types	There are multiple wetland types – they are often nested within land cover types. Each will be weighted according to their relative methylation potential. For example, lakes or rivers generally have lower methylation potential than bogs, peatlands, and swamps.
	Habitat characteristics	Additional factors related to habitats that influence methylation potential that are not mapped by land cover or wetland type. For example, areas with higher soil organic carbon (SOC) have higher methylation potential than areas with lower SOC levels.
Mercury Inputs/Risk	ASGM activity	*Artisanal and Small-scale Gold Mining (ASGM)* – is the largest source of global mercury emissions. ASGM activity is highly

		variable in its distribution and relative amount of activity, and is of global and local importance.
	Power Generation	*Stationary Power Sources* – Globally, stationary power sources contribute the most mercury emissions of all large-scale industrial activities. Of these stationary power sources, coal-fired power plants are the largest single contributor. Quantifying the amount of industrial activity within watersheds is an important indicator of local mercury emission potential.
	Large-Scale Industry	*Large Scale Mining* (LSM) – LSM is the largest mercury emitter of all large-scale industrial activities, with non-ferrous metal production at the top of the list. Cement production is a close second. Tracking the amount of large-scale industrial activity in each watershed is an important indicator of mercury contamination of air and water.
	Intentional Uses	*Waste Disposal* – Emissions associated with mercury-added product waste is the largest source of emissions in this sector. Documentation the location of waste disposal sites within watersheds is important for guiding monitoring of potential contamination.
	Agricultural activity	*Rice paddy fields* – certain agricultural activities, such as planting rice in paddy fields, can change Hg input. Rice paddy fields are a dominant agricultural land use throughout Asia and have been identified as important sites for methylmercury production in the terrestrial ecosystem and a primary pathway of MeHg exposure to humans in mercury mining areas.
	Degradation caused by unsustainable land management practices	*Deforestation* – is likely the most important process driving mercury releases to the water in tropical forest regions. Incorporating the amount of deforestation in a watershed is an important indicator of disturbance and potential mercury releases.
		Soil Erosion – is the primary process that carries mercury from the land into freshwater ecosystems. There are many factors influencing soil erosion, and it is responsible for releasing mercury into the air and water. It is particularly pronounced where ASGM activity and deforestation occur but is not limited to these areas. Soil erosion is a good proxy for habitat degradation, and an important indicator of the mercury transport process in terrestrial and freshwater environments.
		Fire Frequency – 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.

1055 The information available to create a robust global threat assessment requires combining a mix
1056 of discrete categorical and continuous data. Methods are being developed that ensure both
1057 consistency and transparency in this approach, as well as the ability to down-scale this
1058 approach for application at regional and local-levels to make use of critical information not
1059 available at a global scale (e.g., point-source data). As water is a major pathway for mercury

through ecosystems, and watersheds offer a justifiable, hierarchical approach to assessment across many spatial scales, evaluating the threat of mercury via watersheds has emerged as an important part of this approach (Evers and Sunderland 2019). Creating new assessments of watershed risk, sensitivity and threat of mercury impacts to nature and people 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 synergistic 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).

While the precision of modeling the sensitivity of ecosystems is still in its early stages and therefore introduces its own uncertainty, there will be a level of contribution for decision-making such as prioritization of sites for effectiveness evaluation of the Minamata Convention. At its most basic level, it will be possible to at least categorically identify ecosystems or areas that tend to have relatively low (e.g., deserts) or high sensitivity (e.g., wetlands), low or high inputs (e.g., from global Hg deposition), and low or high threat (e.g., low sensitivity and low inputs vs. high sensitivity and high inputs).

6.3.4 Ancillary measurements

Ancillary measurements often collected with biota mercury data include species, total body length, weight, fat levels, and stable isotopes of carbon ($\delta^{13}C$) and nitrogen ($\delta^{15}N$). Stable isotope measurements in biota assist with identifying changes in food web structure and 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 efforts related to the Minamata Convention or driven by large-scale factors such as changes in food web complexity, trophic position of biota, climate change, overfishing, and biogeochemical conditions. AMAP's mercury monitoring programs for biota include these ancillary measurements and 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). A range of low- and high-resolution approaches to collect ancillary data can be maintained depending on the needs of the monitoring program (**Table 6.3**).

Table 6.3. Example of low- and high-resolution approaches for sampling and analyzing biota in conjunction with ancillary measurements.

Resolution Tier	Matrix	Ancillary measurement examples
Low Resolution	Biota (e.g., muscle tissue)	Species, body length, mass, spatial coordinates
High Resolution	Seawater	pH, dissolved oxygen, salinity, temperature, depth, mercury ($\delta^{202}Hg$) stable isotopes
	Surface Sediments (e.g., top 2 cm)	Carbon ($\delta^{13}C$), nitrogen ($\delta^{15}N$), mercury ($\delta^{202}Hg$) stable isotopes, depth, temperature
	Biota (e.g., muscle tissue, blood, keratin-based tissue)	Species, body length, mass, spatial coordinates, carbon ($\delta^{13}C$), nitrogen ($\delta^{15}N$), mercury ($\delta^{202}Hg$) stable isotopes

It is likely a tiered approach will be beneficial for evaluating the effectiveness of the Minamata Convention when considering the existing and new data for biota as they will have different levels of resolution and quality assurance and quality control (QA/QC).

Specifically, 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. 2019b) 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.

Braaten et al. (2019) argue that to evaluate the effectiveness of the Minamata Convention, there is a need for identification of legacy Hg sources and for separating these sources from long-range atmospheric sources of Hg. Abiotic media, such as water and sediment/soils, could be included with biomonitoring to provide further support in understanding temporal trends and spatial gradients. Abiotic media should not be used exclusively because of interpretative limitations and uncertain connectivity with associated biota, especially high trophic level species.

For each terrestrial location, this should include lake and catchment morphology, pollution deposition patterns, and local pollution history. For each animal species data must include length, weight, and age. Samples (i.e., fish muscle) for determination of total Hg concentrations, should also be analyzed for stable isotopes of nitrogen (δ15N) and carbon (δ13C) for a better understanding of trophic position and energy sources.

6.3.5 Tiered approach to monitoring

For effectiveness evaluation of the Minamata Convention, an integrated tiered approach to monitoring of mercury in biota is proposed: tier 1) documentation of environmental levels with a wide spatial representation; tier 2) intensive monitoring of temporal trends in well-documented sites that also focus on collection of supporting data; and tier 3) data collection to test cause-effect relationships.

For a low resolution tier, a minimum set of parameters can be defined (e.g., lake coordinates, fish Hg concentration (muscle, wet weight; fish sampling restricted to the same season), fish weight and length; preferably fish age, sex and maturity stage, and supplementary information on the lake and catchment (i.e. size, elevation, and land cover and use (natural, agriculture, developed). Higher resolution interests would include time series that would preferably be established for lakes with well-known pollution loading (local catchment sources, or from long-range transported air pollution). Inclusion of lakes with only external (long-range transported) pollution loads is crucial for effect evaluation in remote areas (boreal, subarctic, arctic). In addition to low resolution parameters further ancillary data should be included such as stable carbon and nitrogen isotopes for fish, water chemistry (including Hg species, total and methyl-Hg, TOC, pH (all mandatory), and nutrients (preferred)). For high resolution scenarios, information on diet (from fatty acids, for example), stable isotopes of lower foodweb organisms, to provide data on food web structure.

The tiered approach would rely on 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 stations would be needed in order to represent ecosystem sensitivity (which can be mapped with a certain level of precision and certainty) and have a mixture of background/reference sites together with hotspot sites with well-known local sources. Then, proper 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 programmes, is shown in section 3 of the **Annex** to this guidance.

6.4 Possible frameworks for implementing the tiered approach

In view of the current gaps in mercury monitoring data in biota, the need to understand source types and their ecosystem linkages, spatial gradients, and temporal trends, and the interest of using bioindicators relevant for human exposure assessment and ecological health, two overarching biotic Hg monitoring frameworks are proposed herein– one for landscapes (continents) and another for seascapes (oceans) – to implement the tiered approach and support global biomonitoring efforts under the Minamata Convention.

The oceanic and continental frameworks would integrate existing and ongoing data collection efforts with new sampling with a view to providing relevant information in a cost-effective way. This would be achieved by building upon existing long-term efforts and identifying data gaps so further global assessments may be conducted. Below is a brief overview of the main aspects of two proposed frameworks. Further information may be found in the technical background document on biota monitoring (Evers and Sunderland 2019).

6.4.1 Oceanic Framework

The cycling and movement of Hg in the world's oceans varies by hemisphere, basin and juxtaposition with the continental land masses. Therefore, Hg concentrations in fish, sea turtles, birds, and marine mammals vary significantly across geographic areas.

The proposed Oceanic Framework for global monitoring of mercury in biota is composed of three steps, each further divided in three sub-steps. Step 1 starts with a characterization of ocean basins and is followed by collecting data on fish and other biota from existing mercury monitoring programmes. Step 2 is focused on the selection of key bioindicators (e.g., Drevnick et al. 2015, 2017), and Step 3 focuses on analyzing existing data, determining optimal sample sizes, and collecting and analyzing samples (**Figure 6.3**; Evers and Sunderland 2019).

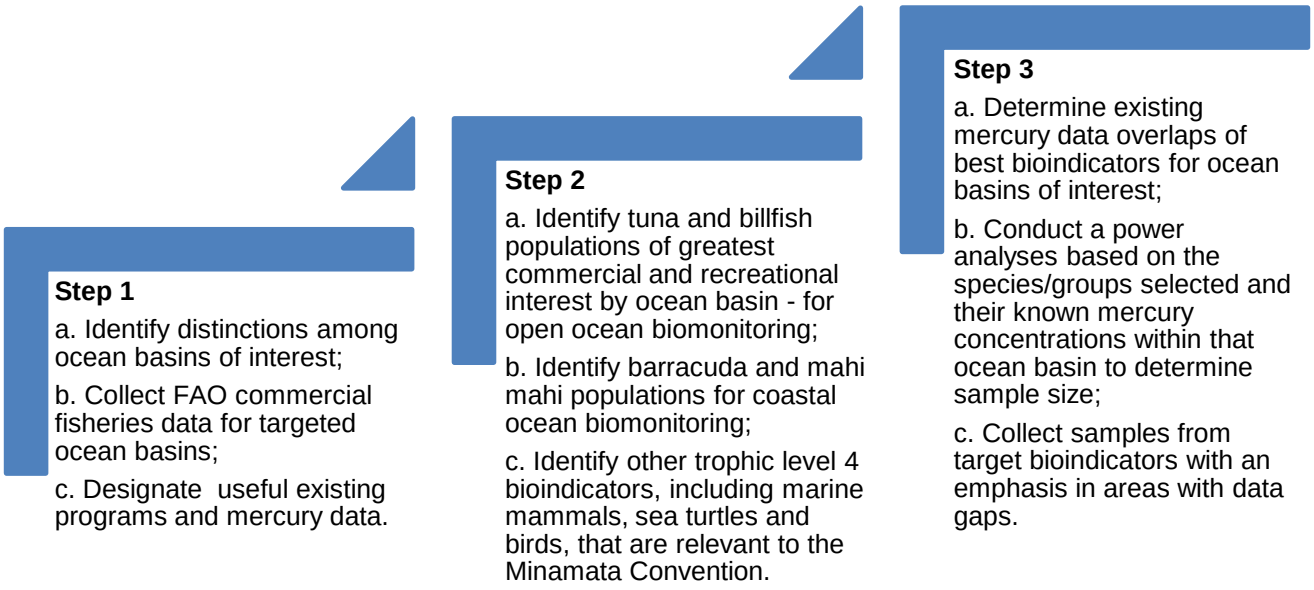


Figure 6.3. Stepwise components for developing an oceanic approach for mercury biomonitoring.

6.4.2 Continental Framework

To understand the complexities of a landscape and its ability to methylate mercury and make it available in the food web, biota monitoring under the Continental Framework would require multiple defined steps (**Figure 6.4**). In this proposed framework, Step 1 would focus on mapping the best locations for global Hg monitoring, with a focus on wetlands and other sensitive ecosystems. Mercury methylation is highest in wetlands – and, potentially greatest in estuarine wetlands such as mangroves, and peatland bogs that are generally acidic. Forested areas are also an important habitat for increasing dry deposition rates of atmospheric Hg in temperate (Driscoll et al. 2007) and tropical (Gerson et al. In Review) ecosystems, while agricultural areas tend to dampen methylation rates (Chen et al. 2008) with the exception of rice paddies (Abeysinghe et al. 2017). Step 2 would focus on identifying overlap in existing data with areas that are of particular relevance for monitoring, including ASGM sites, important freshwater fishing areas and key biodiversity areas, and identify potential bioindicators for harmonized sampling and comparison (global pilot study example is Buck et al. 2019). Step 3 would select the best ecosystem sensitivity spots based on the two previous steps to establish spatial patterns and temporal trends, as well as linkages to Hg sources.

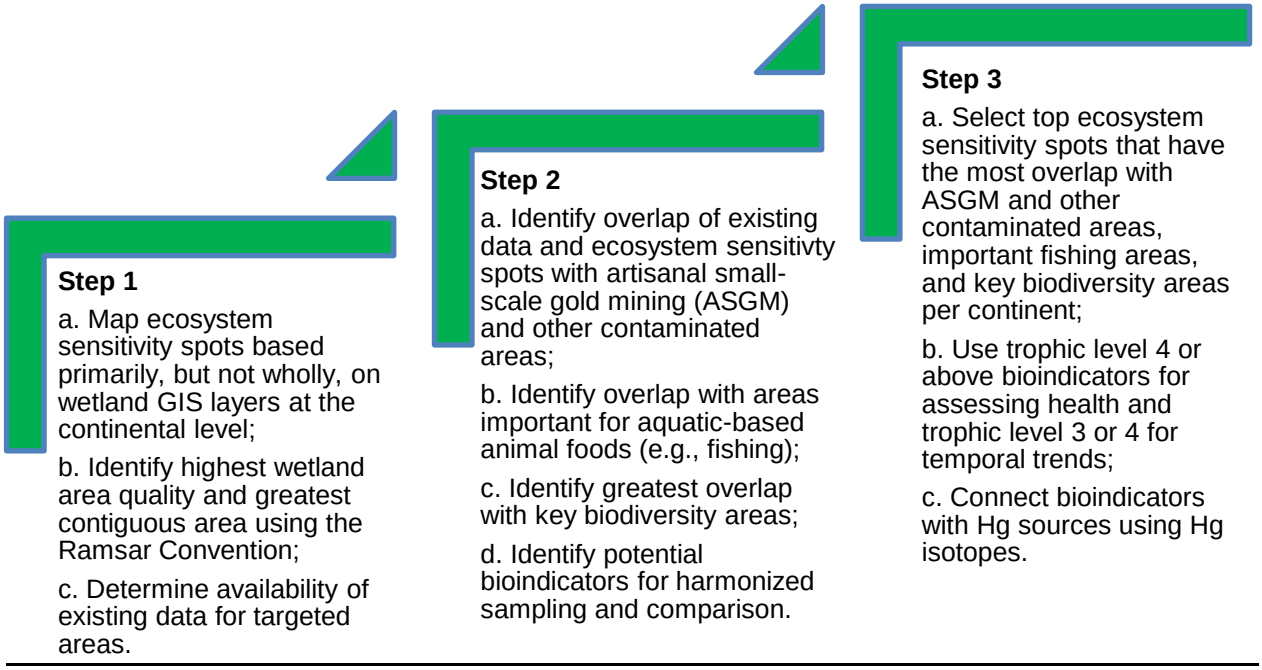


Figure 6.4. Stepwise components for developing a continental approach using biota for Hg monitoring.

6.5 Field sampling, laboratory analysis and data management

Field sampling and laboratory analysis depend on a number of factors, including the taxa chosen, their habitats and abiotic conditions, objectives of the monitoring, etc. Protocols for collecting samples from biota are available for all tissue types, and examples of the proper selection of tissue type are well-established with associated information about the percent methylmercury content in the tissue and the preferred type of tissue preparation. Most muscle, blood, egg and keratin-based (e.g., scutes, feathers, and fur) tissues primarily contain methylmercury. This is important for the simpler and more cost-effective laboratory analyses of total mercury concentrations that can assume 95% or more methylmercury content (with some exceptions).

The timing of biota sampling at monitoring locations varies according to the objectives, habitats/ecosystems, and chosen bioindicators. The fraction of mercury retention in the atmosphere, soils, and waters can vary over days to centuries (**Figure 6.5**). Therefore, knowledge of mercury retention in habitats that biota is sampled from will be important for understanding spatial gradients, temporal trends, 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 spatial and temporal patterns. Likewise, knowledge of mercury retention in habitats that biota is sampled from will be important for understanding spatial gradients, temporal trends, and linkages with mercury sources. For linkages to mercury source types, mercury isotopes are important. 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 every other year for one or two decades (Riget et al. 2011; Eagles-Smith et al. 2016a; Sunderland et al. 2018; Evers et al. 2020). Even longer time scales are needed to examine changes in biota mercury concentrations in large water bodies, such as in bluefin tuna in the North Pacific Ocean (Sunderland et al. 2009).

A more in-depth overview of the requirements for optimal sampling, analytical measurements, and data collection and analysis is available as supplementary material.

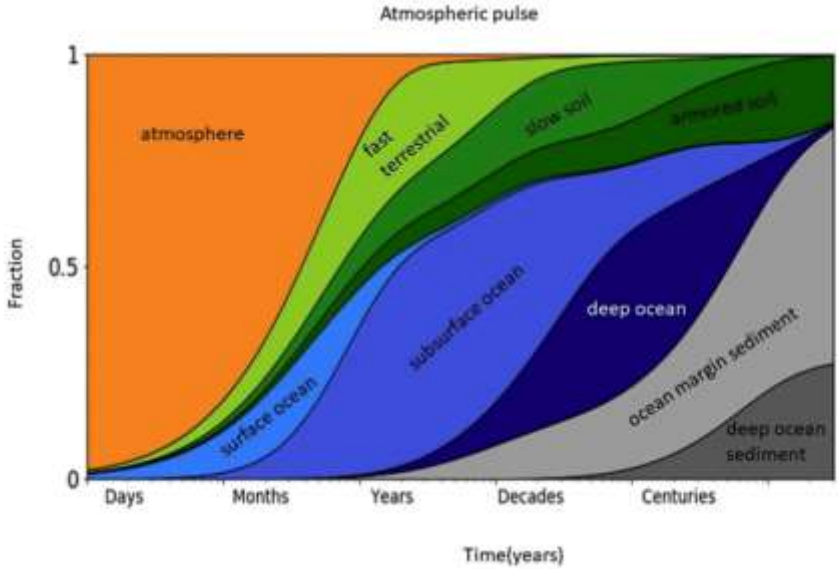


Figure 6.5. Retention of mercury fraction (0 to 100%) over time (days to centuries) in various compartments of the atmosphere, landscape (e.g., soils), and waterscape (e.g., ocean waters and sediments). Graph from Elsie Sunderland, Harvard University, Massachusetts, United States.

7.0. Human biomonitoring

7.1. Introduction

Understanding human exposures to chemical hazards through biomonitoring activities is important for scientific and regulatory purposes (WHO Regional Office for Europe 2015; Louro et al. 2019). 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 (World Health Organization (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 minimal cost (e.g., field work, participant sampling, and biomarker analysis can likely be done for under $100 USD/participant (Evers et al. 2016).

Human biomonitoring data can help address core policy questions that drive effectiveness evaluation (Table 4.1). 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. These core policy questions that drive effectiveness evaluation can 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 and inter-sectoral 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 7.1**). Chapter 7 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 effectiveness evaluation, 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 Convention Articles (e.g., Articles 4, 7, 14, 16-19, 21).

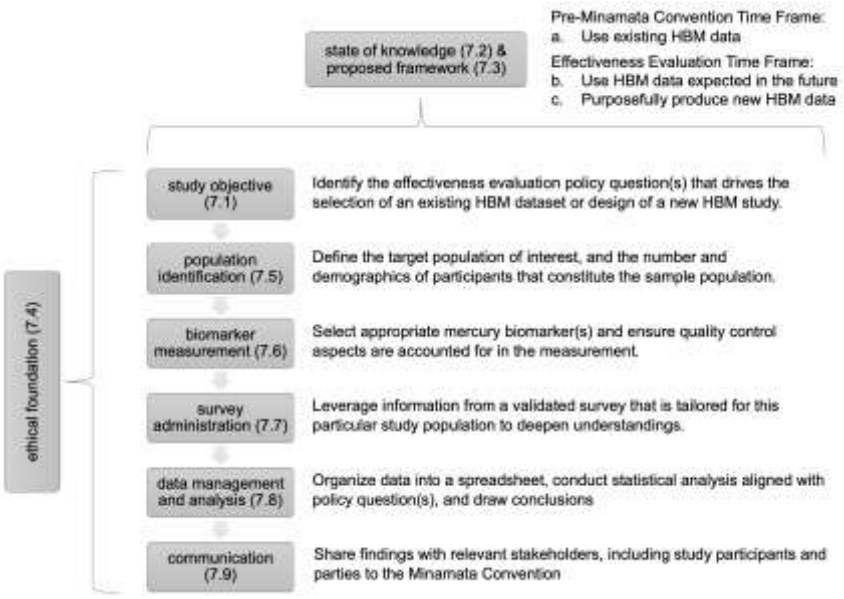


Figure 7.1. Proposed approach for using human biomonitoring data (HBM) for the purposes of effectiveness evaluation. The proposed approach lists key elements that need to be considered when using existing HBM data or when planning a new HBM study. The numbers in parenthesis in the shaded boxes refer to chapter sections that offer more details.

7.2 State of Knowledge

7.2.1 Existing Data

To assess our current understanding of human exposures to mercury, a systematic search of the recent (2000 to 2018) literature identified 312 high-quality 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 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 of which parties need 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 limited to high 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 Program (AMAP) and the DEMOnstration of a study to COordinate and Perform Human biomonitoring on a European Scale (DEMOCOPHES) effort.

7.2.2 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).

7.2.3 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 Health Survey (GerES), Republic of Korea's National Environmental Health Survey (KoNEHS), US National Health and Nutrition Examination Survey (NHANES)), and these will be dependable programs for effectiveness evaluation. 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 dependent 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 without existing data sources should consider a harmonized approach to launch new biomonitoring studies. A good starting point is the recent guidance from the WHO to characterize prenatal mercury exposures (World Health Organization (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. Each country would own its data. 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",[8] this WHO protocol was piloted between 2015 and 2017 in diverse settings. Examples of mercury sources include 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 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.

7.3 Proposed Framework

This section outlines a proposed framework in which parties can use human biomonitoring data for effectiveness evaluation. Driven by policy questions that drive effectiveness evaluation

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

(Table 4.1), there are three main components to the proposed framework that the parties should 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 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 in a harmonized way so that they are purposefully designed to fill data gaps, build capacity, and address effectiveness evaluation.

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.

Information in this chapter provides essential guidance (and links to key resources) for parties to consider in terms of using existing and generating new human biomonitoring data for effectiveness evaluation. The guidance presented in this chapter 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 overarching policy relevant questions, albeit on different scales.

7.4 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. Given that human biomonitoring may focus on vulnerable populations, participatory engagement of pertinent stakeholders (e.g., study participants, community leaders, health care providers, regional authorities) is necessary not only for ethical 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 (Council for International Organizations of Medical Sciences) 2016). In addition, parties may consult papers on legal, ethical, and social issues pertaining to human biomonitoring from the European Human Biomonitoring Initiative (HBM4EU) (HBM4EU 2018a) the Canadian Health Measures Survey (Day et al. 2007), and the World Health Organization's recent guidance on ASGM (WHO 2021a).

7.5 Human Population Group

7.5.1 Identification of Target Population

All human populations worldwide are exposed to some amount of mercury (UNEP and World Health Organization (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 in a given country is a decision of the responsible party, and should be driven by their interests in consideration of policy questions that drive effectiveness evaluation (Chapter 4). For example, some parties may choose to focus on the general population while others may choose to focus on a specific vulnerable group (e.g., pregnant women, ASGM, 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 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 (UNEP and World Health Organization (WHO) 2008).

Human exposures 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 (UNEP and World Health Organization (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 6). Sampling of freshwater fish and seafood has found widespread methylmercury contamination, with some widely-consumed predatory species (e.g., tuna, swordfish, grouper, mackerel; GEMS/Food Contaminants n.d.) being among the most highly contaminated. 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 exposure for some communities to both organic and inorganic mercury (Rothenberg, Windham-Myers, and Creswell 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 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 and local communities) 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 fish consumers including sport fishers and Indigenous

Peoples and local communities, and users of mercury-containing skin-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 5 and 6).

7.5.2 Identification of Sample Population

Upon identifying a target population for investigation (7.5.1), the researchers would ideally sample all individuals from this target population, though achieving this is impractical (e.g., too many individuals to sample, and this is prohibitively expensive and takes too much time). Instead, researchers will sample a subset of the target population to realize a sample population. 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.

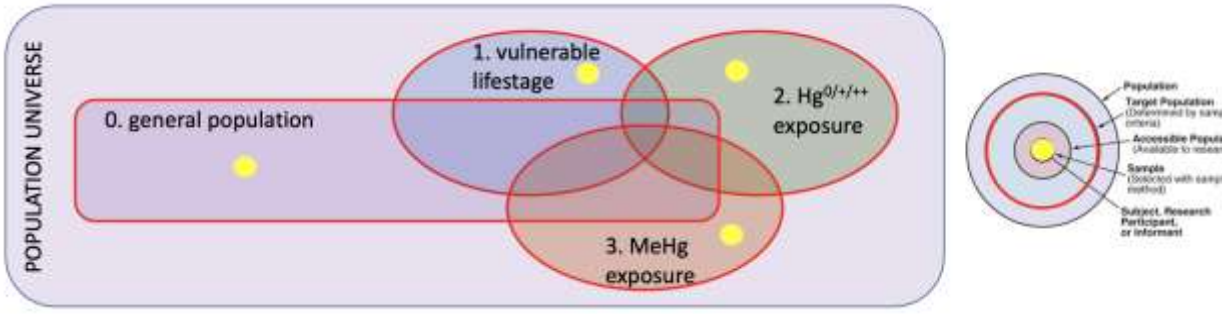


Figure 7.2. Population groups to consider. Within a country, exposures to mercury will be realized by all inhabitants (i.e. population universe), including members of the general population as well as members who are deemed vulnerable because of their lifestage or exposure situation. These population 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). Once parties select a specific target population to focus upon (driven by their interests in consideration of the policy questions that drive effectiveness evaluation), steps need to be taken to help ensure that the sample population (yellow circle) 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 (World Health Organization (WHO) 2018a), HBM4EU (HBM4EU 2018b, 2018c, 2017b), 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 (World Health Organization (WHO) 2018a). In addition, the HBM4EU statistical plan (HBM4EU 2017b) 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 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.

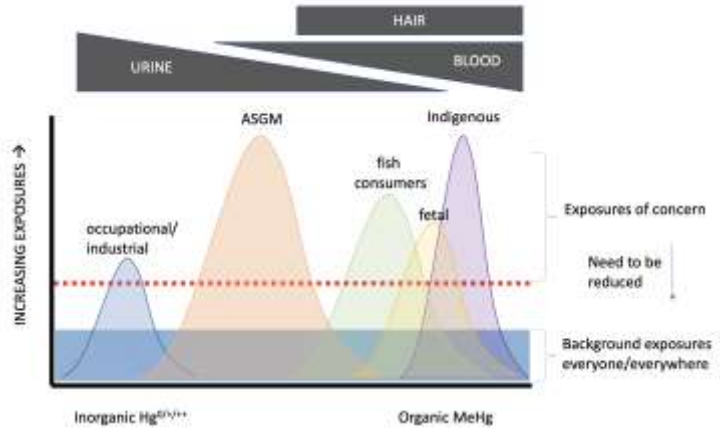
The nature by which participants are recruited and studied should be carefully detailed following guidance from the STROBE (Strengthening the Reporting of Observational Studies in Epidemiology) initiative (Vandenbroucke et al. 2007). Ideally the sampling process is free from any biases, and participants 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.

7.6 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 (World Health Organization (WHO) 2018b) and HBM4EU (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 7.3. 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.



7.6.1 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. 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 contaminated sites care is needed in the interpretation of total mercury levels in hair. In such settings carefully analyzing the hair for methylmercury, rather than total mercury, gives a better measure of dietary exposure especially when coupled with quality survey instruments.

7.6.2 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 a relatively easy, non-invasive, and cost effective, and there are good protocols available from WHO (World Health Organization (WHO) 2018b) and HBM4EU (HBM4EU 2018d, 2019).

7.6.3 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). 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 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.

7.6.4 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). When multiple biomarker measurements are taken from a given individual, a deeper exposure assessment can be performed. 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 (World Health Organization (WHO) 2018b) and HBM4EU (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 7.7) that collect pertinent information on the study population and exposure sources.

7.6.5 Biomarker Measurements

A number of analytical methods (e.g., cold vapour atomic absorption spectrometry and cold vapour atomic fluorescence spectrometry 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 (World Health Organization (WHO) 2018b) and by HBM4EU(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. 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, measures of total mercury content in a given biomarker will suffice in nearly all 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. It is estimated that setting up such a laboratory infrastructure can be had for under $100K USD, that annual operation costs could be ~$50-100K USD (for analysis costs and lab technician salary), and that such a lab could yield thousands of mercury measurements for both human biomarkers and fish/wildlife samples. Stakeholders with deeper interests, understandings, and resources can take advantage of more sophisticated yet complex approaches to human biomonitoring that integrate mercury speciation or mercury isotope analysis.

7.7 Survey Protocol

Combining the results of mercury biomarker measurements (Section 7.6) 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 (World Health Organization (WHO) 2018a)) and HBM4EU (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. The survey data should also be amenable for capture into an electronic format that can be integrated into a data management program

7.7.1 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. 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 could consider human biomonitoring efforts in geographic sites where biota are being sampled to maximize efficiencies and data quality (Chapter 6). Detailed protocols for developing dietary surveys are available from the UN/WHO (UN Environment and World Health Organization (WHO) 2008) and the US EPA (US EPA 2016), and the HBM4EU has a comprehensive dietary questionnaire available for parties to adapt (HBM4EU 2020a).

7.7.2 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 (UN Environment and World Health Organization (WHO) 2008; Eagles-Smith et al. 2017; Ha et al. 2017; ATSDR 1999).

Identification of a target population based on these particular exposures (Section 7.3.1) should trigger the need to include screening level assessment surveys to deepen understanding of potential exposures. Example of relevant screening level assessments for mercury are available from WHO (World Health Organization (WHO) 2018a) and HBM4EU (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 (World Health Organization (WHO) 2021b). There is also a survey from a UNIDO/UNDP/GEF-sponsored initiative that is often used (Veiga and Baker 2003), 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).

7.8 Data Management and Analysis

7.8.1 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 the "baseline" based on input by parties and other Minamata Convention stakeholders. 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 are encouraged to consider recent guidance from the WHO on a harmonized approach for conducting new biomonitoring activities (World Health Organization (WHO) 2018a).

7.8.2 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 7.5 (Populations) and 7.7 (Surveys) should be consulted, along with resource documents from the WHO ((UNEP and World Health Organization (WHO) 2008; World Health Organization (WHO) 2018a) and the STROBE (Strengthening the Reporting of Observational Studies in Epidemiology) 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 7.6 (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.

7.8.3 Database

Paragraph 1 (d) of article 17 of the Convention calls for parties to facilitate the exchange of epidemiological information, in close cooperation with the World Health Organization and other relevant organizations, as appropriate. In line with that article of the Convention, the compilation and exchange of data via a centralized database should be considered. This database must abide by the FAIR principles (Wilkinson 2016). The HBM4EU offers guidance on managing such databases (HBM4EU 2017a)

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, and precision), and mercury values (count (n)), percentiles including 10^{th}, 25^{th}, 50^{th}, 75^{th}, 90^{th}, and 95^{th} 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 2017b)). Strategies for dealing with missing data and measures below detection limits are provided in the HBM4EU statistical analysis plan (HBM4EU 2017b). Key information from surveys (e.g., dietary intake values; occupational practices; other exposure sources) need 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. The HBM4EU statistical analysis plan (Section 11.2.2) 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 2017b).

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 human subjects research 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 (World Health Organization (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 2017b).

7.8.4 Data Analysis

Statistical analysis of human biomonitoring data may help address core policy questions that drive effectiveness evaluation (Section 4.1). 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 2017b). Five key statistical analyses are listed below that align with the overarching policy-relevant questions. More sophisticated aspects of Data Analysis (especially modelling) are provided in Chapter 8, 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. The HBM4EU statistical analysis plan (Section 6) provides detailed guidance on temporal trends analysis (HBM4EU 2017b).

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. While challenging, there are success stories to draw inspiration from. For example, and as discussed in the Global Mercury Assessment 2018 report: 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: The ultimate goal 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.

7.9 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 (World Health Organization (WHO) 2018a). Parties may also decide on if (and how) the data is used for risk management.

Chapter 8. Cross-media data management, modelling and analysis

8.0 Introduction – use of monitoring data from different media

Chapters 5, 6 and 7 provide guidance on the collection, management and analysis of data in air, biota and from human biomonitoring. By analyzing monitoring data, spatial and temporal 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. 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 spatial and temporal trends and 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 largely on known or suspected physical processes, mechanistic models explicitly representing these processes are a valuable tool for interpretation of monitoring results and can thereby contribute to the effectiveness evaluation. 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. When statistical models are used together with mechanistic models, improved predictive capacity can be obtained. This is particularly important when models are used for policy evaluation applications and uncertainties need to be minimized. In most cases, attribution of observed trends to specific drivers such as direct anthropogenic mercury releases, legacy mercury, natural process-driven releases, and non-mercury environmental or behavioral 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 be influenced by human behavior. This makes cross-media analysis involving both mechanistic and statistical modeling in all relevant media necessary in order to fully evaluate effectiveness of the Convention.

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 trends can also be used as input to models which can explicitly connect those trends to outcomes in other media, or to models which can attribute those trends to specific sources or drivers. **Tables 8.1** and **8.2** summarize, for monitoring data and ancillary observational data respectively, the data, metadata, and other information that can facilitate cross-media analysis and modeling.

Table 8.1. 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	Total or gaseous elemental mercury levels; measurement/ method uncertainty; Wet deposition,	Latitude; longitude; altitude; Sampling time, frequency, duration; averaging methods; sampling method;	Proximity to known point sources; type (urban/regional/background); meteorological variables; measurement/method uncertainty;	● Spatial variations ● Temporal trends ● Atmospheric model evaluation ● Input for local-scale modeling ● Back-trajectory analysis ● Bottom-up attribution analysis
Air - Tier 2	Air - Tier 1 and Speciated reactive mercury; high-resolution PBM and GOM; Dry deposition of mercury; mercury throughfall	Air - Tier 1	Air - Tier 1 and deposition of Sulfate; Land Cover; Land Use; Leaf Area Index; Air Quality Tracers (e.g., SO2, CO2, CO, PM2.5, O3)	Air - Tier 1 and ● Estimate air-ocean and air-terrestrial mercury exchange ● Covariate profiling ● “Top-down” SRR attribution
Air - Tier 3	Air - Tier 2 and mercury isotopes; additional speciation measurements	Air - Tier 1	Air - Tier 2	Air - Tier 2 and ● Combined “top-down” and “bottom-up” 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; ethics board	diet info; age; sex; known occupational and other exposures; education, socioeconomic status, amalgam status; additional measures of central tendency (variance) including mean (SD) and geometric mean (95% CI); indication of data normality; measurement/method uncertainty	● Spatial variations ● Temporal trends ● Exposure model evaluation ● Input for local health impact / risk assessment modeling ● Guideline value exceedance statistics
Human - Tier 2	Human - Tier 1 or cord blood (10^{th}, 25^{th}, 50^{th},	Human - Tier 1	Human - Tier 1 and associated relevant biota measurements	Human - Tier 1 and ● “Top-down” exposure attribution

	75th, 90th, and 95th percentiles); optionally methylmercury ; mercury isotopes			
Human Tier 3	Human Tier 2	Human - Tier 1	Human - Tier 2	Human - Tier 2 and ● Combined “Top-down” and “bottom-up” exposure attribution
Biota - Tier 1	Tissue/organ mercury and/or methylmercury levels; measurement/ method uncertainty; distribution statistics or quantiles	Geolocation or water body name; Spatial coverage; sampling time period; method info; tissue/organ type; habitat; wet or dry weight	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; water temperature, DOC, pH, TSS, nutrient concentrations;	● Spatial variations ● Temporal trends ● Input for local exposure modeling
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 in sediment; GEM in air; wet deposition; meteorological data	Biota - Tier 1 and ● “Top-down” biota mercury attribution ● Food web model evaluation
Biota - Tier 3	Biota - Tier 1	Biota - Tier 1	Biota - Tier 2 and 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

Table 8.2. 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; measurement/method uncertainty; surface fluxes;	Latitude; longitude; depth; Sampling time, frequency, duration; averaging methods;	Presence of known point sources; soil horizon; land use; carbon concentrations	• Terrestrial model evaluation • Input for local-scale modeling
Vegetation	Mercury levels; measurement/method uncertainty; exchange fluxes; litterfall fluxes;	latitude; longitude; Sampling time, frequency, duration; averaging methods;	vegetation type; NDVI; carbon fluxes	• Terrestrial model evaluation • Input for local-scale modeling
Food items and other products	Methylmercury and total mercury levels; statistical distribution information	country/region; Sampling time period;	Consumer population; exposure type (diet, skin, etc.)	• Input for exposure modeling
Freshwater	Mercury levels; methylmercury levels; measurement/method uncertainty;	latitude; longitude; depth; Sampling time, frequency, duration; averaging methods; water body name	dissolved and particulate carbon concentrations; temperature;	• Input for food web modeling
Ocean	Mercury levels; methylmercury levels; measurement/method uncertainty;	latitude; longitude; depth; Sampling time, frequency, duration; averaging methods; water mass name	dissolved and particulate carbon concentrations; nutrient concentrations; temperature; salinity; dissolved oxygen	• Ocean model evaluation • Input for food web modeling
Sediment	Mercury levels; methylmercury levels; mercury accumulation rates; measurement/method uncertainty;	latitude; longitude; depth; Dating info; dating method	accumulation rates; total organic carbon; grain size	• Input for watershed modelling • Input for food web modelling • Mass balance model evaluation
Snowpack	Mercury levels; exchange fluxes; measurement/method uncertainty	latitude; longitude; Sampling time, frequency, duration; averaging methods; sampling methods	snow depth; accumulation rates; snow density	• Atmospheric model evaluation • Input for local-scale modeling

8.1 General Considerations

Some media-specific analyses are outlined in Chapters 5, 6, and 7, and can be useful tools to inform effectiveness evaluation via single-medium monitoring data. Chapter 5 discusses 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. Chapter 6 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 7 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 to evaluate Convention effectiveness.

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 modeled 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.

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 effectiveness evaluation and provide more robust answers to questions of trend identification and attribution.

8.1.2 Types of models that can support effectiveness evaluation

8.1.2.1 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 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 emission inventories (small-scale spatial and temporal effects)
- *Readiness:* Multiple available models

8.1.2.2 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:* Reliance on atmospheric model input
- *Readiness:* Multiple available models

8.1.2.3 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, toxicokinetic parameters describing human mercury metabolism can be used, with important uncertainties arising from differences in methylmercury uptake and elimination across individuals. Regression-based models (including GAM/GLM 8.1.2.8) 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 modeled 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

8.1.2.4 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 modeling.

- *Strengths:* Bottom-up attribution, specificity
- *Weaknesses:* Some parameters difficult to obtain (e.g. food web structure, water biogeochemistry)
- *Readiness:* Further study is necessary to determine how current food-web model would be best used in the effectiveness evaluation processes

8.1.2.5 Mass balance models (AKA 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 modeling 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

8.1.2.6 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. This type of modeling 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 modeling 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

8.1.2.7 Terrestrial models

Terrestrial models represent the exchange of mercury between vegetation and soil reservoirs via processes associated with biochemical transformations of carbon 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. 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
- *Readiness:* Emerging applications for multi-media model coupling

8.1.2.8 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 predictive 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. To most effectively implement this type of analysis, monitoring data should share common comparable ancillary data across sites, and separate training and testing data subsets should be used to avoid overfitting.

- *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

8.1.3 Role of Coupled-Media Modeling and Analysis

Models or modeling 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 Section 8.1.2 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. The uncertainty from each model in the chain will be propagated into the next, meaning that relative uncertainty will increase with more models in the chain. Additionally, when models are chained in this way, the time scales of responses in each medium need to be considered, since a mercury response propagating through many media and through a food web can be delayed by years or longer.

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. Moreover, the responses of the terrestrial and ocean systems feed back 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 modeling 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 can not yet observe those benefits, coupled-media models can be used to estimate them.

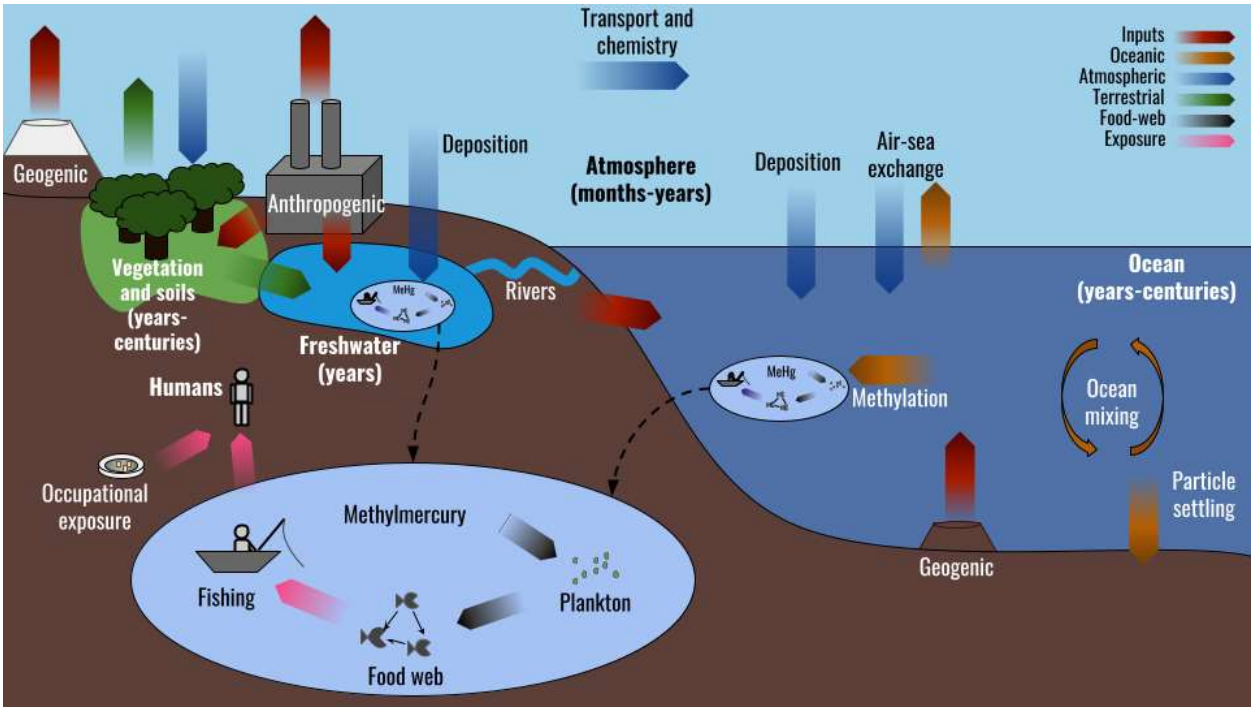


Figure 8.1: Flows of mercury between modeled media. Arrow colors indicate which models simulate those flows (red - model inputs, blue - atmospheric models, green - terrestrial models, orange - ocean models, black - food web models, magenta - exposure models), while values in parentheses indicate relevant time scales in each medium.

8.2 Model evaluation and comparison to observations

8.2.1 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 policy-relevant questions have clear answers and which require further monitoring/analysis to answer with a given degree of confidence. Combining multiple modeling 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 behavior of modeled 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.

In addition to model uncertainty, the comparison of modeled and observed quantities requires consideration of the uncertainty introduced to the comparison itself 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.

Recommendation

- For model evaluation and analyses, model output uncertainties should be quantified.
- Uncertainties introduced by model-observation comparisons should be included in those comparisons

8.2.2 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. Consistency between the model and measured values of the quantities of interest means that when accounting for the uncertainties in the model, the measurement, and the manipulation of each for the purpose of comparison, the model and measured values are not statistically distinguishable. In order to draw conclusions from the applied model, the uncertainty in its results must be smaller than the magnitude of the result itself.

Recommended 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 modeled quantity is consistent with available comparable monitoring results
3. Whether the uncertainty in the modeled 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.

8.2.3 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 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 are the accumulation of the types of uncertainty discussed in Section 8.2.1, and the model outputs are not necessarily the same quantity that monitoring efforts are measuring, but the two often overlap closely.

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 and 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. This type of top-down approach can be useful for attribution on its own, especially when used in Bayesian networks that avoid many of the challenges in multiple-regression based analysis (collinearity between independent variables, incomplete independent variables with respect to underlying mechanism, over fitting, multiple strong predictors, etc.). Top-down approaches 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, whereby bottom-up analysis (representing process-level knowledge) provides a prior estimate independent from the monitoring data itself, and top-down analysis provides the likelihood of those same estimates given the monitoring observations (representing the weight of evidence). By incorporating the quantified uncertainties of each model 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.

8.2.4 Standardized visualization and analyses

A standardized method of visualization and summary analysis would facilitate communication of combined monitoring and modeling findings. The most common form of visual comparison for spatially resolved models with collections of observational data is a colored map of modeled values with the corresponding observations of the same quantity overlaid as colored dots using

the same color scale. A standardized choice of color 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. Color maps should be chosen with consideration of viewers with color 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.

To facilitate quantitative model-monitoring comparisons, model output that is comparable to monitoring findings can be tabulated along with relevant uncertainty and other information for comparison. For spatially resolved models, this information can be extracted at each monitoring location. Relevant information for tabulation includes model uncertainty and/or confidence intervals for the output, spatial and temporal averaging info, and modeled trends and their confidence intervals.

8.2.5 Information from modeling

Table 8.3 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 4, would greatly facilitate both single-medium and multi-media analyses. A common, self-describing 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.

Table 8.3. Information from modeling 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	Seawater mercury	Input sources; circulation	Latitudes, longitudes,	Gross evasion fluxes to air or

	and temporal trends	concentrations, temporal trends	source; processes represented;	depths	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
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 models	Bulk mercury levels across media and their temporal trends	Media-averaged mercury levels, temporal trends	Input sources; rates/timescales represented	Represented spatial extent	Time-evolution of mercury levels across media and fluxes between media

8.3 Recommended Analysis Methods for Addressing Questions in Table 4.1

8.3.1 Characterize 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, and aggregating results by monitoring site type (e.g. urban/point-source-influenced/background) can show high-level source influences. 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 elevate 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, smaller-scale modeling could be significant.

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 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).

Recommendations

- Use several atmospheric chemical transport models 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 in areas with sparse monitoring data
- Use spatially resolved models to interpolate levels and trends of mercury while accounting for the drivers of spatial and temporal differences
- Any estimate should provide a detailed discussion / presentation of the associated uncertainty and factors that have determined this uncertainty

8.3.2 Identify 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.

Statistical analyses can be performed on time series data from air monitoring sites (Tier 1, 2, and 3) to identify observed trends which take into account sources of temporal variability. The Mann-Kendall test can identify the significance of an upward or downward trend across a time period and its standard error. The magnitude of the trend and its confidence interval can be obtained using this information and Sen's slope, which is robust to outliers and data which deviate from linear time behavior.

For biota monitoring (Tier 1, 2, and 3), if individual sample data are available, generalized linear modeling 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 modeling can be performed to quantify expected trends or relative trends in locations, forms, 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. Site-specific modeling can extend the observed trends in monitoring media to other media and to exposure and health impacts through cause-effect relationships.

Modeled 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.

Recommendations

- Identification of temporal trends should account for variability and uncertainty to obtain:
  - Trend magnitude
  - Confidence interval of trend
  - Measure of statistical significance of trend

8.3.3 Quantify 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 modeling can best represent the physical and chemical processes that determine the transport and fate of mercury. Better representation of these processes will increase the applicable scope of the modeling and contribute to an iterative process where future modeling will better answer the effectiveness evaluation questions of interest.

Recommendations

- Use top-down analysis to identify key environmental drivers
- Perform large-scale measurement/model intercomparisons to identify key processes
- Iterative approach: use available information to improve application of models

8.3.4 Inform 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 modeled 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 than either method individually.

Biota monitoring and ancillary data (Tier 2 and 3) can be used in top-down modeling 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 modeling 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 modeling, 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 behavior 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 behavioral changes influenced by the convention or changing mercury concentrations in the diet

When adding bottom-up modeling to the above-described approach, improved explanatory power that includes even more factors influenced by the convention can be obtained. 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.

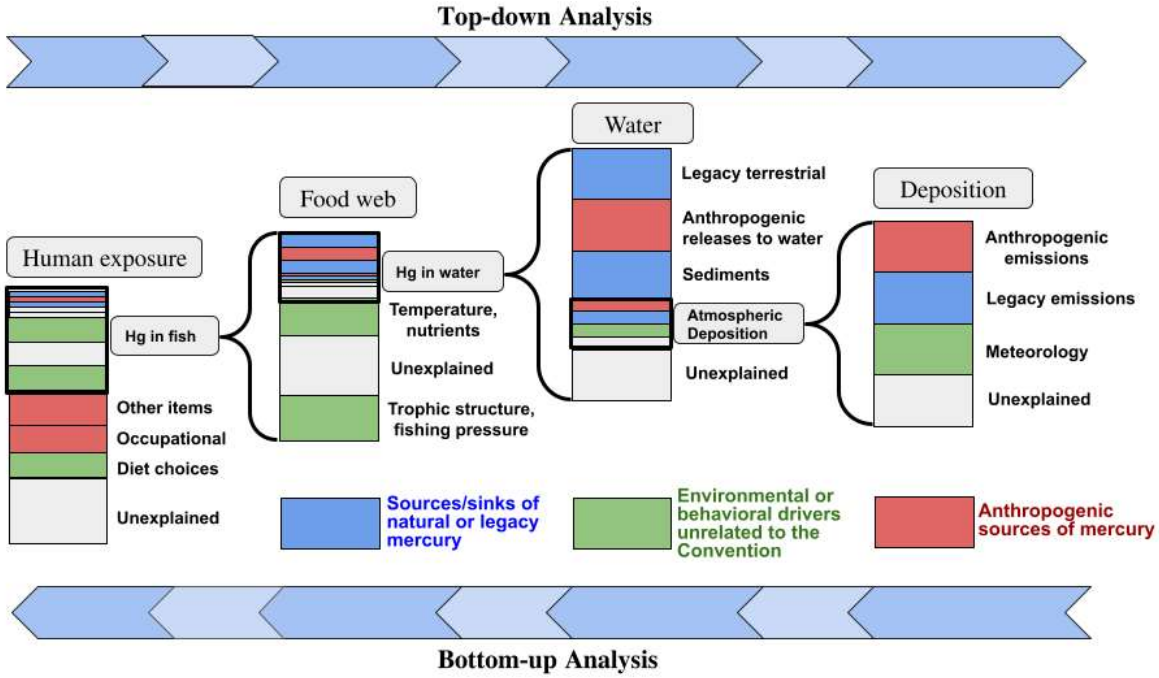


Figure 8.2. Illustration of attribution across media for hypothetical contributions of selected drivers at a hypothetical location. The colored bars represent the fractional contributions of different drivers to observed mercury trends/variability in each medium. The drivers of variability/change in a given medium can in turn be attributable to drivers in other media.

Recommendations

- Perform bottom-up analyses with atmospheric models for source attribution
- Perform top-down analyses with GLM/GAM for air and biota attribution where sufficient ancillary data is available
- Perform top-down analysis of changes in exposure pathways to attribute changes in human biomarkers to measures influenced by the convention
- At intensive monitoring sites, perform combined top-down and bottom-up attribution analyses for air, biota and human biomarkers
- To quantify legacy impacts, use coupled-media approaches where possible

8.3.5 Estimate 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.

Through the combination of this type of analysis and the source attribution discussed in the previous section, impacts and avoided impacts can be attributed to Convention-influenced actions as well as other contributing sources, in areas and subpopulations where confounding sources of variability have been accounted for. This type of analysis can require the combination of a range of modeling analyses, to connect emissions changes via air and environmental media to biota and humans and then to exposure and health impacts, and would require rigorous consideration of uncertainties.

Recommendations

- Where possible, estimate exposure based on specific sources and use exposure attribution information to estimate marginal health impacts/costs of individual drivers
- Estimate trends in risk associated with trends in exposure and/or biomarker benchmark values

Annex

Tiered approach to monitoring mercury and mercury compounds in the context of the effectiveness evaluation of the Minamata Convention on Mercury

1. INTRODUCTION

A tiered approach for programmes to monitor mercury and mercury compounds is proposed to support Parties who may wish to develop new monitoring programs, and/or improve existing ones, with a view to contributing to the effectiveness evaluation.

Tier 1 is intended to provide guidance to Parties that are seeking to create a monitoring program, or expand a minimal program, but that may not have sufficient resources to implement the actions in Tier 2. Following guidance by the COP, the methods in Tier 1 are cost effective, practical, feasible, and sustainable. The Tier 1 methods are intended to provide information that are useful to Parties in identifying and characterizing gaps and needs of national, regional, or local interest and to provide information that is useful to the collective effectiveness evaluation effort. While the implementation of Tier 1 actions may not fully address the four policy questions (see Chapter 3), 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 all policy questions identified by the COP and create a basis particularly for source attribution at the local, national, and global scales. The methods and approaches may be more expensive or complex than those under Tier 1. It would not be necessary that all Parties implement the Tier 2 actions for the effectiveness evaluation to be performed at a global level, but the basis for the effectiveness evaluation will be more robust with greater implementation.

Tier 3 identifies research methods and approaches that may play a vital role in supporting the Tier 2 programs and the effectiveness evaluation, primarily by improving our understanding of key processes that link sources of mercury to environmental concentrations and exposures, and ultimately impacts on human health and the environment. Not all Parties would be expected to employ Tier 3 methods or approaches, but where Tier 3 efforts and results are available, the information may be taken into consideration in the effectiveness evaluation.

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

2. AIR MONITORING

Hg Measurement	Ancillary Measurements	Location/Spatial Distribution	Frequency	Contribution to information categories[9]	Modeling/Analysis[10]
TIER 1					
Total Gaseous Mercury or Gaseous Elemental Mercury (a range of methods may be used depending on objectives, resources, and other constraints): - continuous mercury analyzers (recommended where possible) - manual trap 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.	Precipitation, meteorological data (where available, may be from nearby sites).	Sites should be selected in a mix of locations that include a)remote, background, b) regionally representative, and c) source impacted locations. 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.	(1) Baselines and Spatial Patterns; (2) Temporal Trends; (5) Estimating Exposure and Adverse Effects.	(1a) TGM/GEM concentrations from comparable measurements (present aggregate results per site type); supplemented by modelled TGM/GEM concentrations; (1b) Wet deposition concentrations from comparable measurements (present aggregate results per site type); supplemented by modelled wet deposition for areas without measurements; (2) Mann-Kendall, Sen's slope; (4a) "Bottom-up" CTMs for TGM/GEM. (4b) "Bottom-up" CTMs for wet deposition. Measurement / model intercomparisons for ambient mercury

[9] See chapter 4.

[10] See chapter 8.

					concentrations and wet deposition.
TIER 2					
Speciated Reactive Mercury: - high resolution measurements of PBM2.5, GOM using existing network SOPs; - cation exchange membranes. Dry deposition of mercury: - Total Hg and MeHg litterfall and throughfall (select forest ecosystems).	Emission inventories, atmospheric deposition of sulfate, land cover, land use, leaf area index, meteorology, air quality tracers (including SO_2, CO_2, CO, PM2.5),	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 and Spatial Patterns; (2) Temporal Trends; (3) Key Environmental Processes; (4) Source Attribution; (5) Estimating Exposure and Adverse Effects.	(3) Estimate air-ocean and air-terrestrial total Hg flux; (4c) "Top-down" modelling of SSR using ancillary measurements (aggregate results for different site-locations). Analysis using mechanistic and statistical models at the national, regional, and global scales.
TIER 3					
Mercury Isotopes: - e.g. multi-collector inductively coupled plasma; - mass spectrometry (MC-ICP-MS) Additional speciation methods Applications of Tier 1 and Tier 2 methods in intensive research contexts to support process understanding	Same as above.	Expected to be opportunistic siting, collocated at long-term monitoring and research sites. Aircraft campaigns, ocean surveys, flux towers, etc.	Data is collected through research programs and will be provided when available for the purposes of supporting the effectiveness evaluation.	(2) Temporal Trends; (3) Key Environmental Processes; (4) Source Attribution; (5) Estimating Exposure and Adverse Effects.	(4) Calibration of "Bottom-up" CTMs input parameters with "top-down" approaches; Including isotopic fingerprints, where applicable. Measurement-model fusion to estimate total mercury deposition; dry deposition modeling (local)

2338
2339

2340 **3. BIOTA MONITORING**

Hg Measurement	Ancillary Measurements	Location/Spatial Distribution	Frequency	Contribution to information categories[11]	Modeling/Analysis[12]
TIER 1					
Tissue from bird and fish can be used to monitor total mercury levels in biota. Species to be 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 4 species are used in a number of existing programs and are a reasonable starting point.	Spatial Coordinates (Latitude/Longitude), Species Name, Body Length & Mass, Age, Sex and Maturity Stage; catchment description (size of lake, elevation, landcover and use)	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. Where little or no prior information exists, the possibility of using ecosystem sensitivity modelling for the chosen sites may be explored.	Annual measurements, with a consistent sampling season over time for each core fixed site.	(2) Temporal Trends; (5) Estimating Exposure and Adverse Effects.	(2) Time series analysis of individual observations within a generalized linear model (GLM) or framework (i.e. mixed effects model) that can account for variation associated with species and site Note that individual sample data is most useful for analysis rather than aggregated values.
TIER 2					
In Tier 2, a consistent taxon would be sampled in different sites over time. While it is important to sample as consistent a taxon as possible across locations, if that is not possible, sampling several taxa in the	Additional ancillary measurements may include: In biota: carbon (δ13C) & nitrogen (δ15N) stable isotopes; In water: DOM/DOC/TOC, TSS, salinity, DO, (pH). N	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 type of habitat type and then	Yearly monitoring rotating across sites added at Tier 2 (in such a manner that each particular site would only be monitored	(1) Baselines and Spatial Patterns; (2) Temporal Trends; (3) Key Environmental Processes;	(4a) "Top-down" approach: Use generalized additive model (GAM) or Bayesian models to estimate the relative contribution of different influxes, sources and large-scale drivers with the help of the ancillary

[11] See chapter 4.

[12] See chapter 8.

multiple sites would help in accounting for species effects statistically. However, it is noted that monitoring novel species that have not been previously monitored elsewhere would be less informative for the effectiveness evaluation.	and P, phytopigments (e.g. chlorophyl-a); In sediment: THg In Air: GEM, wet deposition, and meteorological data. Description of local hydrologic catchment.	either rotated or randomly sampled within each habitat category. 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.	every few years).	(4) Source Attribution; (5) Estimating Exposure and Adverse Effects.	measurements (assumed to be affected by them). Local and regional influxes / sources / drivers can be presented in a combined form to form a global picture; (4b) "Bottom-up" approach: Use mechanistic models to distinguish between sources of deposited mercury.
TIER 3					
Sampling as above, but consideration may be given to species in 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.	Mercury (δ202Hg) stable isotopes in biota and suspected source-matrixes of interest; other chemical tracers related to known drivers (i.e. changes in CO2 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 food web structure, as well as associated land cover data.	Intensively monitor selected areas (e.g. catchments), with a primary site (supersite) for collocated measurements and secondary (or satellite) sites to capture variability across the catchment. Catchments 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 and Spatial Patterns; (2) Temporal Trends; (3) Key Environmental Processes; (4) Source Attribution; (5) Estimating Exposure and Adverse Effects.	(4) Calibrate "Bottom-up" CTMs input parameters with "top-down" approaches, including isotopic fingerprints, where applicable.

3. HUMAN BIOMONITORING

Hg Measurement	Ancillary Measurements	Location/Spatial Distribution	Frequency	Contribution to information categories[13]
TIER 1				
Blood, urine, or hair THg depending on sampled population.	WHO Survey or HBM4EU Instruments. Relevant biota measurements, where possible.	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.	(1) Baselines and Spatial Patterns; (2) Temporal Trends; (5) Estimating Exposure and Adverse Effects.
TIER 2				
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 biota measurements.	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 and Spatial Patterns; (2) Temporal Trends; (3) Key Environmental Processes; (4) Source Attribution; (5) Estimating Exposure and Adverse Effects.
TIER 3				
same as above for Tier 2	WHO Survey or HBM4EU Instruments or National/Regional population survey instruments. Relevant biota measurements.	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 and Spatial Patterns; (2) Temporal Trends; (3) Key Environmental Processes; (4) Source Attribution; (5) Estimating Exposure and Adverse Effects.

[13] See chapter 4.

References (to be finalized)

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

Abeysinghe, K.S., Qiu, G., Goodale, E., Anderson, C.W., Bishop, K., Evers, D.C., Goodale, M.W., Hintelmann, H., Liu, S., Mammides, C. and Quan, R.C., 2017. Mercury flow through an Asian rice-based food web. Environmental Pollution, 229, pp.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.

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

AMAP. 2018. AMAP Assessment 2018: Biological effects of contaminants on Arctic Wildlife and Fish. Arctic Monitoring and Assessment Programme (AMAP), Tromso, Norway. 84pp.

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

Angot, H., Barret, M., Magand, O., Ramonet, M., Dommergue, A. A 2-year record of atmospheric mercury species at a background Southern Hemisphere station on Amsterdam Island. Atmospheric Chemistry and Physics, 14, 11461-11473, 2014.

Arnold, R.D. and Wade, J.P., 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).

Basu, Niladri, Milena Horvat, David C Evers, Irina Zastenskaya, Pál Weihe, and Joanna Tempowski. “A State-of-the-Science Review of Mercury Biomarkers in Human Populations Worldwide between 2000 and 2018.” Environmental Health Perspectives 126, 1–14, 2018.

Basu, Niladri, Milena Horvat, David C Evers, Irina Zastenskaya, Pál Weihe, and Joanna Tempowski. 2018. “A State-of-the-Science Review of Mercury Biomarkers in Human Populations Worldwide between 2000 and 2018.” Environmental Health Perspectives 126 (10): 1–14. https://doi.org/10.1289/EHP3904.

Bignert A, Riget F, Braune B, Outridge P, Wilson S. 2004. Recent temporal trend monitoring of mercury in Arctic biota - how powerful are the existing data sets? Journal of Environmental Monitoring 6:351-355.

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

Braaten, H.F.V., Gundersen, C.B., Sample, J.E., Selvik, J.R. and de Wit, H. 2018. Atmospheric deposition and lateral transport of mercury in Norwegian drainage basins: A mercury budget for Norway. NIVA-rapport.

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

Braune, B.M., Gaston, A.J. and Mallory, M.L., 2016. Temporal trends of mercury in eggs of five sympatrically breeding seabird species in the Canadian Arctic. Environmental Pollution, 214, pp.124-131.

Buck, DG, DC Evers, E Adams, J DiGangi, B Beeler, J Samanek, J Petrlik, MA Turnquist, O Speranskaya, K Reagan. 2019. A global-scale assessment of fish mercury concentrations and the identification of biological hotspots. Science of the Total Environment 687:956-966.

Carvan, M.J., T.A. Kalluvila, R.H. Klingler, J.K. Larson, M. Pickens, F.X. Mora-Zamorano, V.P. Connaughton, I. Sadler-Riggleman, et al. 2017. Mercury-induced epigenetic transgenerational inheritance of abnormal neurobehavior is correlated with sperm epimutations in zebrafish. PLoS ONE 12:e0176155.

Chaves-Ulloa, R., Taylor, B.W., Broadley, H.J., Cottingham, K.L., Baer, N.A., Weathers, K.C., Ewing, H.A. and Chen, C.Y. 2016. Dissolved organic carbon modulates mercury concentrations in insect subsidies from streams to terrestrial consumers. Ecological Applications 26:1771-1784.

Chen, C.Y., Pickhardt, P.C., Xu, M.Q. and Folt, C.L., 2008. Mercury and arsenic bioaccumulation and eutrophication in Baiyangdian Lake, China. Water, Air, and Soil Pollution, 190:115-127.

Chételat, J., Richardson, M.C., MacMillan, G.A., Amyot, M. and Poulain, A.J. 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 (Council for International Organizations of Medical Sciences). 2016. “International Ethical Guidelines for Health-Related Research Involving Humans, 4th Edition.”

Cransveld, A., Amouroux, D., Tessier, E., Koutrakis, E., Ozturk, A.A., Bettoso, N., Mieiro, C.L., Bérail, S., Barre, J.P., Sturaro, N. and 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., and D.C. Evers. 2020. The impact of mercury on North American songbirds: effects, trends, and predictive factors. Ecotoxicology

D’Amore, F., Bencardino, M., Cinnirella, S., Sprovieri, F., and Pirrone, N.: Data quality through a web-based QA/QC system: implementation for atmospheric mercury data from the Global Mercury Observation System, Environmental Science: Processes 17, 1482-1493, 2015.

Day, Brent, Renée Langlois, Mark Tremblay, and Bartha Maria Knoppers. 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. C.N., Hedgecock. I.M., and Pirrone, A Modeling Comparison of Mercury Deposition from Current Anthropogenic Mercury Emission Inventories Environ. Sci. Technol. 50, 5154–5162, 2016.

Depew, D.C., Basu, N., Burgess, N.M., Campbell, L.M., Devlin, E.W., Drevnick, P.E., Hammerschmidt, C.R., Murphy, C.A., Sandheinrich, M.B. and Wiener, J.G. 2012a. Toxicity of dietary methylmercury to fish: derivation of ecologically meaningful threshold concentrations. Environmental Toxicology and Chemistry 31:1536-1547.

Depew, D.C., Basu, N., Burgess, N.M., Campbell, L.M., Evers, D.C., Grasman, K.A. and Scheuhammer, A.M., 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.

Dietz, R., Letcher, R.J., Desforges, J.P., Eulaers, I., Sonne, C., Wilson, S., Andersen-Ranberg, E., Basu, N., Barst, B.D., Bustnes, J.O., Bytingsvik, J., Ciesielski, T.M., Drevnick, P.E.,

Gabrielsen, G.W., Haarr, A., Hylland, K., Levin, M., Mckinney, M.A., Nørregaard, R.D., Pedersen, K.E., Provencher, J., Tartu, S., Aars, J., Ackerman, J.T., Rosing-Asvid, A., Barrett, R., Bignert, A., Born, E.W., Branigan, M., Braune, B., Bryan, C.E., Eagles-Smith, C.A., Evans, M., Evans, T.J., Fisk, A.T., Gamberg, M., Gustavson, K., Hartman, C.A., Helander, B., Herzog, M.P., Hoekstra, P·F., Houde, M., Hoydal, K., Jackson, A.K., Kucklick, J., Lie, E., Loseto, L., Mallory, M.L., Miljeteig, C., Mosbech, A., Muir, D.C.G., Nielsen, S.T., Peacock, E., Pedro, S., Peterson, S·H., Polder, A., Rigét, F·F., Roach, P., Saunes, H., Sinding, M.H·S., Skaare, J.U., Søndergaard, J., Stenson, G., Stern, G., Treu, G., Schuur, S.S., Víkingsson, G., 2019. Current state of knowledge on biological effects from contaminants on arctic wildlife and fish. Sci. Total Environ. 696: 133792.

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

DiMento, B.P., Mason, R.P., Brooks, S. and Moore, C. The impact of sea ice on the air-sea exchange of mercury in the Arctic Ocean. Deep Sea Research Part I: Oceanographic Research Papers 144, 28-38, 2019.

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

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

Driscoll, C.T., Mason, R.P, Chan, H.M., Jacob, D.J and Pirrone, N. Mercury as a global pollutant: sources, pathways and effects. Environ. Sci. Technol. 47, 4967-4983, 2013.

Driscoll, C.T., R.P. Mason, H.M. Chan, D.J. Jacob, and N. Pirrone. 2013. Mercury as a global pollutant: sources, pathways, and effects. Environmental Science and Technology 47:4967-4983.

Driscoll, C.T., Y.J. Han, C.Y. Chen, D.C. Evers, K.F. Lambert, T.M. Holsen, N.C. Kamman, and R. Munson. 2007. Mercury contamination in remote forest and aquatic ecosystems in the northeastern U.S.: Sources, transformations and management options. Bioscience 57:17-28.

Durnford, D., Dastoor, A., Ryzhkov, A., Poissant, L., Pilote, M., and Figueras-Nieto, D. How relevant is the deposition of mercury onto snowpacks? – Part 2: A modeling study, Atmos. Chem. Phys., 12, 9251–9274, 2012.

Eagles-Smith C.A., J.T. Ackerman, J.J. Willacker, M.T. Tate, M.A. Lutz, J.A. Fleck, A.R. Stewart, J.G. Wiener, D.C. Evers, J.M. Lepak, and J.A. Davis. 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, C. A.; Silbergeld, E. K.; Basu, N.; Bustamante, P.;Diaz-Barriga, F.; Hopkins, W. A.; Kidd, K. A.; Nyland, J. F. Modulators of mercury risk to wildlife and humans in the context of rapid globalchange. Ambio 47, 170−197, 2018.

Eagles-Smith, C.A., Wiener, J.G., Eckley, C.S., Willacker, J.J., Evers, D.C., Marvin-DiPasquale, M., Obrist, D., Fleck, J.A., Aiken, G.R., Lepak, J.M. and Jackson, A.K. 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, EK Silbergeld, N. Basu, P. Bustamante, F Diaz-Barriga, WA Hopkins, KA Kidd and JF Nyland. 2018. Modulators of mercury risk to wildlife and humans in the context of rapid global change. Ambio 47:170-197.

Eagles-Smith, Collin A., Ellen K. Silbergeld, Niladri Basu, Paco Bustamante, Fernando Diaz-Barriga, William A. Hopkins, Karen A. Kidd, and Jennifer F. Nyland. 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. https://doi.org/10.1007/s13280-017-1011-x.

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

Evers D.C., J.G. Wiener, N. Basu, R.A. Bodaly, H.A. Morrison, K.A. Williams. 2011a. Mercury in the Great Lakes region – Bioaccumulation, spatial and temporal patterns, ecological risks, and policy. Ecotoxicology 20:1487-1499.

Evers, D.C. 2018. The Effects of Methylmercury on Wildlife: A Comprehensive Review and Approach for Interpretation. Encyclopedia of the Anthropocene. Elsevier.

Evers, D.C. and E. Sunderland. 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.

Evers, D.C., Keane, S.E., Basu, N. and 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, D.C., S.E. Keane, N. Basu, and D. Buck. 2016. Evaluating the effectiveness of the Minamata Convention on Mercury: Principles and recommendations for next steps.

Evers, D.C., Sauer, A.K., Burns, D.A., Fisher, N.S., Bertok, D.C., Adams, E.M., Burton, M.E.H. and Driscoll, C.T. 2020. A synthesis of patterns of environmental mercury inputs, exposure and effects in New York State. Ecotoxicology 29:1565-1589.

Evers, D.C., Y.J. Han, C.T. Driscoll, N.C. Kamman, M.W. Goodale, K.F. Lambert, T.M. Holsen, C.Y. Chen, T.A. Clair, and T. Butler. 2007. Identification and Evaluation of Biological Hotspots of Mercury in the Northeastern U.S. and Eastern Canada. Bioscience 57:29-43.

Evers, David C., Susan Egan Keane, Niladri Basu, and David Buck. 2016. “Evaluating the Effectiveness of the Minamata Convention on Mercury: Principles and Recommendations for next Steps.” Science of the Total Environment 569–570: 888–903. https://doi.org/10.1016/j.scitotenv.2016.05.001.

FAO. 2020. The State of World Fisheries and Aquaculture 2020. Sustainability in Action. https://doi.org/10.4060/ca9229en.

Fielding, R., Kiszka, J.J., Macdonald, C., McCormack, M.A., Dutton, J., Ollivierre, A.D., Arnett, J.A., Elkins, M., Darby, N.A., Garcia, H.M. and 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: An International Journal 2021:1-25.

Gabriel, M.C., Howard, N. and Osborne, T.Z. 2014. Fish mercury and surface water sulfate relationships in the Everglades Protection Area. Environmental Management 53:583-593.

GEMS/Food Contaminants. n.d. “Global Environment Monitoring System—Food Contamination Monitoring and Assessment Program.” 2021. Accessed January 7, 2021. https://extranet.who.int/gemsfood/Default.aspx#.

Gerson, JR, N Szponar, AA Zambrano, B Bergquist, E Broadbent, CT Driscoll, GE Watsa, ME Watsa, DC Evers, LE Fernandez, H Hsu-Kim, G Inga, A Martinez, C Moore, W Pan, M Silman, EA Ury, C Vega, ES Bernhardt. 2021. Amazon forests capture high levels of atmospheric mercury pollution from artisanal gold mining. Nature Communications.

Goodrich, Jaclyn M J.M., Hwai-Nan H.-N. Chou, Stephen E Gruninger, Alfred Franzblau, and Niladri Basu. 2016. “Exposures of Dental Professionals to Elemental Mercury and Methylmercury.” Journal of Exposure Science & Environmental Epidemiology 26 (1): 78–85. https://doi.org/10.1038/jes.2015.52.

Gorski, P.R., Armstrong, D.E., Hurley, J.P. and Krabbenhoft, D.P. 2008. Influence of natural dissolved organic carbon on the bioavailability of mercury to a freshwater alga. Environmental Pollution 154:116-123.

Gustin MS, Dunham-Cheatham SM, Zhang L. Comparison of 4 Methods for Measurement of Reactive, Gaseous Oxidized, and Particulate Bound Mercury. Environmental Science & Technology 53 14489-14495, 2019.

Gustin, M., D.C. Evers, M. Bank, C.R. Hammerschmidt, A. Pierce, N. Basu, J. Blum, P. Bustamante, C. Chen, C.T. Driscoll, M. Horvat, D. Jaffe, J. Pacyna, N. Pirrone, and N. Selin. 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.S., H.M. Amos, J. Huang, M.B. Miller, and K. Heidecorn. 2015. Measuring and modeling mercury in the atmosphere: A critical review. Atmospheric Chemistry and Physics 15:5697–5713.

Ha, Eunhee, Niladri Basu, Stephan Bose-O’Reilly, José G. Dórea, Emeir McSorley, Mineshi Sakamoto, and Hing Man Chan. 2017. “Current Progress on Understanding the Impact of Mercury on Human Health.” Environmental Research 152: 419–33. https://doi.org/10.1016/j.envres.2016.06.042.

HBM4EU. 2017a. “Data Management Plan (DMP).” https://www.hbm4eu.eu/work-packages/deliverable-10-1-data-management-plan-july-2017/.

HBM4EU. 2017b. “Statistical Analysis Plan.” https://www.hbm4eu.eu/work-packages/deliverable-10-5-statistical-analysis-plan/.

HBM4EU. 2018a. “Legal and Ethics Policy Paper Update August 2018.” https://www.hbm4eu.eu/?mdocs-file=4263.

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

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

HBM4EU. 2018d. “SOP 3: Procedure for Obtaining Human Samples.” 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.” 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.” 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.” 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.” https://www.hbm4eu.eu/?mdocs-file=6084.

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

Horowitz, H. M., Jacob, D. J., Zhang, Y., Dibble, T. S., Slemr, F.,Amos, H. M., Schmidt, J. A., Corbitt, E. S., Marais, E. A., and Sunderland, E. M.: A new mechanism for atmospheric mercury redox chemistry: Implications for the global mercury budget, Atmos. Chem. Phys. 17, 6353-6371, 2017.

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

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

Huang JY, Choi HD, Landis MS, Holsen TM. An application of passive samplers to understand atmospheric mercury concentration and dry deposition spatial distributions. Journal of Environmental Monitoring; 14: 2976-2982, 2012.

Kocman, D, SJ Wilson, HM Amos, KH Telmer, F Steenhuisen, EM Sunderland, RP Mason, P Outridge, M Horvat. 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

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.

Louro, Henriqueta, Milla Heinälä, Jos Bessems, Jurgen Buekers, Theo Vermeire, Marjolijn Woutersen, Jacqueline van Engelen, et al. 2019. “Human Biomonitoring in Health Risk Assessment in Europe: Current Practices and Recommendations for the Future.” International Journal of Hygiene and Environmental Health 222 (5): 727–37. https://doi.org/10.1016/j.ijheh.2019.05.009.

Lukina, Anna O., Mandy Fisher, Cheryl Khoury, John Than, Mireille Guay, Jean François Paradis, Tye E. Arbuckle, and Melissa Legrand. 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. https://doi.org/10.1016/j.chemosphere.2020.128402.

Lusk, J.D., Rich, E. and Bristol, R.S., 2005. Methylmercury and other environmental contaminants in water and fish collected from four recreational fishing lakes on the Navajo Nation, 2004. Completion report for intergovernmental agreement. US Fish and Wildlife Service, New Mexico Ecological Services Field Office, Albuquerque, NM.

Macagnano, A., Papa, P., Avossa, J., Perri, V., Marelli, M., Sprovieri, F., Zampetti, E., Bearzotti, A. and Pirrone, N.: Passive Sampling of Gaseous Elemental Mercury Based on a Composite

Mahaffey, Kathryn R, Elsie M Sunderland, Hing Man Chan, Anna L Choi, Philippe Grandjean, Koenraad Marien, Emily Oken, et al. 2011. “Balancing the Benefits of N-3 Polyunsaturated Fatty Acids and the Risks of Methylmercury Exposure from Fish Consumption.” NUTRITION REVIEWS 69 (9): 493–508. https://doi.org/10.1111/j.1753-4887.2011.00415.x.

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

Mao, H., Talbot, R. Speciated mercury at marine, coastal, and inland sites in New England- Part 1: Temporal variability. Atmospheric Chemistry and Physics, 12, 5099-5112, 2012.

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

Mason, R.P., Hammerschmidt, C.R., Lamborg, C.H., Bowman, K.L., Swarr, G.J., Shelley, R.U. The air-sea exchange of mercury in the low latitude Pacific and Atlantic Oceans. Deep- Sea Research 122: 17-28, 2017.

McLagan, D. S., C. P. J. Mitchell, H. Huang, B. Abdul Hussain, Y. D. Lei, F. Wania. 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, 2016.

McLagan, D.S., C.P.J. Mitchell, A. Steffen, H. Hung, C. Shin, G.W. Stupple, M.L. Olson, W.T. Luke, P. Kelley, D. Howard, G.C. Edwards, P.F. Nelson, H. Xiao, G.-R. Sheu, A. Dreyer, H. Huang, B. Abdul Hussain, Y.D. Lei, I. Tavchunsky, F. Wania. Global evaluation and calibration of a passive air sampler for gaseous mercury. Atmos. Chem. Phys. 18, 5905-5919, 2018.

McLagan, D.S., C.P.J. Mitchell, H. Huang, Y.D. Lei, A.S. Cole, A. Steffen, H. Hung, F. Wania. A high precision passive air sampler for gaseous mercury. Environ. Sci. Technol. Lett. 3, 24-29, 2016.

Miller MB, Dunham-Cheatham SM, Gustin MS, Edwards GC. Evaluation of cation exchange membrane performance under exposure to high Hg0 and HgBr2 concentrations. Atmospheric Measurement Techniques 12, 1207-1217, 2019.

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, G. W., Mitchell, C. P. J., Martin, A. R., Steffen, A., Babi, D., Prestbo, E. M., Sprovieri, F., and Wania, F.: A field intercomparison of three passive air samplers for gaseous mercury in ambient air, Atmos. Meas. Tech. Discuss. In review, 2020.

NADP Site Selection and Installation Manual, http://nadp.slh.wisc.edu/lib/manuals (last visited Feb 2021)

Nerentorp Mastromonaco, M., Gårdfeldt, K., Jourdain, B., Abrahamsson, K., Granfors, A., Ahnoff, M., Dommergue, A., Méjean, G., Jacobi, H.-W. Antarctic winter mercury and ozone depletion events over sea ice. Atmospheric Environment 129, 125-132, 2016.

Obrist, D.; Kirk, J.; Zhang, L.; Sunderland, E.; Jiskra, M.; Selin, N.E. 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, 2018.

Palacios-Torres, Y., Jesus, D. and Olivero-Verbel, J., 2020. Trace elements in sediments and fish from Atrato River: an ecosystem with legal rights impacted by gold mining at the Colombian Pacific. Environmental Pollution, 256, p.113290.

Pirrone, N., Ferrara, R., Hedgecock, I. M., Kallos, G., Mamane, Y., Munthe, J., et al. Dynamic processes of mercury over the Mediterranean region: results from the Mediterranean Atmospheric Mercury Cycle System (MAMCS) Project. Atmospheric Environment, 37 , S21–S39, 2003.

Prestbo, E.M., and D.A. Gay. 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, 2009.

Rauert, C.; Harner, T.; Schuster, J. K.; Quinto, K.; Fillmann, G.; Castillo, L. E.; Fentanes, O.; Ibarra, M. V.; Miglioranza, K.; Rivadeneira, I. M.; Pozo, K.; Puerta, A. P.; Zuluaga, B. H. A. Towards a regional passive air sampling network and strategy for new POPs in the GRULAC region: Perspectives from the GAPS Network and first results for organophosphorus flame retardants. Sci. Total Environ.573, 1294−1302, 2016.

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

Rothenberg, Sarah E., Lisamarie Windham-Myers, and Joel E. Creswell. 2014. “Rice Methylmercury Exposure and Mitigation: A Comprehensive Review.” Environmental Research 133: 407–23. https://doi.org/10.1016/j.envres.2014.03.001.

Rudd, J.W., Bodaly, R.A., Fisher, N.S., Kelly, C.A., Kopec, D. and 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.

Schartup, A.T., A. Qureshi, C. Dassuncao, C.P. Thackray, G. Harding, E.M. Sunderland. 2018. A model for uptake and trophic transfer of methylmercury by marine plankton. Environmental Science & Technology. 52(2):654-662

Scheuhammer, A., Braune, B., Chan, H.M., Frouin, H., Krey, A., Letcher, R., Loseto, L., Noël, M., Ostertag, S., Ross, P. and 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.

Schmeltz, D., Evers, D.C., Driscoll, C.T. et al. MercNet: a national monitoring network to assess responses to changing mercury emissions in the United States. Ecotoxicology 20, 1713–1725, 2011.

Schroeder, W. H., Munthe, J. ATMOSPHERIC MERCURY-AN OVERVIEW. Atmospheric Environment, 32, 809 – 822, 1998.

Sheehan, Mary C, Thomas A Burke, Ana Navas-Acien, Patrick N Breysse, John McGready, and Mary A Fox. 2014. “Global Methylmercury Exposure from Seafood Consumption and Risk of Developmental Neurotoxicity: A Systematic Review.” Bulletin of the World Health Organization 92 (4): 254-269F. https://doi.org/10.2471/blt.12.116152.

Sherman, Laura S, Joel D Blum, Alfred Franzblau, and Niladri Basu. 2013. “New Insight into Biomarkers of Human Mercury Exposure Using Naturally Occurring Mercury Stable Isotopes.” Environmental Science & Technology 47 (7): 3403–9. https://doi.org/10.1021/es305250z.

Sherman, Laura S, Joel D Blum, Niladri Basu, Mozhgon Rajaee, David C Evers, David G Buck, Jindrich Petrlik, and Joseph Digangi. 2015. “Assessment of Mercury Exposure among

Small-Scale Gold Miners Using Mercury Stable Isotopes." Environmental Research 137: 226–34. https://doi.org/10.1016/j.envres.2014.12.021.

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

Skov, H., Christensen, J. H., Goodsite, M. E., Heidam, N. Z., Jensen, B., Wahlin, P., and Geernaert, G.: 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, 2004.

Skov, H., Sørensen, B. T., Landis, M. S., Johnson, M. S., Sacco, P., Goodsite, M. E., Lohse, C., and Christiansen, K. S.: Performance of a new diffusive sampler for Hg0 determination in the troposphere, Environ. Chem., 4, 75–80, 2007.

Slemr, F., Martin, L., Labuschagne, C., Mkololo, T., Angot, H., Magand, O., D Dommergue, A., Garat, P., Ramonet, M., and Bieser, J.: 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, 2020.

Sprovieri, F., Pirrone, N, Bencardino, N., D´Amore, F., Angot, H., Barbante, C., Brunke, E.-G., Arcega-Cabrera, F., Cairns, W., Comero, S., del Carmen Diéguez, M., Dommergue, A., Ebinghaus, R., Feng, X.B., Fu, X., Garcia, P.E., Gawlik, P.M., Hageström, U., Hansson, K., Horvat, M., Kotnik, J., Labuschagne, C., Magand, O., Martin, L., Mashyanov, N., Mkololo, T., Munthe, J., Obolkin, V., Ramirez Islas, M., Sena, F.,Somerset, V., Spandow, P., Vardè, M., Walters, C., Wängberg, I., Weigelt, A., Yang, X., and Zhang, H. Five-year records of mercury deposition flux at GMOS sites in the Northern and Southern hemispheres, Atmos. Chem. Phys., 17, 2689-2708, 2017.

Sprovieri, F., Pirrone, N, Bencardino, N., D´Amore, F., Carbone, F., Cinnirella, S., Mannarino, V., Landis, M., Ebinghaus, R., Weigelt, A., Brunke, E.-G., Labuschagne, C., Martin, L., Munthe, J., Wängberg, I., Artaxo, P., Morais, F., de Melo Jorge Barbosa, H., Brito, J., Cairns, W., et al. Atmospheric mercury concentrations observed at ground-based monitoring sites globally distributed in the framework of the GMOS network, Atmos. Chem. Phys., 16, 11915-11935, 2016.

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

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

Sunderland EM, Krabbenhoft DP, Moreau JW, Strode SA, Landing WM. Mercury sources, distribution, and bioavailability in the North Pacific Ocean: Insights from data and models. Global Biogeochemical Cycles. 2009 Jun;23(2).

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.

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

UNEP and World Health Organization (WHO). 2008. “Guidance for Identifying Populations At Risk From Mercury Exposure.” UNEP and WHO: Geneva, Switzerland.

UNEP. 2016. UNEP Global Review of Mercury Monitoring Networks. United Nations Environment, Geneva, Switzerland. (http://www.mercuryconvention.org/Portals/11/documents/2016%20call%20for%20submissions/UNEP%20-%20Global%20Review%20of%20Mercury%20Monitoring%20Networks_Final.pdf).

UNEP. 2019. “Global Mercury Assessment 2018.” Geneva, Switzerland.

US EPA. 2016. “Guidance for Conducting Fish Consumption Surveys. EPA 823-B-16-002.” Washington, DC.

Vandenbroucke, Jan P, Erik Von Elm, Douglas G Altman, Peter C Gøtzsche, Cynthia D Mulrow, Stuart J Pocock, Charles Poole, James J Schlesselman, and Matthias Egger. 2007. “Strengthening the Reporting of Observational Studies in Epidemiology (STROBE): Explanation and Elaboration.” PLoS Medicine | Www 4. https://doi.org/10.1371/journal.pmed.

Veiga, Marcello M., and Randy Baker. 2003. “Protocols for Environmental and Health Assessment of Mercury Released by Artisanal and Small-Scale Gold Miners.” UNIDO.

Višnjevec, Ana Miklavčič, David Kocman, and Milena Horvat. 2014. “Human Mercury Exposure and Effects in Europe.” Environmental Toxicology and Chemistry 33 (6): 1259–70. https://doi.org/10.1002/etc.2482.

Wängberg, I., Hageström, U., Sommar, J. and Ferm, M.: Development and Testing of a Passive Sampler for Measurement of Gaseous Mercury, Atmos. Chem. Phys. Discuss., 528:1-9, 2016.

Weiss-Penzias, P.S., Gay, D.A., Brigham, M.E., Parsons, M.T., Gustin, M.S., ter Schure, A.: Trends in mercury wet deposition and mercury air concentrations across the US and Canada. Sci. Total Environ. 568, 546–556, 2016.

Whitney, M.C. and D.A. Cristol. 2017. Impacts of Sublethal Mercury Exposure on Birds: A Detailed Review. Reviews of Environmental Contamination and Toxicology, DOI 10.1007/398_2017_4.

WHO Regional Office for Europe. 2015. “Human Biomonitoring: Facts and Figures.”

Wiener, JG, MB Sandheinrich, SP Bhavsar, JR Bohr, DC Evers, BA Monson and CS Schrank. 2012. 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. The FAIR Guiding Principles for scientific data management and stewardship. Sci Data 3, 160018 (2016). https://doi.org/10.1038/sdata.2016.18

Wilkinson, Mark D. 2016. “Comment: The FAIR Guiding Principles for Scientific Data Management and Stewardship.” Nature Publishing Group

Wilkinson, Mark D. 2016. “Comment: The FAIR Guiding Principles for Scientific Data Management and Stewardship.” Nature Publishing Group. https://doi.org/10.1038/sdata.2016.18.

Wintle, B.A., Bekessy, S.A., Keith, D.A., Van Wilgen, B.W., Cabeza, M., Schröder, B., Carvalho, S.B., Falcucci, A., Maiorano, L., Regan, T.J. and Rondinini, C., 2011. Ecological–economic optimization of biodiversity conservation under climate change. Nature Climate Change, 1(7), pp.355-359.

WMO/GAW Report 160. Manual for the GAW Precipitation Chemistry Programme, Nov 2004.

WMO/GAW Report 181. Operations Handbook - Ozone Observations with a Dobson Spectrophotometer: revised Nov 2008.

WMO/GAW Report 192. Guidelines for the Measurement of Atmospheric Carbon Monoxide, Nov 2010.

WMO/GAW Report 201. Quality Assurance and Quality Control for Ozonesonde Measurements in GAW, Oct 2014.

WMO/GAW Report 201. Standard Operating Procedures (SOPs) for Air Sampling in Stainless Steel Canisters for Non-Methane Hydrocarbons Analysis, Sep 2012.

World Health Organization (WHO). 2018a. “Assessment of Prenatal Exposure to Mercury: Human Biomonitoring Survey The First Survey Protocol.” World Health Organization, 1–59. http://www.euro.who.int/__data/assets/pdf_file/0010/386893/survey-mercury-eng.pdf?ua=1&ua=1.

World Health Organization (WHO). 2018b. “Assessment of Prenatal Exposure to Mercury: Standard Operating Procedures Ii.” http://www.euro.who.int/pubrequest.

World Health Organization (WHO). 2021a. “Human biomonitoring in artisanal and small-scale gold mining: ethical and scientific principles.” https://www.who.int/publications/i/item/9789240020702

World Health Organization (WHO). 2021b. “A Step-by-Step Guide of the WHO Framework for Developing a Public Health Strategy for Artisanal and Small-Scale Gold Mining in the Context of the Minamata Convention on Mercury.”

Wright, P.L., Zhang, L., Marsik, F. Overview of mercury dry deposition, litterfall, and throughfall studies. Atmos. Chem. Phys 16, 13399 – 13416, 2016.

Wyn, B, KA Kidd, NM Burgess, RA Curry. 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, J.E., 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.

Zhang L, Zhou P, Cao S, Zhao Y. Atmospheric mercury deposition over the land surfaces and the associated uncertainties in observations and simulations: a critical review. Atmospheric Chemistry and Physics 19, 15587– 15608, 2019.

Zhu, W., Lin, C.J., Wang, X., Sommar, J., Fu, X., Feng, X. Global observations and modeling of atmosphere-surface exchange of elemental mercury: a critical review. Atmos. Chem. Phys. 16, 4451–4480, 2016.